Ergodic properties of certain iterated function systems arising in partially hyperbolic dynamics

Klaudiusz Czudek

PhD dissertation prepared under the supervision of
prof. dr hab. Tomasz Szarek

Institute of Mathematics Polish Academy of Sciences

Abstract

This PhD dissertation is devoted to the study of ergodic properties of Markov processes corresponding to systems of interval increasing homeomorphisms with probabilities. These systems in general arise in connection with fractals and partially hyperbolic dynamical systems. Our objects of interest appeared to be important due to relations to Kan's example of a diffeomorphism possessing two attractors with intermingled basins. This is briefly explained in Chapter 1.

Chapter 2 describes basic facts on the behaviour of the Markov processes under consideration. The dynamics depends on the values of, so called, the average Lyapunov exponents at 0 and 1. It is proved that if both exponents are positive, then there exists a stationary distribution μ with $\mu((0,1)) = 1$. In that case systems appear to be synchronizing, i.e. the distance between corresponding trajectories starting from two arbitrary points tends to 0 almost surely.

In Chapter 3 it is proved that the average distance between two trajectories is diminishing exponentially fast provided the system is consisted of C^2 diffeomorphisms. The proof strongly relies on certain version of the Baxendale theorem proved by Gharaei and Homburg, which says that the volume Lyapunov exponent of the system is negative. The exponential convergence allows us to show the classical probability limit theorems for the stochastic processes under consideration. The method is based on solving the Poisson equation and the Gordin method.

In the general case it is unknown whether the average distance between two trajectories is diminishing exponentially fast. Nevertheless it is possible to exploit certain result of Dominique Malicet from 2014 to estimate the average distance and show the classical limit theorems. Here the method is again based on martingale approximation and uses the Maxwell-Woodroofe criterion. This is the content of Chapter 4.

Chapter 5 is devoted to the study of very specific systems of homeomorphisms with place-dependent probabilities, called Alsedà-Misiurewicz systems. All known methods of proving ergodicity and stability of iterated function systems with place-dependent probabilities rely on the contractivity in average, which for our system is not satisfied. Nevertheless we demonstrate that these properties hold.

Keywords: iterated function systems, Kan's diffeomorphisms, partially hyperbolic dynamics, synchronization, Baxendale theorem, central limit theorem, Poisson equation, g-measures

AMS subject classification: 37A25, 37E05, 60F05, 60G10, 60J05, 60J25, 76N10

Streszczenie

Niniejsza rozprawa doktorska jest poświęcona badaniom ergodycznych własności procesów Markowa stowarzyszonych z układami rosnących homeomorfizmów odcinka z prawdopodobieństwami. Takie systemy w ogólności są powiązane z fraktalami i dynamiką częściowo hiperboliczną. Nasze obiekty zainteresowań stały się ważne z powodu związków z przykładem Kana dyfeomorfizmu z dwoma atraktorami, których baseny atrakcji są wszędzie gęste i mają dodatnią miarę Lebesgue'a. Rozdział 1 wyjaśnia zwięźle powyższe związki.

Rozdział 2 opisuje podstawowe fakty na temat zachowania rozważanych procesów Markowa. Dynamika zależy od wartości tak zwanych średnich wykładników Lyapunowa w 0 i 1. Jest tam dowiedzione, że jeśli obydwa wykładniki są dodatnie, to istnieje rozkład stacjonarny μ, taki że $\mu((0,1)) = 1$. W tym wypadku systemy okazują się być synchronizujące, to znaczy odległość pomiędzy trajektoriami startującymi z dwóch punktów zbiega do zera prawie na pewno.

W Rozdziale 3 dowodzi się, że średnia odległość pomiędzy trajektoriami maleje wykładniczo szybko, o ile system składa się z dyfeomorfizmów klasy C^2. Dowód silnie polega na pewnej wersji twierdzenia Baxendale'a dowiedzionej przez Gharaei i Homburga, które mówi, że średni wykładnik Lyapunova względem miary stacjonarnej jest ujemny. Wykładnicza zbieżność pozwala nam pokazać klasyczne probabilistyczne twierdzenia graniczne. Metoda polega na rozwiązaniu równania Poissona i aproksymacji martyngałem.

W ogólnym przypadku nie wiadomo, czy średnia odległość pomiędzy trajektoriami maleje wykładniczo. Niemniej jednak można wykorzystać pewne wyniki Dominique'a Malicet z 2014 roku do podania oszacowania górnego na średnią odległość i pokazania klasycznych probabilistycznych twierdzeń granicznych. Tutaj metoda polega na aproksymacji martyngałem i wykorzystuje kryterium Maxwell'a-Woodroofe'a. To jest zawartość Rozdziału 4.

Rozdział 5 poświęcony jest studiowaniu pewnego szczególnego systemu homeomorfizmów odcinka z prawdopodobieństwami zależnymi od położenia, zwanymi układami Alsedy-Misiurewicza. Wszystkie znane techniki dowodzenia ergodyczności i stabilności iterowanych układów funkcyjnych z prawdopodobieństwami zależnymi od położenia polega na średnim zwężaniu, które nie zachodzi w rozważanych systemach. Niemniej jednak udało się znaleźć dowody tych własności.

Słowa kluczowe: iterowane układy funkcyjne, dyfeomorfizm Kana, dynamika częściowo hiperboliczna, synchronizacja, twierdzenie Baxendale'a, centralne twierdzenie graniczne, równanie Poissona, g-miary

Klaysfikacja AMS: 37A25, 37E05, 60F05, 60G10, 60J05, 60J25, 76N10

Contents

Preface

This PhD dissertation is the outcome of my research funded from two grants: "Diamond Grant" 0090/DIA/2017/46 (Chapters 4 and 5) and partially "Preludium 18" 2019/35/N/ST1/02363 (Chapter 3). I have been the director of both grants, and I am grateful for funding to, respectively, Polish Ministry of Science and Higher Education and Polish National Science Centre.

Chapter 1 briefly explains the part of the title: "arising in partially hyperbolic dynamics". Chapter 2 presents the results obtained by my predecessors, while my contribution is presented in Chapters 3-5. Chapter 4 is based on two papers (both published in *Israel Journal of Mathematics*), one coauthored with my supervisor Tomasz Szarek, and one with Tomasz Szarek and Hanna Wojewódka-Ściążko. Chapter 5 is based on my paper published in *Nonlinearity*, whereas Chapter 3 contains unpublished results, which I have proven eventually in 2021 (although the first attempt was at the beginning of my PhD studies). Unfortunately, due to the lack of time I could not prepare the proof carefully. This is apparent in the text: the last two chapters are, in my opinion, quite readable in comparison with the (very technical) content of Chapter 3, which is definitely not presented in an optimal way. I insisted to include these results since it makes the dissertation much more complete. I apologize the readers for this inconvenience.

Now I should acknowledge all who influenced me and my research. First of all, I am grateful to all my teachers. I would like to mention particularly Jacek Gulgowski and the whole Division of Real Functions in the Institute of Mathematics, University of Gdańsk, especially Nikodem Mrożek, Adam Kwela and Rafał Filipów: for the invitation to their research and the first taste of scientific work, although having nothing in common with my present interests. I believe that my first papers were crucial in getting funds for my research. Another important teacher during the period of bachelor and master's studies was Michał Stukow. On one hand I attended several of his lectures on important branches of mathematics like Algebra, Galois Theory, Hyperbolic Geometry and Riemannian Surfaces, and on the other hand he affected my attitude towards mathematics, which had even some impact on my choice of the direction of research.

I remember very well the day of my entry exam for PhD studies. I have a plenty of good memories from these four years, for which I would like to thank my friends, the staff of the Dynamical Systems Department in IM PAS (especially to prof. dr hab. Feliks Przytycki), the staff of the Dynamical Systems Department in University of Warsaw, the administrative staff (I need to mention Mrs Anna Poczmańska) and the directory board of IM PAS. I also appreciate the "Wandering Seminar" organized in November 2017 by Michał Rams and Dominik Kwietniak, where the special guest was Victor Kleptsyn. It had a strong impact on my research.

Of course I need to thank my supervisor, Tomasz Szarek, to whom I owe a lot. I am sure it would be impossible to receive the Diamond Grant or be enrolled on PhD studies in IM PAS without his help. He has supported me and advised on each occasion and has been a mentor in every element of mathematical work. I am still learning from him how to be more persevere and diligent. He has always been reminding me that work can never be more important than a family and that my interests cannot focus just on mathematics. I appreciate the freedom he gave me.

I would like to thank my whole family (I mean by this also my wife's family), particularly my parents for their encouragement and care. I have never been said that my career choice is wrong. Their financial support during bachelor and master's studies make me possible to devote fully to science. My grandparents have always been very proud of me. Unfortunately, they passed away in 2019.

During PhD studies my children were born, Paweł and Mateusz. I have learned from them that I should be curious about the world and enjoy my work. Whatever happened I have always been in good mood when they were next to me.

Finally, the most important person I would like to thank is my wife, Karolina: for her understanding of my ambitions, patience during the periods of more intensive work, for her care and encouragement. After the Covid-19 outbreak she did a lot to enable me working from home (it is not so easy with a small child). Scientist's job, although makes me happy, has some disadvantages affecting private life. She is easygoing with all of them and accompanying me after every failure. This dissertation would not be the same without her support.

Chapter 1

Introduction

1.1 Iterated function systems

The main object of my research are iterated function systems. Let X be a Polish space, and let $f_1, \ldots, f_m$ be continuous transformations from X to X. The m-tuple $(f_1, \ldots, f_m)$ is called an iterated function system. If a probability distribution $\big(p_1(x), \ldots, p_m(x)\big)$ is assigned to $(f_1, \ldots, f_m)$, then the system is called an iterated function system with probabilities. Probabilities are called to be place-independent if p_i's are constant functions. Otherwise the probabilities are called place-dependent.

Iterated function systems appeared in mathematics as processes with complete connections [OM35]. The authors considered a process on $\{0, 1\}$ in which the position in the next step depends on the whole past. This kind of process may be equivalently represented as an iterated function system with place-dependent probabilities on the Cantor set. The work has been continued in [DR37], [ITM48], [ITM50], [DF66]. It should be pointed out that a relation to machine learning has been found [BM53], [Kar53].

The golden era of iterated functions systems started in eighties, when they were exploited to code and generate fractals [Hut81]. Later it was discovered also that iterated function systems with probabilities may generate a fractal image as well (see [DS86] and [DHN85]). This launched more exhaustive research on ergodicity and stability of Markov processes corresponding to iterated function systems with probabilities. The most important papers in this matter are probably [BDEG88] and [LY94]. A comprehensive survey on elaborated methods is [Ste12].

The connections to the ergodic theory of smooth dynamical systems comes through g-measures [Kea72]. There is an important connection of g measures to thermodynamical formalism: a g-measure μ is an equilibrium measure for the potential $\log g$. The explanation of relations between g-measures, equilibrium measures and Gibbs measures is the content of [BFV19].

An interesting survey on iterated function systems is [DF99].

1.2 Smooth dynamical systems

In the modern theory of smooth dynamical systems a huge effort is being made to answer the questions what the behaviour of a typical dynamical system is and how much chaotic it is. A representative example of chaotic dynamical system are hyperbolic diffeomorphisms whose dynamics is well understood.

Definition. *The set Λ is hyperbolic provided there exists its open neighbourhood U of Λ along with a Riemannian metric g on it, and a splitting $T_xM = E^s_x \oplus E^u_x$ for $x \in \Lambda$*

- *which is invariant, thus $D_xf(E^s_x) \subseteq E^s_{f(x)}$ and $D_xf(E^u_x) \subseteq E^u_{f(x)}$ for every $x \in \Lambda$, and*
- *there exist constants $C > 0$, $\lambda \in (0,1)$ such that for any $x \in \Lambda$, $u \in E^s_x$, $v \in E^u_x$ and $n \geq 1$,*
$$\|D_xf^n(u)\| < c\lambda^n \text{ and } \|D_xf^n(v)\| > c^{-1}\lambda^{-n}.$$

The most fundamental theorems about topological and ergodic aspects of hyperbolic dynamics are provided in [KH95].

It may be proven ([AS70]) that hyperbolic diffeomorphisms do not form a dense set in the space of all C^1 diffeomorphisms on a given manifold M. Therefore to understand the behaviour of a typical dynamical system some conditions in the definition of hyperbolicity must be relaxed. To this end partially hyperbolic dynamical systems were introduced.

Definition. *A partially hyperbolic set Λ for diffeomorphisms f is a compact, invariant set for which there exists a continuous splitting of the tangent spaces $T_x\Lambda = E^s_x \oplus E^c_x \oplus E^u_x$, which*

- *is invariant for Df and*
- *there exist constants $c > 0$ and $\lambda \in (0,1)$ such that for every $x \in \Lambda$ and vectors $u \in E^s_x$, $w \in E^c_x$, $v \in E^u_x$ and for every $n \geq 1$,*
$$\|D_xf^nu\| \leq c\lambda^n\|D_xf^nw\| \text{ and } \|D_xf^nw\| \leq c\lambda^n\|D_xf^nv\|.$$

See [CP15] for an excellent treatment of partially hyperbolic dynamics.

Partially hyperbolic dynamical systems are mentioned here since iterated function systems may serve as a model of partial hyperbolicity (a good example here is the porcupine-like horseshoe [DG12]). Then the space on which an iterated function system is defined corresponds to the central direction in the definition of partial hyperbolicity. This is exactly the reason why I was studying iterated function systems as explained in the next section.

Let $f : M \to M$ be a dynamical system, and let φ be a smooth observable. Let us pick a point $x \in M$ randomly according to some distribution. Then $(\varphi(x), \varphi(f(x)), \ldots)$ becomes a stochastic process. It is a feature of chaotic dynamical systems that this process satisfies classical probability limit theorems for some "good" choice of the distribution.

This way of looking at dynamical systems has been started by Sinai and developed by many mathematicians. Nowadays this is a central branch of smooth ergodic theory. See [Den89], [Liv96], [Dol08], [DSL15], [Gou15] for good surveys.

1.3 Kan's example

In 1963 Lorenz proposed a simplified model of atmospheric convection ([Lor63]). He observed that a small change of initial condition may cause a considerable change in qualitative behaviour of solutions. He found also numerically an attractor with fractal structure bearing now his name. Later this class of attractors was called "strange attractors" [RT71]. The presence of a strange attractor may be considered as a kind of chaos as well.

In the history of dynamical systems some other definitions of attractors appeared. For example, in [Kan94] an attractor was defined as a closed invariant subset containing the ω limit set of the set of points of positive Lebesgue measure. The set points with this property is called the basin of attraction. Kan has found an example of diffeomorphism for which there exist two attractors with basins which are intermingled, i.e. every open subset of the space contains points from both basins.

Theorem 1.1 ([Kan94]). *For $k \geq 1$ there exists an open set of C^k diffeomorphisms of $\mathbb{T}^2 \times \mathbb{I}$ (in C^k topology) for which there are two coexisting attractors whose basins are intermingled, and the union of both basins has full Lebesgue measure.*

The set in the statement is some C^k neighbourhood of

$$f(x, y, z) = (3x + y, 2x + y, z + \cos(2\pi x)\frac{z}{32}(1 - z)), \tag{1.1}$$

where we used the correspondence $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. This is a partially hyperbolic diffeomorphism with the central direction being $\mathbb{I}$. The levels $[0, 1] \times \{0\}$ and $[0, 1] \times \{1\}$ are attractors. Kan only announces the result postponing the proof to another paper, and he gives the idea of the proof in the case of noninvertible map (and thus much simpler to deal with) given by $f(x, z) = (3x, z + \cos(2\pi x)\frac{z}{32}(1 - z))$ on $\mathbb{S}^1 \times \mathbb{I}$. However, to my best knowledge Kan has never published the proof.

The proof of the existence of two intermingled basins for (1.1) itself was given in [BM08]. In fact more general theorem was proven. Recall that the Schwarzian derivative of C^3 transformation is given by the formula

$$Sf(y) = \frac{f'''(y)}{f'(y)} - \frac{3}{2}\left(\frac{f''(y)}{f'(y)}\right)^2.$$

Theorem 1.2 ([BM08]). *Let $T : X \to X$ be a continuous transformation of a compact metric space X with an ergodic invariant Borel measure μ. Let f be a skew product of a form $f : X \times [0, 1] \to X \times [0, 1]$, $f(x, z) = (Tx, f_x(z))$. If f is C^3, $Sf_x < 0$ $\mu \times$ Leb almost everywhere, and the levels $[0, 1] \times \{0\}$ and $[0, 1] \times \{1\}$ are attractors (hence if their basins have positive $\mu \times$ Leb measure), then there exists a measurable function $\sigma : X \to [0, 1]$ such that (x, z) us in the basin of attraction of $[0, 1] \times \{0\}$ if $z < \sigma(x)$ and in the basin of attraction of $[0, 1] \times \{1\}$ if $z > \sigma(x)$.*

The authors define also quantities

$$\Lambda_0 = \int_X \log f_x'(0) d\mu \text{ and } \Lambda_1 = \int_X \log f_x'(1) d\mu,$$

which are the average Lyapunov exponents at 0 and 1. It is proved that the negative Lyapunov exponent implies that the corresponding level is an attractor.

For an invertible f we can inverse the dynamics. Then the graph σ in the statement appears to attract the trajectory of $\mu \times$ Leb almost every point. The graph carries the measure ν being the pullback of μ by the projection to the base X. This measure is a physical measure if X is a manifold and μ is the volume.

If the transformation in the base is $T : \mathbb{S}^1 \to \mathbb{S}^1$, $T(x) = kx$, then the system is noninvertible. However it is still possible to show the existence of SRB measure by representing this dynamical systems as a projection of a suitable invertible system defined on solenoid. We can apply the reasoning to obtain the measure $\widetilde{\nu}$ which is SRB for the extended system, and then project it back to $\mathbb{S}^1 \times \mathbb{I}$ to obtain SRB measure ν for f.

In both papers continuity played an important role. This assumption has been dropped in [AM14], where the case of positive average Lyapunov exponents has been treated.

Theorem 1.3. *Let $F : X \times [0, 1] \to X \times [0, 1]$ be an invertible skew-product $F(x, y) = (Tx, f_x(y))$, where T possesses an ergodic stationary measure μ. If*

(I) $\Lambda_0, \Lambda_1 > 0$,

(II) $\{f_x : x \in X\}$ *is finite, or f_x are C^2 diffeomorphisms with $f_x''/((f_x')^2$ bounded uniformly in x and y,*

(III) F is essentially contracting, i. e. $|F(x,y_1) - F(x,y_2)| \to 0$ for every y_1, $y_2 \in [0,1]$ and μ almost every $x \in X$.

Then there exists a measurable function $\sigma : X \to [0,1]$ whose graph is F invariant and such that

1. *$|F^n(x,y) - F^n(x,\sigma(x))| \to 0$ as n goes to infinity,*
2. *$|F^{-n}(x,y) - (T^{-n}(x),0)| \to 0$ as n goes to infinity if $y < \sigma(x)$,*
3. *$|F^{-n}(x,y) - (T^{-n}(x),1)| \to 0$ as n goes to infinity if $y > \sigma(x)$.*

Although two first assumptions appears to be easy to check, the third one is rather mysterious. In the second part of the paper an example of a nontrivial system satisfying the hypothesis is provided. In the base there is two-sided Bernoulli shift ($X = \{0,1\}^{\mathbb{Z}}$, $T\big((x_i)_{i\in\mathbb{Z}}\big) = (x_{i+1})_{i\in\mathbb{Z}}$ μ is the product measure corresponding to the probability vector $(1/2, 1/2)$). In the fiber there are two transformations depending only on x_0. Namely we fix parameter $c \in (0, 1/2)$ and define $f_0(y) = \frac{1}{2c}y$ if $y \le c$ and $f_0(y) = \frac{1}{2(1-c)}(y-1)+1$ if $y > c$. The second transformation f_1 is defined by $f_1(y) = 1 - f_0(1-y)$ for $y \in [0,1]$ (see Figure 1.1). Then f_x if defined by f_{x_0} thus there are finitely many transformations in the fiber.[1] The first assumption is trivially satisfied but the proof of the third one is technical and quite complicated.

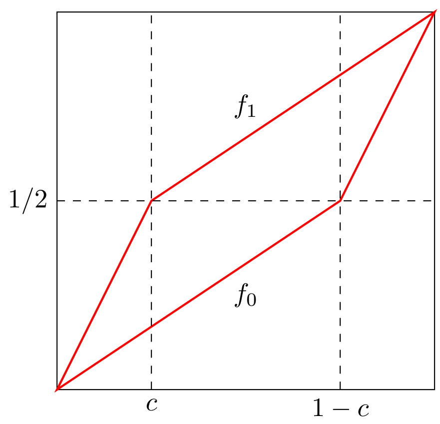

Figure 1.1: The fiber transformations in the Alsedà, Misiurewicz's paper.

This work was later developed by two independent groups of mathematicians: M. Gharaei, A. J. Homburg ([GH17]) and T. Szarek, A. Zdunik ([SZ16]). Both groups were focused on step skew products with Bernoulli shift in the base. In the first paper (which was published first) stronger assumptions are imposed that the transformations in the fiber are C^2 diffeomorphisms, and in the latter one it is only assumed that the transformations are homeomorphisms differentiable at 0 and 1. In both papers it was proven (among other theorems) that if the Lyapunov exponents Λ_0, Λ_1 are positive, then there exists a function σ from the assertion in the theorem from [AM14]. In other words, it has been proven that if in the theorem from [AM14] we restrict to step skew-products over Bernoulli shift then one can drop the assumption of being essentially contractive. What is interesting, in [GH17] (thus under the stronger assumptions) it was proven that a system

[1] A skew product over Bernoulli shift in which the fiber transformation depends only on the first coordinate is called a step skew-product.

satisfying these assumptions must necessarily be essentially contractive. Actually [Mal17] combined with [SZ16] implies the same for the systems of homeomorphisms.

Kan's work attracted much more attention. In [MW05] the authors prove that for every positive integer k or even $k = \infty$ there exists a diffeomorphism of $\mathbb{T}^2 \times \mathbb{S}^2$ with exactly k intermingled basins. In [IKS08] and [Ily08] the authors give examples of open subset of C^1 diffeomorphisms with C^1 topology where all diffeomorphisms have two intermingled basins (in fact the latter paper contains the first published proof of Kan's result).

1.4 The content of the dissertation

The goal of the dissertation is to launch the study of ergodic properties of Kan's type transformations with the emphasis on limit theorems. Our research was restricted to the case of step skew products over Bernoulli shift (thus to the setting of [GH17] and [SZ16]). This case reduces to the study of the corresponding Markov processes, and to use some standard techniques from probability theory.

Chapter 2 contains basic info about this kind of processes. This is a collection of results from previous papers, mainly [GH17] and [SZ16]. Although the form of presentation is rather new, no result is due to myself in this chapter.

The main theorem of mine is that the Markov processes under consideration have exponential decay of correlations assuming that diffeomorphisms are of class C^2. The whole Chapter 3 is devoted to the proof and consequences of this result and is due to me, however one should take into account that some ideas are borrowed from previous paper [CS20b] without explicit mentioning.

Chapter 4 is devoted to the analysis of systems of homeomorphisms. It is the result of work of mine and Tomasz Szarek [CS20b] and mine, Hanna Wojewódka-Ściążko and Tomasz Szarek [CWSS20].

In Chapter 5 we study the case when transition probabilities are place-dependent. The situation arise when some g-measures are considered on Kan's diffeomorphisms. The content is a part of my paper [Czu20].

The results of Chapter 3 are based on the result that volume Lyapunov exponents are negative (Lemma 4.1 in [GH17]). Since in [GH17] only a sketch of proof is given, an appendix is included here in which the details are filled.

Chapter 2

Basic facts about systems of homeomorphisms with place-independent probabilities

2.1 The definition of the process and the average Lyapunov exponents

Let $f_1, \ldots, f_m$ be increasing homeomorphisms of the interval $(0,1)$, and let $p_1, \ldots, p_m$ be positive numbers summing up to 1 (in the sequel every m-tuple of numbers with this property will be referred to as a probability vector). Let us fix also a Borel probability measure μ with $\mu\big((0,1)\big) = 1$.

The present chapter is intended to study stochastic processes defined as follows. In the first step draw a point X_0 with respect to the probability distribution μ. After that, pick one of the homeomorphisms randomly according to the assigned probability vector and independently of the choice of X_0, and move to the point $X_1 := f_i(X_0)$, where f_i is the outcome of the drawing. Then repeat the latter step: pick homeomorphisms with respect to the assigned probability vector and independently of the previous drawings. If the result at the step n is f_i, move to the point $X_{n+1} := f_i(X_n)$.

It is evident that the designed process is a Markov process with the initial distribution μ. Its transition probabilities, which may be easily found, are of the form

$$\sum_{i=1}^{m} p_i \delta_{f_i(x)}. \tag{$*$}$$

Our purpose is to give a general picture of how the behaviour of this random walk depends on the transformations f_i. The theorems presented in this chapter are the result of work of several mathematicians: Alsedà, Misiurewicz, Volk, Kleptsyn, Gharaei, Homburg, Szarek and Zdunik. In Section 2.8 more exact description is provided.

First of all, let us observe that it may happen that transformations are chosen in such a way that the investigation of the behaviour of the process is actually reduced to the study of one or more simpler systems as explained below.

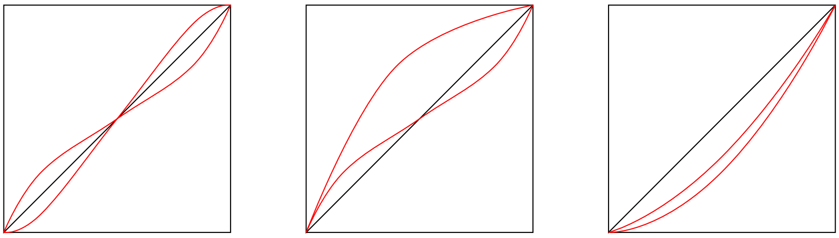

Figure 2.1

Example 1. Let us consider two interval homeomorphisms each with exactly one fixed point at 1/2 (see the box on the left on Figure 2.1). In that case there are two invariant subintervals, and the analysis of the random walk is reduced to the analysis of the system on each of the invariant subintervals.

Example 2. Let f_1 has exactly one fixed point at 1/2 and satisfies $f_1(x) < x$ for $x > 1/2$ and $f_1(x) > x$ for $x < 1/2$. Let f_2 satisfies $f_2(x) > x$ for $x \in (0,1)$ (see the middle box in Figure 2.1). It is easy to see that whatever the starting point is, the sequence eventually gets to $[1/2, 1)$ and stays there forever. Therefore, similarly to the previous example, investigation of the random walk on the whole interval $(0,1)$ is reduced to the investigation of the random walk on $(1/2, 1)$. Note the latter system is not a system of homeomorphisms anymore.

Example 3. This time we request both homeomorphisms to be under diagonal (see the right box on Figure 2.1). It is easy to show that $X_n \to 0$ almost surely.

To exclude situations like above we assume in the sequel that

$$\text{for every } x \in (0,1) \text{ there exist } i, j \text{ with } f_i(x) < x < f_j(x). \tag{A1}$$

It will turn out that the behaviour of the process depends strongly on the properties of the transformations f_i close to zero and one. To formulate a suitable condition we introduce the second assumption that[1]

$$\text{every } f_i \text{ is differentiable at 0 and 1, and all derivatives are nonzero.} \tag{A2}$$

This assumption allows us to define quantities called the average Lyapunov exponents

$$\Lambda_0 := \sum_{i=1}^{m} p_i \log f_i'(0) \quad \text{and} \quad \Lambda_1 := \sum_{i=1}^{m} p_i \log f_i'(1),$$

which for short we shall refer to as Lyapunov exponents at 0 or 1, respectively. Observe that $\log f_i'(0) < 0$ when 0 is an attractive fixed point of f_i and $\log f_i'(0) > 0$ when repelling. Therefore $\Lambda_0 < 0$ means intuitively that 0 is attracting in average, whereas $\Lambda_0 > 0$ means 0 is repelling in average. Keeping that in mind, the statements of the main theorems in the present section should not be surprising. The analysis will be preceded with two propositions concerning the behaviour of the process in the neighbourhood of 0 when $\Lambda_0 < 0$ and when $\Lambda_0 > 0$.

[1] We state it by abuse of notation. The homeomorphisms transform $(0,1)$ onto itself and formally the derivative at 0 or 1 cannot be defined. However, it is clear that any homeomorphisms of $(0,1)$ may be uniquely extended to a homeomorphism of $[0,1]$ and this correspondence is one to one. This gives the derivatives at 0 and 1 the strict meaning.

2.2 The behaviour in the neighbourhood of 0 when $\Lambda_0 < 0$

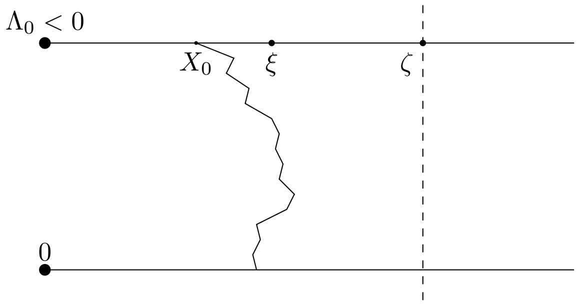

Figure 2.2: If (X_n) starts from $(0, \xi)$, then at least half of its total mass never escapes $(0, \zeta]$.

Proposition 2.1. *Let $\Lambda_0 < 0$. Then for every $\zeta > 0$ there exists $0 < \xi < \zeta$ such that a Markov process (X_n) defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with transition probabilities $(*)$ and $X_0 \le \xi$ a.s. satisfies*

$$\mathbb{P}\left(\bigcap_{n\ge 0}\{X_n \le \zeta\}\right) \ge 1/2.$$

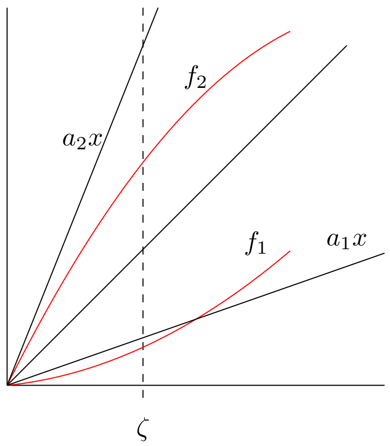

Figure 2.3

Proof. Without loss of generality we may assume ζ to be as close to zero as we wish. In particular we may assume that one can find (by $\Lambda_0 < 0$ and the definition of derivative) positive numbers $a_1, \dots, a_m$ with (see Figure 2.3)

1. $\sum_{i=1}^m p_i \log a_i < 0$ and
2. $f_i(x) \le a_i x$ for $i = 1, \dots, m$ and $x \le \zeta$.

Consider a function of α defined by $\alpha \to a^\alpha$, where a is certain fixed positive number. The application of the Taylor formula (at 0) to this function yields

$$a^\alpha = 1 + \alpha \log a + o(\alpha).$$

Therefore by $\sum_{i=1}^m p_i \log a_i < 0$ there exists α such that

$$\sum_{i=1}^m p_i a_i^\alpha =: c < 1. \tag{2.1}$$

Pick k_0 so large that $c^{k_0} =: p < 1/4$, and take positive ξ with $\xi < \zeta$ which is so close to zero that the transition from $(0, \xi]$ to $[\zeta, 1)$ is impossible in less than $k_0 + 1$ steps. Let (X_n) be a Markov process with transition probabilities $(*)$ and $X_0 \le \xi$ a.s., and denote $Y_n := X_{nk_0}$.

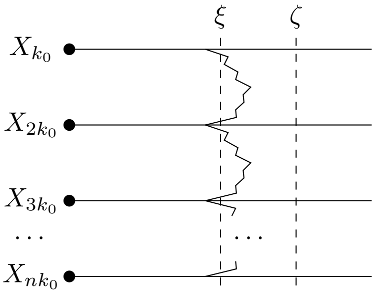

Figure 2.4: The set C_n.

By the choice of k_0 it suffices to prove

$$\mathbb{P}\Big(\bigcap_{n\geq 0} \{Y_n \leq \xi\} \Big) \geq 1/2.$$

To this end observe that for every n and for almost every ω one can find a sequence of numbers $i_1, \dots, i_{k_0} \in \{1, \dots, m\}$ such that $Y_n(\omega) = f_{i_{k_0}} \circ \dots \circ f_{i_1}(Y_{n-1}(\omega))$. In this way we can define a random variable $A_n(\omega) := a_{i_{k_0}} \cdots a_{i_1}$. Notice the choice of A_n is not necessarily unique[2], but surely it can be defined to be measurable with respect to the σ-field $\sigma(Y_1, \dots, Y_n)$.

By (2.1) we obtain

$$\mathbb{E}(A_n^\alpha | Y_{n-1}) = \sum_{i_1,\dots,i_{k_0}} p_{i_{k_0}} \cdots p_{i_1} a_{i_{k_0}}^\alpha \cdots a_{i_1}^\alpha = \sum_{i=1}^m p_i a_i^\alpha \sum_{i_1,\dots,i_{k_0-1}} p_{i_{k_0-1}} \cdots p_{i_1} a_{i_{k_0-1}}^\alpha \cdots a_{i_1}^\alpha$$

$$= c \sum_{i_1,\dots,i_{k_0-1}} p_{i_{k_0-1}} \cdots p_{i_1} a_{i_{k_0-1}}^\alpha \cdots a_{i_1}^\alpha.$$

Proceeding in this manner gives

$$\mathbb{E}(A_n^\alpha | Y_{n-1}) = c^{k_0} = p < 1/4.$$

Now, define $B_n := \{Y_n \leq \xi\}$ and $C_n := B_1 \cap \dots \cap B_n$ (see Figure 2.4). By the definition of the numbers a_i and ζ we have $\{Y_{n-1} < \zeta\} \subseteq \{Y_n < A_n Y_{n-1}\}$. Hence

$$\{Y_n > \xi\} \cap C_{n-1} \subseteq \{A_n Y_{n-1} > \xi\} \cap C_{n-1}$$

for every n, which implies

$$\mathbb{P}\big(\{Y_n > \xi\} \cap C_{n-1}\big) \leq \mathbb{P}\big(\{A_n Y_{n-1} > \xi\} \cap C_{n-1}\big) \leq \mathbb{P}\big(\{A_n A_{n-1} Y_{n-2} > \xi\} \cap C_{n-1}\big)$$

$$\leq \dots \leq \mathbb{P}\big(\{A_n \cdots A_1 Y_0 > \xi\} \cap C_{n-1}\big) \leq \mathbb{P}\big(\{A_n \cdots A_1 \xi > \xi\} \cap C_{n-1}\big)$$

$$= \mathbb{P}\big(\{A_n \cdots A_1 > 1\} \cap C_{n-1}\big) \leq \int_\Omega (A_n \cdots A_1)^\alpha \mathrm{d}\mathbb{P},$$

where the last step is the Chebyshev inequality.

To proceed recall that $\mathbb{E}(A_n^\alpha | Y_{n-1}) = p$. Thus

$$\int_\Omega (A_n \cdots A_1)^\alpha \mathrm{d}\mathbb{P} = \int_\Omega \mathbb{E}\Big((A_n \cdots A_1)^\alpha | Y_{n-1}, \dots Y_1 \Big) \mathrm{d}\mathbb{P} = \int_\Omega (A_{n-1} \cdots A_1)^\alpha \mathbb{E}\big(A_n^\alpha | Y_{n-1}\big) \mathrm{d}\mathbb{P}$$

$$= p \int_\Omega (A_{n-1} \cdots A_1)^\alpha \mathrm{d}\mathbb{P}.$$

[2]Since it may happen that $a_{i_{k_0}} \cdots a_{i_1} = a_{j_{k_0}} \cdots a_{j_1}$ for two different sequences $(i_1, \dots, i_{k_0})$, $(j_1, \dots, j_{k_0})$

The induction argument combined with the preceding estimate gives $\mathbb{P}(\{Y_n > \xi\} \cap C_{n-1}) < \sum_{n=1}^{\infty} p^n < \sum_{n=1}^{\infty} \frac{1}{4^n}$. Thus

$$\mathbb{P}\Big(\bigcap_{n\geq 0}\{Y_n \leq \xi\}\Big) = 1 - \mathbb{P}\Big(\bigcup_{n\geq 1}\{Y_n > \xi\}\Big)$$

$$= 1 - \sum_{n=1}^{\infty} \mathbb{P}(\{Y_n > \xi\} \cap C_{n-1}) > 1 - \sum_{n=1}^{\infty} 1/4^n = 1/2.$$

□

Remark 1. Obviously, we can assume that $\Lambda_1 < 0$. Then the symmetric version of the statement remains true.

2.3 The behaviour in the neighbourhood of 0 when $\Lambda_0 > 0$

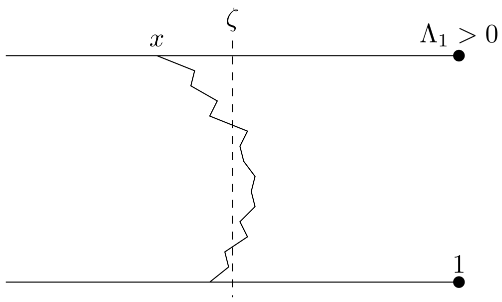

Figure 2.5: Trajectory comes back to $(0, \zeta)$ almost surely.

Proposition 2.2. *If $\Lambda_0 > 0$, then for every $\zeta \in (0,1)$ a Markov chain (X_n) with transition probabilities $(*)$ and starting from some point $x \in (0,1)$ visits the set $(\zeta, 1)$ infinitely many times almost surely.*

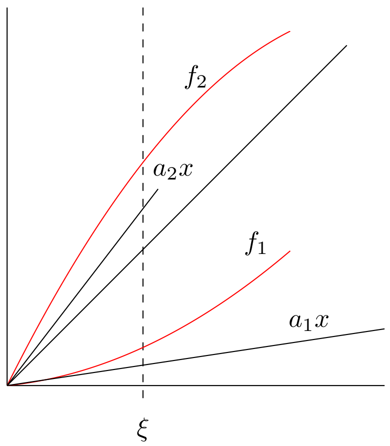

Figure 2.6

Proof. Exactly like in the proof of the preceding proposition we can find (using $\Lambda_1 > 0$ and the definition of derivative) numbers $a_1, \ldots, a_m$ and $\xi > 0$ with

1. $\sum_{i=1}^{m} p_i \log a_i > 0$ and
2. $f_i(x) \geq a_i x$ for $i = 1, \ldots, m$ and $x \leq \xi$.

The first condition may be written in the form $\sum_{i=1}^{m} p_i \log a_i^{-1} < 0$, thus again we can find $\alpha > 0$ with

$$\sum_{i=1}^{m} p_i a_i^{-\alpha} =: c < 1.$$

Fix $x \in (0,1)$. Let (X_n) denote the Markov process with transition probabilities $(*)$ and starting from the point x. For positive integers n, k put

$$C_n^k := \{X_n < \xi\} \cap \cdots \cap \{X_{n+k} < \xi\}.$$

For every n and almost every ω one can find an index $i \in \{1, \ldots, m\}$ such that $X_n(\omega) = f_i(X_{n-1}(\omega))$. In this way we can define a random variable $A_n(\omega) := a_i$. The choice of A_n is not necessarily unique but it may be done to be measurable with respect to the σ-field $\sigma(X_1, \ldots, X_n)$. It is readily seen from the definition of the numbers a_i that

$$\{X_{n-1} < \xi\} \subseteq \{X_n > A_n X_{n-1}\}$$

and consequently

$$\begin{aligned} C_n^k = C_n^{k-1} \cap \{X_{n+k} < \xi\} \subseteq C_n^{k-1} \cap \{\xi > A_{n+k} X_{n+k-1}\} \\ = C_n^{k-2} \cap \{X_{n+k-1} < \xi\} \cap \{\xi > A_{n+k} X_{n+k-1}\} \subseteq C_n^{k-2} \cap \{\xi > A_{n+k} A_{n+k-1} X_{n+k-2}\} \\ \subseteq \cdots \subseteq \{\xi > A_{n+k} \cdots A_{n+1} X_n\}. \end{aligned} \tag{2.2}$$

The proof of the assertion for the particular value $\zeta = \xi$ is completed by showing that $\mathbb{P}(C_n^k) \to 0$ as $k \to \infty$ for any fixed n. This implies the assertion as

$$\mathbb{P}(\liminf\{X_n > 1 - \xi\}) = \mathbb{P}\Big(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} C_n^k\Big) \le \sum_{n=1}^{\infty} \mathbb{P}\Big(\bigcap_{k=n}^{\infty} C_n^k\Big) = 0.$$

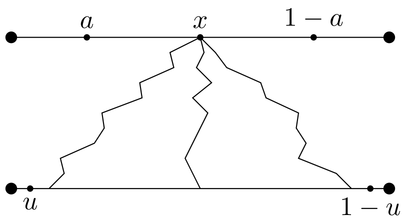

Figure 2.7: X_n is contained in $[u, 1-u]$ a.s.

Fix $n \ge 1$. The random variable X_n is contained in some set $(u, 1-u)$, where $u > 0$, since $X_0 = x$ almost surely (see Figure 2.7). By this fact and (2.2) we have

$$\begin{aligned} \mathbb{P}(C_n^k) \le \mathbb{P}\big(\xi > A_{n+k} \cdots A_{n+1} X_n\big) \le \mathbb{P}\big(\xi > A_{n+k} \cdots A_{n+1} u\big) \\ \le \mathbb{P}\Big((A_{n+k} \cdots A_{n+1})^{-\alpha} > \big(\xi/u\big)^{-\alpha}\Big) \le \big(\xi/u\big)^{\alpha} \mathbb{E}(A_{n+k} \cdots A_{n+1})^{-\alpha}. \end{aligned}$$

The last inequality is the Chebyshev inequality. It remains to estimate the last expression. For $k > 1$ we have

$$\begin{aligned} \int_\Omega A_{n+k}^{-\alpha} \cdots A_{n+1}^{-\alpha} \mathrm{d}\mathbb{P} = \int_\Omega \mathbb{E}\big(A_{n+k}^{-\alpha} | X_{n+k-1}, \cdots, X_n\big) A_{n+k-1}^{-\alpha} \cdots A_{n+1}^{-\alpha} \mathrm{d}\mathbb{P} \\ = \int_\Omega \Big(\sum_{i=1}^{m} p_i a_i^{-\alpha}\Big) A_{n+k-1}^{-\alpha} \cdots A_{n+1}^{-\alpha} \mathrm{d}\mathbb{P} = c \int_\Omega A_{n+k-1}^{-\alpha} \cdots A_{n+1}^{-\alpha} \mathrm{d}\mathbb{P}. \end{aligned}$$

Continuing in this fashion yields

$$\mathbb{P}(C_n^k) \leq \big(\xi/u\big)^{-\alpha} c^k. \tag{2.3}$$

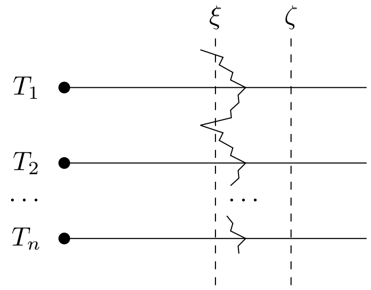

Figure 2.8

It is evident that the statement remains true for any ζ less or equal to ξ. Assume contrary to the claim that there exists $\zeta \in (\xi, 1)$ and $r > 0$ such that $X_n \leq \zeta$ for $n \geq r$ is an event of positive probability. We find, by assumption (A1), some k_0 and some $\beta > 0$ depending only on ξ and ζ with $\mathbb{P}(X_{n+k_0} > \zeta | X_n > \xi) > \beta$. Let (T_n) be a sequence of increasing stopping times greater than r with $X_{T_n} > \xi$ and $T_{n+1} > T_n + k_0$ for every positive integer n. It is possible to find such sequence as the process (X_n) visits $(\xi, 1)$ infinitely many times from the first part of the proof (see Figure 2.8). Again a simple conditioning argument yields

$$\mathbb{P}(X_n \leq \zeta \text{ for all } n \geq r) \leq \mathbb{P}\Big(\bigcap_{n=1}^{k} \{X_{T_n+k_0} \leq \zeta\}\Big) = \mathbb{E}\mathbb{E}\Big(\prod_{n=1}^{k} \mathbb{1}_{\{X_{T_n+k_0} \leq \zeta\}} \Big| \mathcal{F}_{T_k}\Big)$$

$$= \mathbb{E}\Big(\prod_{n=1}^{k-1} \mathbb{1}_{\{X_{T_n+k_0} \leq \zeta\}} \mathbb{E}(\mathbb{1}_{\{X_{T_k+k_0} \leq \zeta\}} | \mathcal{F}_{T_k})\Big) \leq \mathbb{P}\Big(\bigcap_{n=1}^{k-1} \{X_{T_n+k_0} \leq \zeta\}\Big)(1-\beta) \leq (1-\beta)^k$$

for every k, which leads to a contradiction. □

The estimation of the measure of C_n^k holds also when $n = 0$. Therefore the following proposition is a by-product of (2.3).

Proposition 2.3. *If $\Lambda_0 > 0$, then there exists $\alpha \in (0, 1)$ such that for every $\zeta > 0$ sufficiently small and for a Markov process (X_n) with transition probabilities $(*)$ and starting point $x < \zeta$ we have*

$$\mathbb{P}\Big(\bigcap_{k=1}^{n} \{X_k < \zeta\}\Big) \leq \zeta^\alpha / x^\alpha c^n.$$

Proposition 2.3 will be used in Chapters 3 and 4.

Remark 2. As in the previous section, in both propositions we may assume that $\Lambda_1 > 0$. Then the symmetric versions of the statements remain true.

2.4 The behaviour of the random walk when at least one Lyapunov exponent is negative and both are non zero

We are in position to formulate the first of two main theorems of the present chapter.

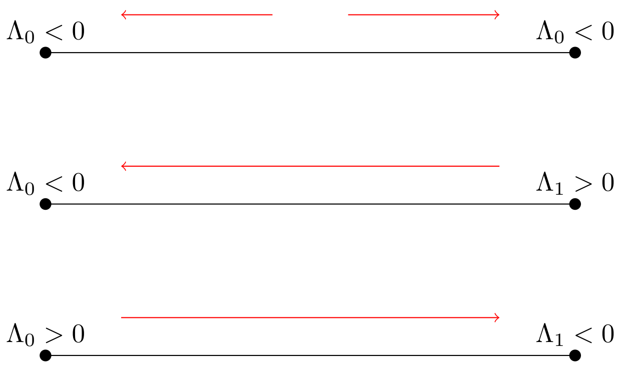

Figure 2.9: The qualitative behaviour under various combinations of values of the Lyapunov exponents

Theorem 2.1. *Let $f_1, \dots, f_m$ be a system of homeomorphisms satisfying assumptions (A1), (A2), and let $(p_1, \dots, p_m)$ be a probability vector. Let (Z_n^x) denote a Markov process with transition probabilities $(*)$ and starting from the point x in $(0,1)$.*

1. *If $\Lambda_0 < 0$ and $\Lambda_1 > 0$, then $Z_n^x \to 0$ a.s. for an arbitrary $x \in (0,1)$.*
2. *If $\Lambda_0 < 0$ and $\Lambda_1 < 0$, then $Z_n^x(\omega) \to 0$ or $Z_n^x(\omega) \to 1$ a.s. for an arbitrary $x \in (0,1)$. If $x \leq y$, then $\mathbb{P}(Z_n^x \to 1) \leq \mathbb{P}(Z_n^y \to 1)$. Moreover,*

$$\lim_{x\to 0} \mathbb{P}(Z_n^x \to 0) = 1 \quad \text{and} \quad \lim_{x \to 1} \mathbb{P}(Z_n^x \to 1) = 1.$$

Proof. The idea of the proof of the first point in the statement is very clear and simple. Fix $\zeta > 0$. Let $\xi \in (0, \zeta)$ stands for the constant in Proposition 2.1, and let $x \in (0,1)$. Since $\Lambda_1 > 0$ the process visits $(0, \xi]$ infinitely many times, by Proposition 2.2. By Proposition 2.1 after every visit at most $1/2$ of the mass escapes from $(0, \zeta]$. This two facts combined yield that the process eventually stays in $(0, \zeta]$ forever almost surely.

In order to explain the details set A_k to be the event that the number of transitions from $(0, \xi]$ to $(\zeta, 1)$ is at least k. We aim to show that $\mathbb{P}(A_{k+1}) \leq 1/2\mathbb{P}(A_k)$ for every positive integer k. Fix k, and denote by T the moment of $(k+1)$th return of (Z_k^x) to $(0, \xi]$, i.e. the kth number n such that $Z_{n-1}^x > \xi$ and $Z_n^x \leq \xi$ (we define the moment of the first visit to be zero in the case when $x \leq \xi$). It is immediate to show that $A_k \in \mathcal{F}_T$, where $\mathcal{F}_T$ is the stopping time σ-algebra. We have

$$\mathbb{P}((\Omega \setminus A_{k+1}) \cap A_k) = \mathbb{E}\mathbb{P}\Big(\bigcap_{n=1}^{\infty} \{Z_{T+n}^x \leq \zeta\} \cap A_k | \mathcal{F}_T \Big)$$

$$= \mathbb{E}\Big(\mathbb{1}_{A_k} \mathbb{P}\Big(\bigcap_{n=1}^{\infty} \{Z_{T+n}^x \leq \zeta\} | \mathcal{F}_T \Big)\Big).$$

The stopping time T is finite a.s. on the set A_k. By the strong Markov property,

$$\mathbb{P}\Big(\bigcap_{n=1}^{\infty}\{Z^x_{T+n} \le \zeta\}|\mathcal{F}_T\Big) = \mathbb{P}\Big(\bigcap_{n=1}^{\infty}\{Z_n^{Z^x_T} \le \zeta\}\Big),$$

almost surely on the set A_k. This random variable is greater than $1/2$ almost surely by Proposition 2.1. Therefore

$$\mathbb{P}((\Omega \setminus A_{k+1}) \cap A_k) > 1/2\mathbb{E}\mathbb{1}_{A_k} = 1/2\mathbb{P}(A_k).$$

Consequently

$$\mathbb{P}(A_{k+1}) = \mathbb{P}(A_{k+1} \cap A_k) = \mathbb{P}(A_k) - \mathbb{P}\big((\Omega \setminus A_{k+1}) \cap A_k)\big) \le 1/2\mathbb{P}(A_k).$$

Therefore $\mathbb{P}(\bigcap_{n=1}^{\infty} A_n) = 0$, and the number of transitions from $(0, \xi]$ to $[\zeta, 1)$ is finite almost surely. It has already been observed that $Z^x_n \le \xi$ for infinitely many n's almost surely by Proposition 2.2, thus $Z^x_n \ge \zeta$ for finitely many n's almost surely. Taking the intersection of such events for $\zeta = 1/k$ completes the proof.

To show the second part let us choose $\zeta > 0$ and apply Proposition 2.1 (and its symmetric version) to get a number $\xi \in (0, \zeta)$ such that

$$\mathbb{P}\Big(\bigcap_{n=1}^{\infty}\{Z^y_n \le \zeta\}\Big) \ge 1/2 \text{ and } \mathbb{P}\Big(\bigcap_{n=1}^{\infty}\{Z^{1-y}_n \ge 1-\zeta\}\Big) \ge 1/2$$

provided $y \le \xi$. Let $x \in (0, 1)$. From the assumption (A1) we conclude that

$$\mathbb{P}(\liminf\{\xi \le Z^x_n \le 1 - \xi\}) = 0,$$

which means that $Z^x_n \in (0, \xi) \cup (1 - \xi, 1)$ infinitely many times a.s. The reasoning from the first part of the proof shows that the number of transitions from $(0, \xi)$ to $(\zeta, 1)$ and from $(1 - \xi, 1)$ to $(0, \zeta)$ is finite almost surely. Thus

$$\mathbb{P}(\liminf\{Z^x_n \le \zeta\} \cup \liminf\{Z^x_n \ge 1 - \zeta\}) = 1.$$

Taking the intersection of the above events for $\zeta = 1/k$ yields the assertion.

We are left to show the properties from the second part of the statement. To prove the first one fix two points $x < y$, and consider a sequence of two dimensional random vectors on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ defined in the following way: pick randomly a transformation f_i with respect to the distribution $(p_1, \dots, p_m)$ and move from $(X_0, Y_0) := (x, y)$ to $(X_1, Y_1) := (f_i(X_0), f_i(Y_0))$. Then repeat the procedure with respect to a random vector (X_1, Y_1). Denote the process by (X_n, Y_n) (Figure 2.10).

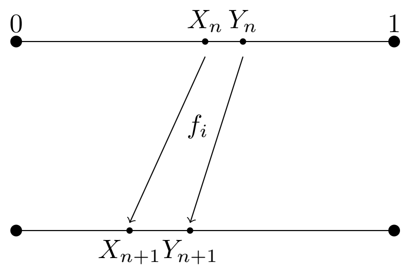

Figure 2.10: At each step one f_i is chosen for both X_n and Y_n.

Notice that the one dimensional distributions of this two dimensional process are exactly the distributions of (Z^x_n) and (Z^y_n). Furthermore, the probability that $X_n < Y_n$ for every positive integer is equal to one since f_i's are increasing. This implies that $\mathbb{P}(Y_n \to 1) \ge \mathbb{P}(X_n \to 1)$. The observation that the assertion does not depend on a probability space but on distribution only completes the proof of the first property.

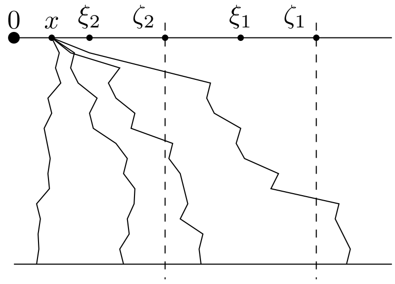

Figure 2.11: At most 1/2 of total mass gets to $(\zeta_2, 1)$, and at most 1/4 of total mass gets to $(\zeta_1, 1)$.

To show the second property fix arbitrary $\zeta_1 > 0$ and pick $\xi_1 \in (0, \zeta_1)$ given by Proposition 2.1. Then take $\zeta_2 < \xi_1$ so close to zero that $f_i(\zeta_2) < \xi_1$ whatever i is (and hence the transition from $(0, \zeta_2]$ to $(\xi_1, 1)$ is impossible in one step). Once again denote by ξ_2 the number given in Proposition 2.1 for $\zeta = \zeta_2$. Let $x < \xi_2$ (see Figure 2.11). By Proposition 2.1 the probability that Z_n^x visits $[\zeta_2, 1)$ at least one time is less that 1/2. Let T be the moment of the first visit in $[\zeta_2, 1)$.

We have

$$\mathbb{P}\Big(\bigcup_{n=1}^{\infty}\{Z_n^x > \zeta_1\}\Big) = \mathbb{E}\mathbb{P}\Big(\bigcup_{n=1}^{\infty}\{Z_{T+n}^x > \zeta_1\} \cap \{T < \infty\} | \mathcal{F}_T\Big)$$
$$= \mathbb{E}\Big(\mathbb{1}_{\{T<\infty\}}\mathbb{P}\Big(\bigcup_{n=1}^{\infty}\{Z_{T+n}^x > \zeta_1\} | \mathcal{F}_T\Big)\Big).$$

By the strong Markov property

$$\mathbb{P}\Big(\bigcup_{n=1}^{\infty}\{Z_{T+n}^x > \zeta_1\} | \mathcal{F}_T\Big) = \mathbb{P}\Big(\bigcup_{n=1}^{\infty}\{Z_n^{Z_T^x} > \zeta_1\}\Big),$$

on $\{T < \infty\}$. The transition from $(0, \zeta_2)$ to $(\xi_1, 1)$ is impossible in one step thus $Z_T^x \le \xi_1$ a.s. Hence by Proposition 2.1 we have

$$\mathbb{P}\Big(\bigcup_{n=1}^{\infty}\{Z_n^{Z_T^x} > \zeta_1\}\Big) < 1/2$$

almost surely on $\{T < \infty\}$. Hence

$$\mathbb{P}\Big(\bigcup_{n=1}^{\infty}\{Z_n^x > \zeta_1\}\Big) < 1/2 \cdot \mathbb{P}(T < \infty) < 1/4,$$

as $\mathbb{P}(T < \infty) < 1/2$ again by Proposition 2.1.

One can proceed with the construction of the sequence $\zeta_1 > \xi_1 > \zeta_2 > \xi_2 > \zeta_3 > \xi_3 > \dots$ in the same way. Then $\mathbb{P}(Z_n^x \to 0) > 1 - 1/2^k$ provided $x \le \xi_k$ a.s. The fact that $\lim_{x\to 1}\mathbb{P}(Z_n^{1-x} \to 1) = 1$ may be proved in the same manner. □

2.5 Stationary processes

In order to explore the dynamics in the case when $\Lambda_0, \Lambda_1 > 0$ we introduce the notion of a stationary process.

Definition. *A stochastic process $(X_n)_n$ (indexed by non-negative integers) defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a measurable space $(S, \mathcal{B})$ is called stationary if for every positive integers k, n and measurable subsets $A_0, \dots, A_k$ of S we have the equality*

$$\mathbb{P}(X_0 \in A_0, \dots, X_k \in A_k) = \mathbb{P}(X_n \in A_0, \dots, X_{n+k} \in A_k).$$

Remark 3. The process in the definition is indexed by non negative integers but it may be assumed to be indexed by integers as well. The only change then is that n is an arbitrary integer.

In particular for an arbitrary stationary process (X_n) the value of $\mathbb{P}(X_n \in A)$ does not depend on n, where A is a fixed measurable subset of S. The significance of stationary processes follows from the Birkhoff ergodic theorem, which we are going to invoke. To this end it is necessary to define the σ-algebra of invariant subsets of S^∞. Denote by θ the shift transformation on S^∞.

Definition. *A measurable*[3] *subset A of S^∞ is said to be invariant if $\theta^{-1}(A) = A$. The set of all invariant subsets forms a σ-algebra (see [Kal02], the beginning of Chapter 9) denoted here by $\mathcal{I}'$.*

Denote by X the transformation $\omega \to (X_1(\omega), X_2(\omega), \ldots) \in S^\infty$. It is quite simple to show that X is a measurable transformation[4], therefore it defines the σ-algebra $\mathcal{I}$ of measurable subsets of Ω which are preimages of the subsets of $\mathcal{I}'$ by X.

Definition. *A process (X_n) is said to be ergodic if all subsets of $\mathcal{I}$ have probability either zero or one.*

We are in position to formulate the Birkhoff ergodic theorem.

Theorem (The Birkhoff ergodic theorem, Theorem 10.6 in [Kal02])**.** *Let (X_n) be a stationary stochastic process, and let ψ be a real valued, measurable function on $\mathcal{S}^\infty$. Then*

$$\frac{1}{n}\big(\psi(\theta^{n-1}X) + \cdots + \psi(X)\big) \to \mathbb{E}(\psi(X)|\mathcal{I}) \quad a.s.$$

In particular ψ may be taken to be a function which takes value one if the first coordinate belongs to some measurable subset A of S and zero otherwise. Then Birkhoff ergodic theorem yields

$$\frac{\#\{i \le n : X_i \in A\}}{n} \to \mathbb{E}(\mathbb{1}_A(X)|\mathcal{I}) \quad \text{a.s.}$$

This gives a description of the statistical behaviour of the stationary process (X_n).

The present chapter is devoted to the study of Markov processes. It is natural to ask whether a Markov process may be a stationary process and when it occurs.

Proposition. *Let (X_n) be a Markov process with values in some measurable space $(S, \mathcal{B})$ with the transition probabilities $p(x, \cdot), x \in S$. Then (X_n) is stationary if and only if the one-dimensional distributions are the same, i.e. the law of X_n is independent of n.*

Proof. It is clear that the one-dimensional distributions of a stationary process are the same. Let $A_0, \ldots, A_k$ be measurable subsets of S. Then

$$\mathbb{P}(X_n \in A_0, \ldots, X_{n+k} \in A_k)$$
$$= \int_{A_0} \mu_n(dx) \int_{A_1} p(x_0, dx_1) \int_{A_2} p(x_1, dx_2) \ldots \int_{A_{k-1}} p(x_{k-1}, A_k),$$

where μ_n denotes the distribution of X_n. It is evident from the formula that if μ_n is independent of n, then the value of $\mathbb{P}(X_n \in A_0, \ldots, X_{n+k} \in A_k)$ is independent of n as well. ☐

[3]Measurable means measurable with respect to the product σ-algebra of the σ-algebra $\mathcal{B}$.
[4]The preimage of a measurable subset of S^∞ is $\mathcal{F}$-measurable.

For a Markov process (X_n) the distribution of $\mathbb{E}(X_{n+1}|X_n = x)$ is $p(x,\cdot)$, $x \in S$, thus $\mathbb{E}(X_{n+1}|X_n = x)$ is independent of time. This implies that all one-dimensional distributions are the same provided the distributions of X_0 and X_1 are the same. It is reasonable then to introduce an operator P on the space of all measures on $(S, \mathcal{B})$ with the property that the distribution of X_2 is $P\mu$ provided the distribution of X_1 is μ.

Definition. *Let $\mathcal{M}$ denote the space of all probability measures on $(S, \mathcal{F})$. The Markov operator $P : \mathcal{M} \to \mathcal{M}$ corresponding to the family of transition probabilities $p(x,\cdot), x \in S$ is defined by the formula*

$$P\mu(A) = \int_S p(x, A)\mu(dx).$$

In particular, if the transition probabilities are given by $(*)$, then

$$P\mu(A) = \sum_{i=1}^{m} p_i \mu(f_i^{-1}(A)). \tag{2.4}$$

When a Markov operator P is given it is convenient also to introduce the dual operator U acting on the space of bounded measurable functions. This operator is sometimes called a transfer operator.

Definition. *The operator U acting on the space $B(S)$ of measurable real functions on $(S, \mathcal{B})$ defined by the formula*

$$U\psi(x) = \int_S \psi(y)p(x, dy), \quad x \in S,$$

is called a dual operator of the Markov operator P corresponding to transition probabilities $p(x,\cdot)$, $x \in S$. In particular if the transition probabilities are given by $(*)$, then it takes the form

$$U\psi(x) = \sum_{i=1}^{m} p_i \psi(f_i(x)).$$

Moreover,

$$U\psi(x) = \mathbb{E}(\psi(X_1) \mid X_0 = x)$$

and

$$U^n\psi(x) = \mathbb{E}(\psi(X_n) \mid X_0 = x)$$

for $n \geq 1$.

The name "dual" is due to the following property.

Proposition. *If $\psi \in B(S)$ and $\mu \in \mathcal{M}$, then*

$$\int_S \psi dP\mu = \int_S U\psi d\mu.$$

Proof. By the Fubini theorem

$$\int_S \psi(y)P\mu(dy) = \int_S \int_S \psi(y)p(x, dy)\mu(dx) = \int_S U\psi(x)\mu(dx).$$

$\square$

It is trivial that U is a bounded operator on $B(S)$ with the supremum norm (moreover, its norm is equal to 1). We finish the section with the following definition.

Definition. *The Markov operator P corresponding to transition probabilities $p(x, \cdot), x \in S$ is a Markov-Feller or Feller operator if its dual operator U preserves the space of continuous functions $C(S)$ on S.*

Since f_i's are homeomorphisms the Markov operators corresponding to the processes with the family of transition probabilities $(*)$ are always Markov-Feller operators.

2.6 The behaviour of the random walk when both Lyapunov exponents are positive

In the case when $\Lambda_0, \Lambda_1 > 0$ one can indicate a measure which turns out to be a stationary distribution. To this end observe that if the system $f_1, \ldots, f_m$ with probability vector $(p_1, \ldots, p_m)$ has positive average Lyapunov exponents at 0 and 1, then the system of inverse functions $(f_1^{-1}, \ldots, f_m^{-1})$ with the same probability vector has negative average Lyapunov exponents at 0 and 1, and Theorem 2.1 holds. Define the function G on the real line by $G(x) = 0$ for $x \leq 0$, $G(x) = 1$ for $x \geq 1$ and

$$G(x) = \mathbb{P}(Y_n^x \to 1) \text{ for } x \in (0, 1),$$

where (Y_n^x) denotes the Markov process corresponding to the system of inverse functions and starting from the point x. From Theorem 2.1 one conclude the following properties of G:

- $0 \leq G(x) \leq 1$,
- $\lim_{x\to 0} G(x) = 0$ and $\lim_{x\to 1} G(x) = 1$,
- G is increasing.

According to the above G is almost a cumulative distribution function (note the lack of right- or left-continuity). Observe that

$$G(x) = \mathbb{P}(Y_n^x \to 1) = \mathbb{E}\mathbb{P}(Y_n^x \to 1 | Y_1^x) = \mathbb{E}G(Y_1^x).$$

Since $Y_1^x = f_i^{-1}(x)$ with probability p_i, $i = 1, \ldots, m$, we obtain

$$G(x) = \mathbb{E}G(Y_1^x) = \sum_{i=1}^{m} p_i G(f_i^{-1}(x)), \tag{2.5}$$

which is exactly the condition (2.4) for $A = (0, x]$. Thus if we manage to show that G is continuous, then the existence of a stationary distribution follows (clearly if two measures are equal on sets of the form $(0, x]$, then the measures are equal).

Theorem 2.2. *Let $f_1, \ldots, f_m$ be a system of homeomorphisms satisfying (A1), (A2). If the probability vector $(p_1, \ldots, p_m)$ is such that $\Lambda_0, \Lambda_1 > 0$, then there exists a unique measure μ such that the Markov process (X_n) with the transition probabilities $(*)$ and with the law of X_0 equal to μ is stationary. The measure μ is atomless. Moreover, a Markov process (Y_n) with transition probabilities $(*)$ and arbitrary initial distribution is stable, which means that the law of Y_n tends to μ in the weak-$*$ topology.*

Remark 4. The existence of a unique stationary distribution is equivalent to the existence of a unique fixed point of P. The stability of the process (Y_n) with the initial distribution ν is equivalent to the convergence of $(P^n\nu)$ to μ in the weak-$*$ topology.

Proof. We start with proving that G is a cumulative distribution function. It will be done by showing its continuity. By (2.5) this implies that μ is stationary and that μ has no atoms.

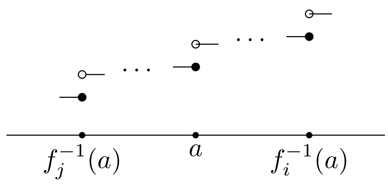

Figure 2.12

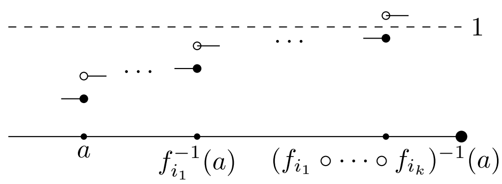

Figure 2.13

Assume contrary that $G(a+) - G(a-) > 0$ for some $a \in (0,1)$. We can assume that a is the number which maximizes the value of $G(x+) - G(x-)$, $x \in (0,1)$ (it exists as G is increasing and bounded). The condition (2.5) leads to the equation

$$G(a+) - G(a-) = \sum_{i=1}^{m} p_i\big(G(f_i^{-1}(a)+) - G(f_i^{-1}(a)-)\big).$$

The right-hand side is a convex combination, and $G(x+)-G(x-)$, $x \in (0,1)$, attains the maximum value at a. Thus it necessarily holds (see Figure 2.12) that

$$G(f_i^{-1}(a)+) - G(f_i^{-1}(a)-) = G(a+) - G(a-)$$

for $i = 1, \ldots, m$. In the same manner we can show that $G((f_j \circ f_i)^{-1}(a)+) - G((f_j \circ f_i)^{-1}(a)-) = G(a+) - G(a-)$ for $i, j = 1, \ldots, m$ and, generally,

$$G((f_{i_1} \circ \cdots \circ f_{i_k})^{-1}(a)+) - G((f_{i_1} \circ \cdots \circ f_{i_k})^{-1}(a)-) = G(a+) - G(a-)$$

for every k and $i_1, \ldots, i_k \in \{1, \ldots m\}$. By assumption (A1) one can find infinitely many finite sequences $(i_1, \ldots, i_k)$ such that the points $\{(f_{i_1} \circ \cdots \circ f_{i_k})^{-1}(a)\}$ are pairwise different (see Figure 2.13), hence it gives infinitely many pairwise different points x with the same positive value of $G(x+) - G(x-)$, which contradicts the fact that G is increasing and $G(1) - G(0) = 1$. It proves the existence of a stationary atomless distribution μ.

Uniqueness of μ will be a consequence of stability. To show the stability, in turn, let us first fix a process (Y_n) with transition probabilities $(*)$ starting from a point $x \in (0,1)$. To show the weak-$*$ convergence of the distribution of Y_n to μ as $n \to \infty$, we first prove that $\mathbb{P}(Y_n \in (a,b)) \to \mu((a,b))$ for any interval $(a,b) \subseteq (0,1)$. This may be deduced from the fact that $\mathbb{P}(Y_n \leq a) \to \mu((0,a))$ and $\mathbb{P}(Y_n < b) \to \mu((0,b))$ as n goes to infinity.

To prove the first statement fix $a \in (0,1)$, and define the Markov process (Z_n) starting from a but corresponding to the system of inverse functions with the same probability vector. Certainly we can assume that (Z_n) and (Y_n) are defined on the same probability space. From the fact that f_i's are increasing we deduce that $\mathbb{P}(Y_n \leq a) = \mathbb{P}(Z_n \geq x)$, which tends to $G(x)$ by the definition of G. This is the desired claim. The proof of $\mathbb{P}(Y_n < b) \to \mu((0,b))$ as $n \to \infty$ is the same. Note the stability of the process starting from a point x is equivalent to the claim that $U^n\psi(x) \to \int \psi d\mu$ pointwise for any $\psi \in C\big((0,1)\big)$. Indeed, $U^n\psi(x) = \int_{(0,1)} U^n\psi(y)\delta_x(dy) = \int_{(0,1)} \psi(y)P^n\delta_x(dy)$.

Finally if (Y_n) has an arbitrary distribution $\nu \in \mathcal{M}\big((0,1)\big)$, then by the Lebesgue convergence theorem

$$\int_{(0,1)} \psi(x)P^n\nu(dx) = \int_{(0,1)} U^n\psi(x)\nu(dx) \to \int_{(0,1)} \psi(x)\mu(dx).$$

This completes the proof. $\square$

2.7 Synchronization

Let us consider a specific probability space on which the processes under consideration are defined (in the sequel we shall use this model frequently). Let $\Omega := \{1,\dots,m\}^{\mathbb{N}}$, let $\mathcal{F}$ be the standard product σ-algebra, and let $\mathbb{P}$ be the product measure of $(p_1,\dots,p_m)$.

Let us assume the system $f_1,\dots,f_m$ with the probability vector $(p_1,\dots,p_m)$ to have positive Lyapunov exponents at 0 and 1. The system of inverse functions $f_1^{-1},\dots,f_m^{-1}$ with the same probability vector has negative Lyapunov exponents at 0 and 1. The processes corresponding to both systems (assumed to start from some point $x \in (0,1)$) may be defined on $(\Omega,\mathcal{F},\mathbb{P})$ by $\omega \to f_{\omega_n}\circ\cdots\circ f_{\omega_1}(x)$ and $\omega \to f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(x)$, $n \ge 0$.

Take S to be an arbitrary dense and countable subset of $(0,1)$. By Theorem 2.1 we can define the subset $\widetilde{\Omega}$ of Ω (of full $\mathbb{P}$ measure) by

$$\widetilde{\Omega} := \bigcap_{x\in S}\Big\{\omega\in\Omega : f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(x)\to 0 \text{ or } f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(x)\to 1\Big\}.$$

Since the intersection of the decreasing sequence of subsets $\{\omega\in\widetilde{\Omega} : f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(1-1/k)\to 0\}$ (indexed by k) has probability 0 (again by Theorem 2.1), we deduce that for every $\omega\in\widetilde{\Omega}$ there exist $x_1,x_2\in S$, $x_1<x_2$, such that $f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(x_1)\to 0$ and $f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(x_2)\to 1$. Let us define a function from $\widetilde{\Omega}$ to $(0,1)$ by the formula

$$x_\omega := \sup\{x\in S : f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(x)\to 0\}.$$

Since S is dense, it may be written equivalently by

$$x_\omega = \inf\{x\in S : f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(x)\to 1\}.$$

Fix $t\in(0,1)$, and observe that $\mathbb{P}(\{\omega\in\Omega : x_\omega\le t\}) = \mathbb{P}(\{\omega\in\Omega : f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(t)\to 1\})$ thus the distribution of the random variable $\omega\to x_\omega$ is μ (cf. Section 2.6).

Take $\varepsilon>0$. By the definition of the function $\omega\to x_\omega$ for every $\xi>0$ we have

$$\mathbb{P}\bigg(\{f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(x_\omega-\varepsilon/2)<\xi\}\cap\{f_{\omega_n}^{-1}\circ\cdots\circ f_{\omega_1}^{-1}(x_\omega+\varepsilon/2)>1-\xi\}\bigg)\to 1$$

as $n\to\infty$. Equivalently

$$\mathbb{P}\bigg(\{x_\omega-\varepsilon/2<f_{\omega_1}\circ\cdots\circ f_{\omega_n}(\xi)\}\cap\{x_\omega+\varepsilon/2>f_{\omega_1}\circ\cdots\circ f_{\omega_n}(1-\xi)\}\bigg)\to 1 \tag{2.6}$$

for every $\xi>0$ and thus

$$\mathbb{P}\big(|f_{\omega_1}\circ\cdots\circ f_{\omega_n}(\xi)-f_{\omega_1}\circ\cdots\circ f_{\omega_n}(1-\xi)|<\varepsilon\big)\to 1$$

as n goes to infinity. Since $\mathbb{P}$ is a product measure we can reverse the order of the sequence $\omega_1,\cdots,\omega_n$ above and rewrite it equivalently

$$\mathbb{P}\big(|f_{\omega_n}\circ\cdots\circ f_{\omega_1}(\xi)-f_{\omega_n}\circ\cdots\circ f_{\omega_1}(1-\xi)|<\varepsilon\big)\to 1$$

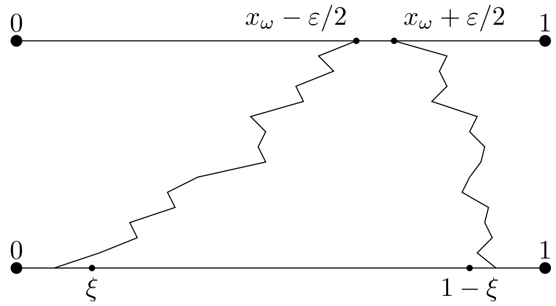

Figure 2.14: For every $\omega \in S$ the dynamics parts interval into to pieces: $(0, x_\omega)$ and $(x_\omega, 1)$.

for every $\xi > 0$ as n goes to infinity.

Hence we have just proven that if a system $f_1, \ldots, f_m$ with $(p_1, \ldots, p_m)$ has positive Lyapunov exponents at 0 and 1, then for every $\xi > 0$ the sequence of random variables

$$(\omega, n) \longmapsto |f_{\omega_n} \circ \cdots \circ f_{\omega_1}(\xi) - f_{\omega_n} \circ \cdots \circ f_{\omega_1}(1-\xi)|$$

converges to zero in probability. We call a system synchronizing if this convergence holds almost surely.

The synchronization of an arbitrary system of C^2 diffeomorphisms with positive Lyapunov exponents has been proven by Gharaei and Homburg [GH17].

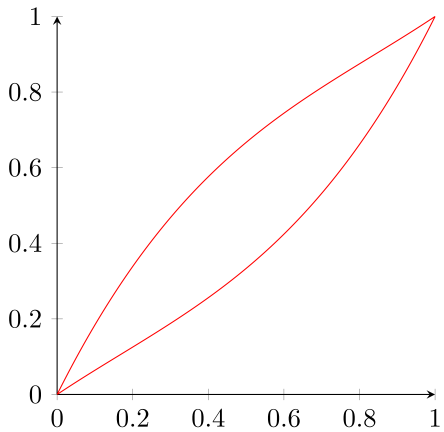

Figure 2.15: The system $f_1(x) = x^2 + \frac{2}{3}x(1-x)^2$ and $f_2(x) = 1 - f_1(1-x)$ with the probability vector $(1/2, 1/2)$.

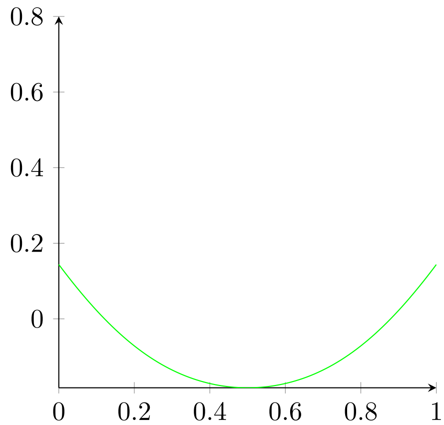

Figure 2.16: The plot of $1/2 \log f_1'(x) + 1/2 \log f_2'(x)$.

Theorem 2.3 (Theorem 4.1 in [GH17])**.** *Let $f_1, \dots, f_m$ be C^2 orientation preserving diffeomorphisms of $[0,1]$ satisfying (A1) and (A2). If $(p_1, \dots, p_m)$ is such that Λ_0, Λ_1 are positive, then*

$$|f_\omega^n(x) - f_\omega^n(y)| \to 0 \quad a.s.$$

for every x, $y \in (0,1)$.

The proof relies on the following theorem (see Figure 2.15 and 2.16)

Theorem 2.4 (Lemma 4.1 in [GH17])**.** *Let $f_1, \dots, f_m$ be C^2 orientation preserving diffeomorphisms of $[0,1]$ satisfying (A1) and (A2). If $(p_1, \dots, p_m)$ is such that Λ_0, Λ_1 are positive, then the volume Lyapunov exponent (with respect to the unique stationary distribution μ)*

$$\sum_{i=1}^{m} p_i \int_{[0,1]} \log f_i'(x) \mu(dx)$$

is negative.

The history of the latter theorem goes back to works of Ledrappier ([Led86]), Baxendale ([Bax89]) and Crauel ([Cra90]). Malicet proved in 2014 an analogue of the Baxendale theorem in non smooth setting [Mal17], which may be used to prove the synchronization for systems of homeomorphisms with positive Lyapunov exponents at 0 and 1.

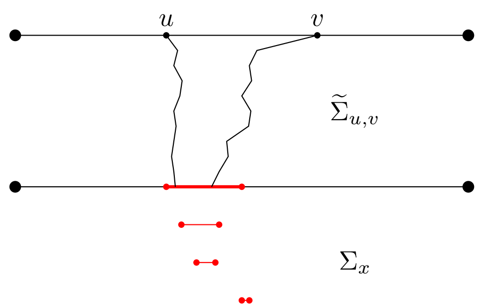

Figure 2.17: The red intervals are the images of I. Their length decreases exponentially fast.

Theorem 2.5 (cf. Corollary 2.13 in [Mal17])**.** *Let $f_1, \dots, f_m$ be interval homeomorphisms. Let $(p_1, \dots, p_m)$ be such that*

- *there exists no nontrivial subinterval of $(0,1)$ which is invariant by all f_i's, and*
- *there exists a measure μ with $\mu((0,1)) = 1$ which is stationary for the random walk,*

then there exist $q < 1$ such that for every $x \in (0,1)$ and for almost every $\omega \in \Omega$ there exits an open neighbourhood I of x such that

$$|f_\omega^n(I)| \leq q^n \text{ for every } n \geq 1. \tag{2.7}$$

Corollary 2.1 (cf. [CWSS20])**.** *Let $f_1, \dots, f_m$ be interval increasing homeomorphisms satisfying (A1) and (A2). If $(p_1, \dots, p_m)$ is such that $\Lambda_0, \Lambda_1 > 0$, then $\sum_{n=1}^{\infty} |f^n(x) - f^n(y)| < \infty$ for every $x, y \in (0,1)$. In particular the system is synchronizing.*

Proof of the Corollary. Recall that the distribution of the random variable $\omega \longmapsto x_\omega$ is μ. Let us take $x \in \mathrm{supp}(\mu)$ and take a neighbourhood I of x such that (2.7) holds for $\omega \in \Sigma_x$, where $\mathbb{P}(\Sigma_x) =: \gamma > 0$.

Set $\beta := \mu(I)\gamma/2$. For any $u, v \in (0,1)$, $u < v$, we may find a set $\Sigma_{u,v} \subset \Omega$ with $\mathbb{P}(\Sigma_{u,v}) \geq \beta$ such that

$$\sum_{n=1}^{\infty} |f_\omega^n(u) - f_\omega^n(v)| < \infty \qquad \text{for } \omega \in \Sigma_{u,v}. \tag{2.8}$$

Indeed, take $u, v \in (0,1)$, $u < v$. By (2.6), the stability (Theorem 2.2), the fact that $\mu(I) > 0$ and $\omega \longmapsto x_\omega$ is μ there exist an integer k_0 and a measurable set $\widetilde{\Sigma}_{u,v}$ of sequences of length k_0 of measure[5] $\mathbb{P}(\widetilde{\Sigma}_{u,v}) > \mu(I)/2$ and such that $f^{k_0}(u), f^{k_0}(v) \in I$. Let $\Sigma_{u,v} = \widetilde{\Sigma}_{u,v} \times \Sigma_x$. Then $\mathbb{P}(\Sigma_{u,v}) \geq \mu(I)\beta/2$ and (2.8) holds.

Fix $x, y \in (0,1)$. Set

$$A := \{\omega \in \Omega : \sum_{n=1}^{\infty} |f_\omega^n(x) - f_\omega^n(y)| < \infty\},$$

and assume, contrary to our claim, that $\mathbb{P}(A) < 1$. Choose a compact subset $A' \subset \Omega \setminus A$ such that $\alpha := \mathbb{P}(A') > 0$. Let $\Sigma_1, \ldots, \Sigma_M$, $M \in \mathbb{N}$, be disjoint cylinders such that $A' \subset \bigcup_{i=1}^M \Sigma_i$ and $\mathbb{P}(\bigcup_{i=1}^M \Sigma_i \setminus A') < \beta\alpha$. Let $\Sigma_i = (\omega_1^i, \ldots, \omega_{n_i}^i) \times \Omega$ for $i \in \{1, \ldots, M\}$. We set $u_i := f_{\omega_{n_i}^i} \circ \cdots \circ f_{\omega_1^i}(x)$ and $v_i := f_{\omega_{n_i}^i} \circ \cdots \circ f_{\omega_1^i}(y)$, and define $\hat{\Sigma}_i = (\omega_1^i, \ldots, \omega_{n_i}^i) \times \Sigma_{u_i, v_i} \subset \Sigma_i$. Obviously,

$$\sum_{n=1}^{\infty} |f_\omega^n(x) - f_\omega^n(y)| < \infty$$

for $\omega \in \hat{\Sigma}_i$. Moreover, $\mathbb{P}(\hat{\Sigma}_i) \geq \beta \mathbb{P}(\Sigma_i)$, and consequently

$$\mathbb{P}(\bigcup_{i=1}^M \hat{\Sigma}_i) \geq \beta \mathbb{P}(\bigcup_{i=1}^M \Sigma_i) \geq \beta \mathbb{P}(A') \geq \beta\alpha.$$

Since $\mathbb{P}(\bigcup_{i=1}^M \hat{\Sigma}_i \setminus A') \leq \mathbb{P}(\bigcup_{i=1}^M \Sigma_i \setminus A') < \beta\alpha$, we finally obtain that $\mathbb{P}(\bigcup_{i=1}^M \hat{\Sigma}_i \cap A') > 0$, which is impossible due to the fact that $\sum_{n=1}^\infty |f_\omega^n(x) - f_\omega^n(y)| < \infty$ for $\omega \in \bigcup_{i=1}^M \hat{\Sigma}_i$. Hence $\mathbb{P}(A) = 1$, and the proof is complete. ☐

All known proofs of synchronization rely on some version of the Baxendale theorem. An elementary proof of this property would be interesting.

2.8 Comments

All theorems and proofs demonstrated so far are the results of work of Alsedà, Misiurewicz, Gharaei, Homburg, Szarek and Zdunik. More specifically, the problems here were initiated in [AM14], which in turn is a generalization of [Kan94] and [BM08]. In [AM14] a very specific symmetric system of two homeomorphisms was considered, and it has been proven that the Lebesgue measure is a stationary distribution of the process and that the system is synchronizing. The last implies the stability and the uniqueness of stationary distribution.

A further research was undertaken in [GH17] and [SZ16] independently and roughly at the same time (prior belongs to [GH17]). Some partial results were obtained by Anna Gordenko in her

[5] By abuse of notation since by the definition $\mathbb{P}$ is a measure on the space of infinite sequences.

master thesis [Gor15] but under very strong assumptions. It is curious that her interest came from other direction (see the remark concerning [DKNP13] below). She was not aware of the existence of [AM14].

The advantage of [SZ16] is that it treats systems consisted of homeomorphisms while in [GH17] the proofs are restricted to systems of C^2 diffeomorphisms. The advantage of [GH17] is it contains a proof that the uniqueness of a stationary distribution implies the synchronization. It also handles some cases where one of the Lyapunov exponents is zero (then the non zero exponent determines the dynamics, see Theorem 5.5 and 6.1 therein; if $\Lambda_0 = 0$ and $\Lambda_1 > 0$, then the process tends to 0 almost surely, if $\Lambda_0 = 0$ and $\Lambda_1 < 0$, then the process tends to 1 almost surely). It is also proved there, that if $\Lambda_0 = \Lambda_1 = 0$, then there is no stationary probability distribution (Theorem 5.4). In other words, $\lim_{n\to\infty} \frac{\#\{i\le n: X_i\in[\xi,1-\xi]\}}{n} = 0$ almost surely for every $\xi > 0$.

When there is no stationary probability distribution, then still some questions about the ergodic behaviour may be asked. Although for any two compact intervals I and J the limits $\lim_{n\to\infty} \frac{\#\{i\le n: X_i\in I\}}{n}$ and $\lim_{n\to\infty} \frac{\#\{i\le n: X_i\in J\}}{n}$ are zero, one can ask about the ratio

$$\lim_{n\to\infty} \frac{\#\{i\le n : X_i\in I\}}{\#\{i\le n : X_i\in J\}}.$$

One can handle this kind of problems using Chacon-Ornstein theorem (see Section 3.8 in [Pet89]), which says that if P (the Markov operator corresponding to the system) has an invariant ergodic[6] measure μ, then the ratio has a limit almost surely being equal to $\mu(I)/\mu(J)$. Note that this implies the Birkhoff ergodic theorem if μ is a probability measure. The existence of ergodic P-invariant measure μ gives the information about the value of the limit for all processes starting from μ almost every point.

In [DKNP13] the existence and uniqueness of infinite invariant measures for symmetric systems was proven. In [BBS20] it was generalized to a vast class of systems.

Recently a complete classification of possible backward and forward behaviour of random systems of homeomorphisms on the real line has been provided by Gordenko [Gor20].

One may ask whether all these results hold when circle homeomorphisms are applied randomly. This appeared to be much more attractive. In eighties Antonov proved that if the system of homeomorphisms is forward and backward minimal (there is no nontrivial closed subset invariant for all homeomorphisms) then exactly one of the following possibilities holds:

- either $|f_{\omega_n}\circ\cdots\circ f_{\omega_1}(x) - f_{\omega_n}\circ\cdots\circ f_{\omega_1}(y)| \to 0$ for $\mathbb{P}$ almost every ω,
- or the system is topologically conjugated to a system of rotations,
- or there exists a circle homeomorphism θ commuting with all homeomorphisms in the circle.

In the last case the identification of the orbits of θ gives another topological space, which is again the circle. The system of homeomorphisms factorizes to a new, synchronizing system. In all cases stationary measure is unique.

Antonov result remained unknown to western mathematicians. It has been rediscovered latter in [DKN07]. Unique ergodicity for systems with infinite number of circle homeomorphisms has been proven recently by Łuczyńska [Łuc21]. The central limit theorem for such systems has been also recently proven by Łuczyńska and Szarek.

[6] Every invariant subset either is of measure zero or its complement is of measure zero.

Chapter 3

Ergodic properties of systems of diffeomorphisms

3.1 The formulation of the main theorem

In Chapter 2 it has been proven that the processes under consideration possess stationary distributions provided Λ_0, $\Lambda_1 > 0$. Usually after establishing basic properties of a stationary process the questions concerning classical limit theorems arise. Among these the central limit theorem, the law of the iterated logarithm and the functional central limit theorem are. Chapter 3 is devoted to this issue.

Take φ to be a measurable real function defined on $[0, 1]$. If (X_n) is a stationary Markov process, then the process of the form

$$\varphi(X_1) + \cdots + \varphi(X_n)$$

is called an additive functional of the process (X_n).

In the classical setting limit theorems hold for partial sums of a sequence of identically distributed independent random variables. Although the process of the form $(\varphi(X_n))$ is stationary, the classical limit theorems are not applicable due to the lack of independence. Nevertheless, Markov processes are memoryless, which implies, intuitively, that X_1 and X_n tend to be independent as n is growing (the process loses the information on the starting position). It should not be surprising that criteria establishing limit theorems rely on proving that the larger $|i - j|$ is, the smaller (in some sense) dependence between $\varphi(X_i)$ and $\varphi(X_j)$ becomes.

To give the strict meaning to this observe that for two independent zero mean random variables Y and Z we have $\mathbb{E}(Z|Y = y) = \mathbb{E}Z = 0$. Thus, given φ with $\int \varphi d\mu = 0$, the $L^2(\mu)$-norm of $\mathbb{E}(\varphi(X_n)|X_1 = x)$ gives a sort of measure of the independence. If X_n and X_0 are close to be independent, then the norm should be small. Recall that $\mathbb{E}(\varphi(X_n)|X_1 = x) = U^n\varphi(x)$ (see Section 2.5). The main result of this chapter is:

Theorem 3.1. *Let $f_1, \ldots, f_m$ be interval orientation preserving C^2 diffeomorphisms with (A1) and (A2). If $(p_1, \ldots, p_m)$ is such that Λ_0, $\Lambda_1 > 0$, φ is a Lipschitz function with $\int \varphi d\mu = 0$, then*

$$\|U^n\varphi\|_{L^2(\mu)} \leq Cq^n$$

for some $q < 1$.

The essential part of the proof of Theorem 3.1 is the following result, whose demonstration constitutes the most part of Chapter 3. In the sequel Theorem 3.1 will be used to deduce the central limit theorem, the law of the iterated logarithm and the functional central limit theorems for additive functional of the Markov processes under consideration.

Theorem 3.2. *If $f_1, \ldots, f_m$ are C^2 diffeomorphisms satisfying (A1) and (A2), $(p_1, \ldots, p_m)$ is such that $\Lambda_0, \Lambda_1 > 0$ and $a \in (0, 1/2)$, then there exist constants $\bar{C}_3 \geq 1$ and $\bar{q}_3 < 1$ such that*

$$\mathbb{E}|Z_n^a - Z_n^{1-a}| \leq \bar{C}_3 \bar{q}_3{}^n$$

for $n \geq 1$.

3.2 Auxillary results

First of all, we shall consider the specific model defined already in Section 2.7. Let $\Omega = \{1, \ldots, m\}^{\mathbb{N}}$, $\mathcal{F}$ denote the standard product σ-algebra on Ω, and let $\mathbb{P}$ be the product measure of the probability vector $(p_1, \ldots, p_m)$. The sequence of random variables $Z_n^x(\omega) := f_{\omega_n} \circ \cdots f_{\omega_1}(x)$ indexed by n, where $x \in (0, 1)$ and $\omega = (\omega_1, \omega_2, \ldots)$, is then a Markov process with transition probabilities $(*)$ and starting from x. Sometimes we shall use also notation more common for skew products: $f_\omega^n(x) := f_{\omega_n} \circ \cdots \circ f_{\omega_1}(x)$ for $\omega = (\omega_1, \omega_2, \ldots)$. Observe that in this particular model we have

$$|U^n\varphi(x) - U^n\varphi(y)| = |\mathbb{E}\big(\varphi(Z_n^x) - \varphi(Z_n^y)\big)| \leq \mathrm{Lip}(\varphi)\mathbb{E}\big|Z_n^x - Z_n^y\big|, \tag{3.1}$$

hence the rate of convergence of $|U^n\varphi(x) - U^n\varphi(y)|$ to zero may be assessed using synchronization (note that $\mathrm{Lip}(\varphi)$ denotes the Lipschitz constant of φ).

In this section we shall use frequently two facts. The first is Proposition 2.3. We have proven there that (since $\Lambda_0 > 0$) there exists $\alpha \in (0, 1)$ and $c < 1$ such that for every $a > 0$ sufficiently small

$$\mathbb{P}\Big(\bigcap_{k=1}^{n}\{Z_k^x < a\}\Big) \leq a^\alpha / x^\alpha c^n \tag{3.2}$$

for $x < a$. Since $\Lambda_1 > 0$, we can also assume α and c to satisfy the analogous property in the neighbourhood of 1, which takes the form

$$\mathbb{P}\Big(\bigcap_{k=1}^{n}\{Z_k^{1-x} > 1 - a\}\Big) \leq a^\alpha / x^\alpha c^n, \tag{3.3}$$

for $x < a$. The number α may be chosen so close to zero to satisfy also another property. Let us define

$$\mathcal{P}_{M,\alpha} = \big\{\mu \in \mathcal{M}\big((0, 1)\big) : \forall_{x \in (0,1)} \mu\big((0, x]\big) \leq Mx^\alpha \text{ and } \mu\big([1 - x, 1)\big) \leq Mx^\alpha\ \big\}.$$

Proposition 3.1 (Gharaei-Homburg, Szarek-Zdunik)**.** *Let $f_1, \ldots, f_m$ be a system of homeomorphisms with (A1) and (A2). If $(p_1, \ldots, p_m)$ is a probability vector such that $\Lambda_0 > 0$ and $\Lambda_1 > 0$, then there exists $\alpha \in (0, 1)$ such that (3.2) and (3.3) are satisfied and, moreover, such that for every $a > 0$ sufficiently small there exists $M \geq 1$ such that the class $\mathcal{P}_{M,\alpha}$ is invariant under the action of the corresponding Markov operator P and every measure supported on $[a, 1 - a]$ belongs to this class.*

Proof. Let us consider the system of inverse functions $f_1^{-1}, \ldots, f_m^{-1}$ with the same probability vector. This system has negative average Lyapunov exponents at 0 and 1. As in the beginning of the proof of Proposition 2.1 and 2.2 we can find numbers $a_1, \ldots, a_m, b_1, \ldots, b_m$ and $\zeta > 0$ with

1. $\sum_{i=1}^m p_i \log a_i < 0$ and $\sum_{i=1}^m p_i \log b_i < 0$
2. $f_i^{-1}(x) \le a_i x$ and $f_i^{-1}(1-x) \ge 1 - b_i x$ for $i = 1, \dots, m$ and $x \le \zeta$.

Using the Taylor formula applied to the function $\alpha \longmapsto a^\alpha$ similarly as in the proof of Proposition 2.1 and 2.2, we find $\alpha > 0$ and $c < 1$ with

$$\sum_{i=1}^m p_i a_i^\alpha < c \text{ and } \sum_{i=1}^m p_i b_i^\alpha < c.$$

Additionally α should satisfy (3.2) and (3.3). Let a be an arbitrary positive number less than ζ. Take M such that $Ma^\alpha = 1$, and take arbitrary $x \in (0,1)$. If $x \ge a$, then $Mx^\alpha \ge Ma^\alpha = 1$, hence $P\mu((0,x]) \le Mx^\alpha$ for every $\mu \in \mathcal{M}$. If $x < a$ and $\mu \in \mathcal{P}_{M,\alpha}$, then

$$\begin{aligned} P\mu((0,x]) = \sum_{i=1}^m p_i \mu((0, f_i^{-1}(x)]) &\le \sum_{i=1}^m p_i \mu((0, a_i x]) \\ &\le \sum_{i=1}^m p_i M a_i^\alpha x^\alpha < Mx^\alpha, \end{aligned}$$

by the choice of α. The analogous computation for $P\mu([1-x,1))$ proves the proposition. □

Corollary 3.1. *Under the assumptions of Proposition 3.1, if μ is the unique stationary distribution, $a \in (0,1/2)$, M and α are the numbers in the assertion, then $\mu \in \mathcal{P}_{M,\alpha}$.*

Proof. First observe that the class $\mathcal{P}_{M,\alpha}$ is weakly-$*$ compact. Indeed, if $\nu_n \in \mathcal{P}_{M,\alpha}$ and $\nu_n \to \nu$ in the weak-$*$ topology, then $\liminf_{n\to\infty} \nu_n\big((0,z)\big) \ge \nu\big((0,z)\big)$ for all $z \in (0,1)$ by the Portmanteau Theorem (Theorem 2.1 in [Bil99]). Hence $\nu\big((0,x]\big) \le \nu\big((0,x+\varepsilon)\big) \le \liminf_{n\to\infty} \nu_n\big((0,x+\varepsilon)\big) \le M(x+\varepsilon)^\alpha$ for every $\varepsilon > 0$, which easily implies the claim. It is immediate to see that $\mathcal{P}_{M,\alpha}$ is convex.

Define $\nu_n := \frac{1}{n}\big(P^{n-1}\delta_{1/2} + \dots + \delta_{1/2}\big)$. Since $\delta_{1/2} \in \mathcal{P}_{M,\alpha}$ and this class is P-invariant and weakly-$*$ compact, the sequence (ν_n) has an accumulation point ν, which must be a stationary distribution by the standard Krylov-Bogoliubov technique: if ψ is an arbitrary continuous function, then

$$\int_{(0,1)} \psi dP\nu_k = \int_{(0,1)} \psi d\nu_k + \frac{1}{n_k}\bigg(\int_{(0,1)} \psi dP^{n_k}\delta_{1/2} - \int_{(0,1)} \psi d\delta_{1/2}\bigg).$$

The modulus of the second summand tends to zero hence $\lim_{k\to\infty} \int \psi dP\nu_k = \lim_{k\to\infty} \int \psi d\nu_k$. On the other hand P is a Feller operator therefore $U\psi$ is continuous. From the definition of the weak-$*$ convergence $\lim_{k\to\infty} \int \psi dP\nu_k = \lim_{k\to\infty} \int U\psi d\nu_k = \int U\psi d\nu = \int \psi dP\nu$. This proves that $P\nu = \nu$. Thus $\mu = \nu$ by the uniqueness of a stationary distribution and $\mu \in \mathcal{P}_{M,\alpha}$. □

The following proposition is crucial in the proof of Theorem 3.2. It relies on a technical lemma, whose proof is postponed to Section 3.3.

Proposition 3.2. *Let $f_1, \dots, f_m$ be increasing homeomorphisms satisfying (A1), (A2). Let $p_1, \dots, p_m$ be such that $\Lambda_0, \Lambda_1 > 0$. If $a > 0$ is sufficiently small and $T^a(x,y)$ is the minimum number k with $a \le Z_k^x < Z_k^y \le 1-a$, $x<y$, then there exist $\gamma > 0$ and $\bar{C}_1 \ge 1$ such that for every $x < y$ we have*

$$\mathbb{E}e^{\gamma T^a(x,y)} \le \bar{C}_1 \max\Big\{\big(a/z\big)^\alpha, 1\Big\},$$

where $z = \min\{x, 1-y\}$.

Proof. Let us take $a > 0$ sufficiently small to satisfy Proposition 3.1. We insist also that the transition from $(0, a)$ to $(1 - a, 1)$ is impossible in one step (this implies that if $x < a$ and $Z_k^x > a$ for some k, then also $Z_{k'}^x \in [a, 1-a]$ for some $k' \leq k$). Eventually let a satisfy the following lemma.

Lemma 3.1. *Let $f_1, \ldots, f_m$ be increasing homeomorphisms satisfying (A1), (A2). Let $p_1, \ldots, p_m$ be such that $\Lambda_0, \Lambda_1 > 0$. Then there exist $\bar{C}_2$ and $\bar{q}_2 < 0$ such that for every $a > 0$ sufficiently small and for every $x \in [a, 1 - a]$ it holds that*

$$\mathbb{P}\left(\frac{\#\{i \leq n : Z_i^x \in [a, 1-a]\}}{n} \leq 3/4 \right) \leq \bar{C}_2 \bar{q}_2{}^n$$

for every n.

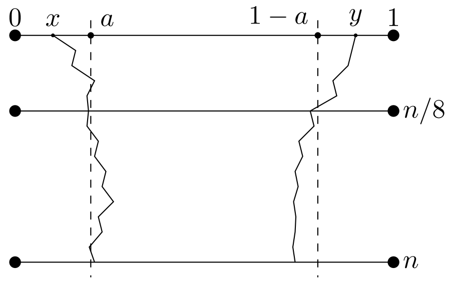

Figure 3.1

Heuristically (cf. Figure 3.1) Proposition 3.1 implies that a vast majority of trajectories visits $[a, 1-a]$ until $n/8$ but we do not know whether it happens at the same moment for x and y. To solve the problem we apply Lemma 3.1, which says that from the moment of the first visit in $[a, 1 - a]$ the trajectory spends at least $3/4$ of time in $[a, 1-a]$ up to the set of measure diminishing exponentially fast.

By (3.2), (3.3) we know that

$$\mathbb{P}\left(\bigcap_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^x < a\} \cup \bigcap_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^y > 1 - a\} \right) \leq 2a^\alpha / z^\alpha c^{\lfloor n/8 \rfloor}$$

for every n, thus we are left with estimating

$$\mathbb{P}\left(T^a(x, y) > n \;\middle|\; \bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^x \geq a\} \cap \bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^y \leq 1 - a\} \right).$$

Since the transition from $(0, a)$ to $(1 - a, 1)$ and from $(1 - a, 1)$ to $(0, a)$ is not possible in one step the above is equal to

$$\mathbb{P}\left(T^a(x, y) > n \;\middle|\; \bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^x \in [a, 1-a]\} \cap \bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^y \in [a, 1-a]\} \right).$$

Fix $k \leq \lfloor n/8 \rfloor$. Clearly there is at least $\lfloor 7n/8 \rfloor$ numbers between k and n. This combined with Lemma 3.1 implies that the conditional probability that $Z_i^x \in [a, 1-a]$ for less than $3/4$ of indices i among $i = k+1, \ldots, \lfloor n/8 \rfloor, \lfloor n/8 \rfloor + 1, \ldots, n$ under the condition that $Z_k^x \in [a, 1-a]$ is less than $\bar{C}_2 \bar{q}_2{}^{\lfloor 7n/8 \rfloor}$. Further, if $Z_i^x \in [a, 1-a]$ for more than $3/4$ of indices i, $i = k+1, \ldots, \lfloor n/8 \rfloor, \lfloor n/8 \rfloor + 1, \ldots, n$, then totally $Z_i^x \in [a, 1-a]$ for at least $\lfloor 7n/8 \rfloor \cdot 3/4$ indices i. At most $\lfloor n/8 \rfloor$ of such i's is less than k, hence $Z_i^x \in [a, 1-a]$ for at least $\lfloor 7n/8 \rfloor \cdot 3/4 - \lfloor n/8 \rfloor > n/2$ of indices $i = k+1, \ldots, n$. The same is true about (Z_n^y). If $Z_i^x \in [a, 1-a]$ for more than $n/2$ indices i between 1 and n

and $Z_i^y \in [a, 1-a]$ for more than $n/2$ indices i between 1 and n then clearly $Z_i^x \in [a, 1-a]$ and $Z_i^y \in [a, 1-a]$ for at least one $i \in [1, n]$. Therefore

$$\mathbb{P}\left(T^a(x,y) > n \;\middle|\; \bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^x \in [a,1-a]\} \cap \bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^y \in [a,1-a]\}\right) \leq 2\bar{C}_2 \bar{q}_2^{\lfloor 7n/8 \rfloor}$$

Further,

$$\mathbb{P}(T^a(x,y) > n) \leq \mathbb{P}\left(\bigcap_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^x < a\} \cup \bigcap_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^y > 1-a\}\right)$$

$$+\mathbb{P}\left(T^a(x,y) > n \middle| \bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^x \geq a\} \cap \bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^y \leq 1-a\}\right) \cdot \mathbb{P}\left(\bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^x \geq a\} \cap \bigcup_{k=0}^{\lfloor n/8 \rfloor} \{Z_k^y \leq 1-a\}\right)$$

$$\leq 2a^\alpha / z^\alpha c^n + 2\bar{C}_2 \bar{q}_2^{\lfloor \frac{7}{8} n \rfloor},$$

Take $\gamma > 0$. Using the preceding estimation yields

$$\mathbb{E}e^{\gamma T^a(x,y)} \leq \sum_{n=0}^{\infty} \mathbb{P}(T^a(x,y) \geq n) e^{\gamma n} = 1 + \sum_{n=1}^{\infty} \mathbb{P}(T^a(x,y) > n-1) e^{\gamma n}$$

$$\leq 1 + \sum_{n=1}^{\infty} \left(2a^\alpha / z^\alpha c^{n-1} + 2\bar{C}_2 \bar{q}_2^{\lfloor \frac{7}{8}(n-1) \rfloor}\right) e^{\gamma n}$$

$$\leq \max\left\{a^\alpha / z^\alpha, 1\right\} \left(1 + \sum_{n=1}^{\infty} \left(2c^{n-1} + 2\bar{C}_2 \bar{q}_2^{\lfloor \frac{7}{8}(n-1) \rfloor}\right) e^{\gamma n}\right).$$

Taking γ sufficiently small makes the series convergent and completes the proof. □

3.3 The proof of Lemma 3.1

For $a \in (0, 1/2)$ and $x \in [a, 1-a]$ let us define s_n to be the moment of the n-th return[1] to $(0, a) \cup (1-a, 1)$ and t_n to be the moment of the n-th return to $[a, 1-a]$. Clearly $s_1 < t_1 < s_2 < t_2 < \dots$ Further, let τ_n be the length of the n-th visit in $[a, 1-a]$, and let σ_n be the length of the n-th visit in $(0, a) \cup (1-a, 1)$ (Figure 3.2). To avoid confusion, the precise definitions are as follows:

$$s_1(\omega, x) = \min\{k \geq 1 : f_\omega^k(x) \in (0, a) \cup (1-a, 1)\},$$

$$t_1(\omega, x) = \min\{k \geq 1 : f_\omega^k(x) \in [a, 1-a]\},$$

$$s_n(\omega, x) = s_{n-1}(\omega, x) + s_1\left(\theta^{s_{n-1}(\omega,x)}\omega, f_\omega^{s_{n-1}(\omega,x)}(x)\right) \text{ for } n \geq 2,$$

$$t_n(\omega, x) = t_{n-1}(\omega, x) + t_1\left(\theta^{t_{n-1}(\omega,x)}\omega, f_\omega^{t_{n-1}(\omega,x)}(x)\right) \text{ for } n \geq 2,$$

$$\tau_1 = s_1 - 1, \quad \tau_n = s_n - t_{n-1}, \quad \sigma_n = t_n - s_n \quad \text{for } n \geq 2,$$

where θ denotes the shift to the left in Ω.

[1]Return means that we insist the existence of j between each pair of consecutive s_n's with $Z_j^x \in [a, 1-a]$. The analogous requirement is made for t_n's.

Define $\rho_n(\omega) = \max\{k : s_k \leq n\}$. We start with the observation that

$$\frac{\#\{i \leq n : Z_i^x \notin [a, 1-a]\}}{n} \leq \frac{\sigma_1 + \cdots + \sigma_{\rho_n}}{\tau_1 + \sigma_1 + \cdots + \tau_{\rho_n} + \sigma_{\rho_n}} = \frac{\#\{i \leq t_{\rho_n} : Z_i^x \notin [a, 1-a]\}}{t_{\rho_n}}. \qquad (3.4)$$

The equality is a direct consequence of the above definitions (cf. Figure 3.2). To show the inequality we observe that $s_{\rho_n} \leq n < s_{\rho_n+1}$, by the definition of ρ_n. If n is greater or equal to t_{ρ_n}, then the trajectory $(Z_i^x)_i$ spends $n - t_{\rho_n}$ more steps in $[a, 1-a]$ up to n in comparison with the same trajectory up to t_{ρ_n} (cf. Figure 3.2). This implies the inequality. In the remaining case $s_{\rho_n} \leq n < t_{\rho_n}$ the trajectory $(Z_i^x)_i$ spends $t_{\rho_n} - n$ more steps outside $[a, 1-a]$ up to t_{ρ_n} in comparison with the same trajectory up to n, which clearly implies (3.4).

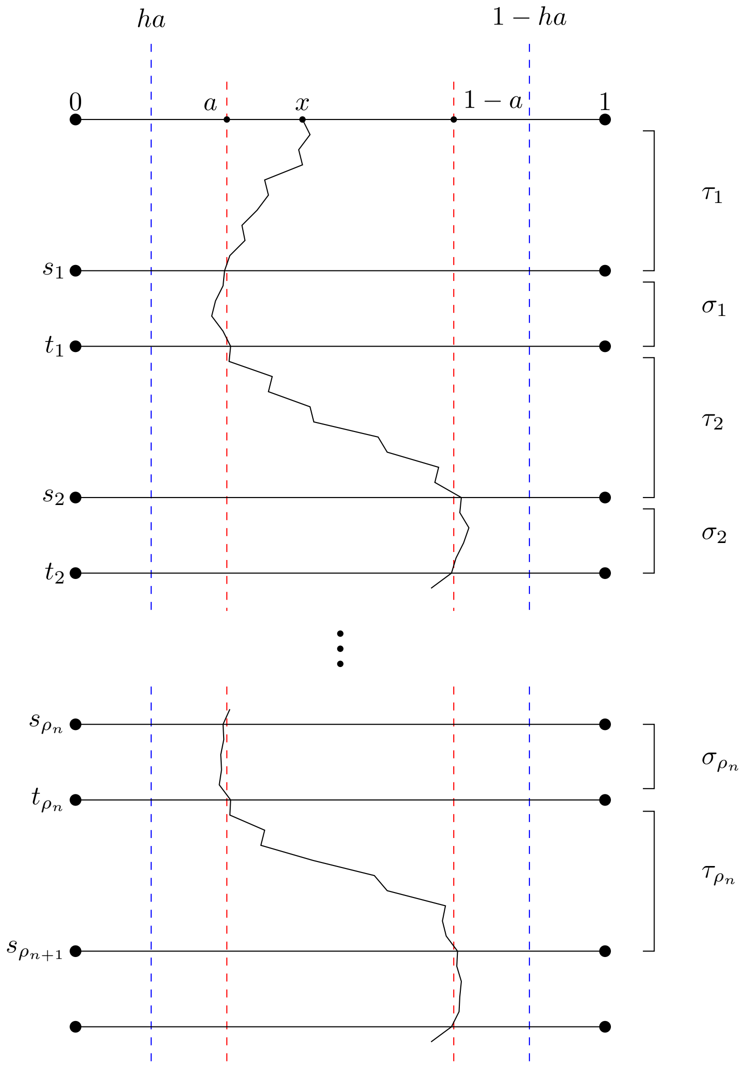

Figure 3.2: n is between s_{ρ_n} and $s_{\rho_{n+1}}$

Put

$$h=\min\left\{\min_{i=1,\dots,m}\frac{f_i'(0)}{2},\min_{i=1,\dots,m}\frac{f_i'(1)}{2}\right\}.$$

Since $Z^x_{s_k-1}\in[a,1-a]$, it cannot happen that $Z^x_{s_k}$ is too close to zero. More precisely, using the definition of the derivative at 0 and 1 we can write that if a is sufficiently small, then (see Figure 3.2)

$$Z^x_{s_k}\geq ha \quad\text{and}\quad Z^x_{s_k}\leq 1-ha$$

for every k. Therefore using Proposition 3.1 we can write

$$\mathbb{P}(\sigma_k>n)\leq\mathbb{P}\bigg(\bigcap_{i=1}^{n}\{Z_i^{ha}<a\}\bigg)\leq\frac{a^\alpha}{(ha)^\alpha}c^n=h^{-\alpha}c^n.$$

Note this inequality is independent of a as long as a is chosen sufficiently small. In view of these remarks, if γ_1 is chosen so that $e^{\gamma_1}c<1$, then by the strong Markov property

$$\mathbb{E}\big(e^{\gamma_1\sigma_k}\mid\mathcal{F}_{s_k}\big)\leq\sum_{n=0}^{\infty}e^{\gamma_1 n}\mathbb{P}(\sigma_k>n)\leq\sum_{n=0}^{\infty}e^{\gamma_1 n}c^n\leq e^{L_1}\quad\text{a.s.}$$

for some L_1 and every k. Therefore

$$\mathbb{E}e^{\gamma_1(\sigma_1+\dots+\sigma_k)}=\mathbb{E}\mathbb{E}\big(e^{\gamma_1(\sigma_1+\dots+\sigma_k)}\mid\mathcal{F}_{s_k}\big)=\mathbb{E}\Big(e^{\gamma_1(\sigma_1+\dots+\sigma_{k-1})}\mathbb{E}\big(e^{\gamma_1\sigma_k}\mid\mathcal{F}_{s_k}\big)\Big)$$

$$\leq e^{L_1}\mathbb{E}e^{\gamma_1(\sigma_1+\dots+\sigma_{k-1})}\leq\dots\leq e^{L_1k} \tag{3.5}$$

and by Chebyshev's inequality

$$\mathbb{P}\bigg(\sigma_1+\dots+\sigma_k>\frac{2kL_1}{\gamma_1}\bigg)\leq e^{-\gamma_1\frac{2kL_1}{\gamma_1}}\mathbb{E}e^{\gamma_1(\sigma_1+\dots+\sigma_k)}\leq e^{-2L_1k+L_1k}=\big(e^{-L_1}\big)^k. \tag{3.6}$$

Put $L_2:=\frac{2L_1}{\gamma_1}$ and $\bar{q_4}:=e^{-L_1}$. With this notation (3.6) may be rewritten as

$$\mathbb{P}(\sigma_1+\dots+\sigma_k>kL_2)\leq\bar{q_4}^k \tag{3.7}$$

for every k. Let us stress once again, this inequality is independent of a as long as it was sufficiently small.

This gives us an upper bound on the time that a trajectory spends outside $[a,1-a]$. One can guess that the next step is to provide a lower bound on the time that a trajectory spends in $[a,1-a]$. This bound will be of the form

$$\mathbb{P}(\tau_1+\dots+\tau_k<L_3k)\leq M_3\bar{q_5}^k,$$

where, which is especially important, the constants L_3, M_3, $\bar{q_5}$ do not depend on a as long as a is sufficiently small, similarly to the previous estimation.

One can see here a sort of large deviation type estimation. However, we cannot just apply the Cramér-Chernoff theorem (Theorem 27.3 in [Kal02]), since (τ_k) is a sequence of neither stationary nor independent random variables. The idea is to define (on the same probability space $(\Omega,\mathcal{F},\mathbb{P})$) a sequence (Y_k) of bounded i.i.d. random variables and satisfying $0\leq Y_k(\omega)\leq\tau_k(\omega)$ for every $\omega\in\Omega$ and every k. What is crucial, the sequence (Y_k) must be independent of a since the constants in the assertion must be independent of a.

To this end, take $b > 0$, $a_1, \ldots, a_m$, $b_1, \ldots, b_m$ such that

- $f_i(x) \geq a_i x$ and $f_i(1-x) \leq 1 - b_i x$ for $x < b$, and
- $\sum_{i=1}^m p_i \log a_i > 0$ and $\sum_{i=1}^m p_i \log b_i > 0$.

Let us consider a random walk (S_n) on $\mathbb{R}$ starting from 0 with i.i.d. steps, which equal $\log a_i$ with probability p_i. Random walks are either recurrent or transient (see Theorem 8.1 in [Kal02]). The strong law of large numbers combined with $\sum p_i \log a_i > 0$ gives that $S_n \to +\infty$ a.s. and the random walk is transient. This implies in particular that $\mathbb{P}(\bigcup_{i=1}^{\infty}\{S_i < 0\}) < 1$ (this again follows by Theorem 8.1 in [Kal02]). In other words, there exists $\eta > 0$ such that for every $A > 0$ the probability that (S_n) enters $[A, +\infty)$ before returning to 0 is greater than η. By this there exist $A > 0$, $r > 0$ such that $\mathbb{E}Y > 40L_2$, where $Y := 0$ if (S_n) visits $(-\infty, 0)$ before the first visit in $[A, +\infty)$, $Y := r \wedge \min\{n \geq 1 : S_n \geq A\}$ otherwise (recall that $L_2 = \frac{2L_1}{\gamma_1}$; the definition is just before (3.7)). Moreover, A and r should also satisfy the analogous property for the random walk defined by p_i and $\log b_i$ with the same constant L_2.

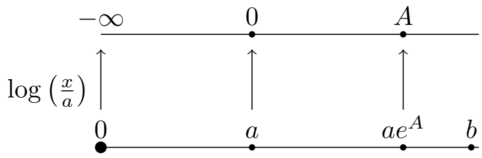

Figure 3.3

Eventually let us fix $a < b$ satisfying so far listed conditions and such that $e^A a < b$ (Figure 3.3). Define $g_i(x) = a_i x$, $i = 1, \ldots, m$. Fix k, put $(i_1, i_2, \ldots) = (\omega_{t_{k-1}}, \omega_{t_{k-1}+1}, \ldots)$. Let us define $Y_k := 0$ if $(g_{i_n} \circ \cdots \circ g_{i_1}(a))_n$ visits $(0, a)$ before the first visit in $(ae^A, 1)$ and $Y_k := r \wedge \min\{n \geq 1 : g_{i_n} \circ \cdots \circ g_{i_1}(a) > ae^A\}$. Observe that the distribution of Y_k is the same as of Y (on Figure 3.3 one can see the correspondence between the random walks). Moreover (Y_k) are independent, bounded by r and satisfy $Y_k \leq \tau_k$ provided $Z^x_{t_{k-1}} < a$. By the Cramér-Chernoff theorem (Theorem 27.3 in [Kal02])

$$\mathbb{P}(Y_1 + \cdots + Y_k < 16kL_2) \leq M_3 \bar{q_5}^k \tag{3.8}$$

for all k's and some $M_3 \geq 1$, $\bar{q_5} \in (0,1)$. Among k returns to $[a, 1-a]$ there are at least $k/4$ returns from $(0, a)$ (denote this event by H_k) or at least $k/4$ returns from $(1-a, 1)$ (denote by G_k). If the first case holds, $Y_{i_1} + \cdots + Y_{i_{\lceil k/4 \rceil}} \leq \tau_1 + \cdots + \tau_k$ for some $i_1 < \ldots < i_{\lceil k/4 \rceil} \leq k$. Hence

$$\mathbb{P}\left(\tau_1 + \cdots + \tau_k < \frac{k}{4} \cdot 16L_2 \,\middle|\, H_k\right) \leq M_3 \bar{q_5}^{k/4}$$

for $k \geq 1$. An analogous computation in the neighbourhood of 1 gives

$$\mathbb{P}\left(\tau_1 + \cdots + \tau_k < \frac{k}{4} \cdot 16L_2 \,\middle|\, G_k\right) \leq M_3 \bar{q_5}^{k/4}$$

for all k as well (after possible amendment of M_3 and $\bar{q_5}$). Both combined yields

$$\mathbb{P}\left(\tau_1 + \cdots + \tau_k < \frac{k}{4} \cdot 16L_2\right) \leq 2M_3 \bar{q_5}^{k/4} \tag{3.9}$$

for every positive integer k.

If $\sigma_1+\cdots+\sigma_{\rho_n} \le \rho_n L_2$ (cf. (3.7)) and $\tau_1+\cdots+\tau_{\rho_n} \ge 16L_2\rho_n/4 = 4L_2\rho_n$ (cf. (3.9)), then by (3.4)

$$\frac{\#\{i \le n : Z_i^x \notin [a,1-a]\}}{n} \le \frac{\sigma_1+\cdots+\sigma_{\rho_n}}{\tau_1+\cdots+\tau_{\rho_n}+\sigma_1+\cdots+\sigma_{\rho_n}} \le L_2\rho_n/(4L_2\rho_n) = 1/4. \tag{3.10}$$

The probability of the remaining part will be estimated using (3.7) and (3.9). Since we are aimed at proving its exponential decay, we need some further information on the growth of ρ_n. Fix $\lambda \in (0,1)$. In a moment we shall need it satisfying

$$L_1\lambda < \gamma_1/4. \tag{3.11}$$

Let us observe that by (3.10), (3.7) and (3.9) the probability of the event

$$\left\{\frac{\#\{i \le n : Z_i^x \in [a,1-a]\}}{n} < 3/4\right\} \cap \left\{\rho_n > \lfloor \lambda n \rfloor\right\}$$

diminishes exponentially fast, and we are left with estimating the probability of

$$\left\{\frac{\#\{i \le n : Z_i^x \in [a,1-a]\}}{n} < 3/4\right\} \cap \left\{\rho_n \le \lfloor \lambda n \rfloor\right\}.$$

To this end we observe that

$$\{\sigma_1+\cdots+\sigma_{\rho_n} > n/4\} \cap \{\rho_n \le \lfloor \lambda n \rfloor\} = \bigcup_{i=0}^{\lfloor \lambda n \rfloor} \{\rho_n = i\} \cap \{\sigma_1+\cdots+\sigma_i > n/4\}$$

$$\subseteq \bigcup_{i=0}^{\lfloor \lambda n \rfloor} \{\sigma_1+\cdots+\sigma_i > n/4\}.$$

Therefore the probability of the event $\{\sigma_1+\cdots+\sigma_{\rho_n} > n/4\} \cap \{\rho_n \le \lfloor \lambda n \rfloor\}$ is, by the Chebyshev inequality, (3.5) and the formula for the sum of geometric sequence, bounded by

$$\sum_{i=0}^{\lfloor \lambda n \rfloor} \mathbb{E}e^{\gamma_1(\sigma_1+\cdots+\sigma_i)}e^{-\gamma_1 n/4} \le e^{-\gamma_1 n/4}\sum_{i=0}^{\lfloor \lambda n \rfloor}(e^{L_1})^i \le e^{-\gamma_1 n/4}\frac{(e^{L_1})^{\lfloor \lambda n \rfloor+1}-1}{e^{L_1}-1}$$

$$\le M_4 e^{L_1\lfloor \lambda n \rfloor}\cdot e^{-\gamma_1 n/4} \le M_4 e^{(L_1\lambda-\gamma_1/4)n},$$

where M_4 is some constant. Thus using (3.11) the probability of

$$\{\sigma_1+\cdots+\sigma_{\rho_n} > n/4\} \cap \{\rho_n \le \lfloor \lambda n \rfloor\}$$

diminishes exponentially fast. Since $\#\{i \le n : Z_i^x \notin [a,1-a]\} = \sigma_1+\cdots+\sigma_{\rho_n}$, this means that the probability of

$$\left\{\frac{\#\{i \le n : Z_i^x \in [a,1-a]\}}{n} < 3/4\right\} \cap \left\{\rho_n \le \lfloor \lambda n \rfloor\right\}$$

diminishes exponentially fast as n goes to infinity, which is the desired assertion.

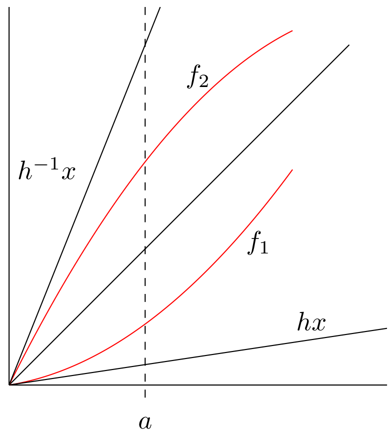

Figure 3.4

3.4 The proof of Theorem 3.2

Let us recall the statement of Theorem 3.2.

Theorem. *If $f_1, \ldots, f_m$ are C^2 diffeomorphisms satisfying (A1) and (A2), $(p_1, \ldots, p_m)$ is such that $\Lambda_0, \Lambda_1 > 0$, and $a \in (0, 1/2)$, then there exist constants $\bar{C}_3 \geq 1$ and $\bar{q}_3 < 1$ with*

$$\mathbb{E}|Z_n^a - Z_n^{1-a}| \leq \bar{C}_3 \bar{q}_3{}^n$$

for $n \geq 1$.

Let $h > 0$ be so small that

$$h \leq \min\left\{ \min_{i=1,\ldots,m} \frac{f_i'(0)}{2}, \min_{i=1,\ldots,m} \frac{f_i'(1)}{2} \right\} \quad \text{and} \quad h^{-1} \geq \max\left\{ \max_{i=1,\ldots,m} 2f_i'(0), \max_{i=1,\ldots,m} 2f_i'(1) \right\}.$$

Clearly we can assume a to be as small as we wish. This simplifies the control of the behaviour of the random walk outside $[a, 1-a]$. Therefore let us take $a > 0$ such that

1. Proposition 3.1 and 3.2 are satisfied,
2. the transition from $(0, a)$ to $(1-a, 1)$ and from $(1-a, 1)$ to $(0, a)$ is impossible in one step,
3. for every $x < a/h$ and $i = 1, \ldots, m$ we have $f_i(x) > hx$ and $f_i(1-x) < 1 - hx$,
4. for every $x \leq a$ and $i = 1, \ldots, m$ we have $f_i(x) \leq h^{-1}x$ and $f_i(1-x) \geq 1 - h^{-1}x$,
5. $(1 + h^{-1})a < 1/2$ and $(1 + h^{-1})aL'' < \gamma$, where L'' is the maximum value of the derivative of $\log f_i'$ on $[0, 1]$, $i = 1, \ldots, m$, and γ is the constant from Proposition 3.2,
6. $\mu\big((0, a)\big) \log \frac{\bar{C}_1}{h^{2+\alpha}} < \frac{|\Lambda|}{8}$ and $\mu\big((1-a, 1)\big) \log \frac{\bar{C}_1}{h^{2+\alpha}} < \frac{|\Lambda|}{8}$, where $\bar{C}_1$ is the constant in Proposition 3.2 and Λ is the volume Lyapunov exponent

$$\Lambda = \sum_{i=1}^{m} p_i \int_{(0,1)} \log f_i'(x) \mu(dx),$$

which is negative by Theorem 2.4.

The proof is an adaptation of the technique from [LP82] (Theorem 1 therein) elaborated for random systems on the circle (and hence on a compact space, see also Proposition 4.18 in [GK20]). The idea here is to define a sequence of stopping times τ_k such that $\mathbb{E}|f_\omega^{\tau_k}(x) - f_\omega^{\tau_k}(y)|^\beta < |x-y|^{k\beta}$ for every $k \geq 1$. By design it shall also satisfy $f_\omega^{\tau_k}(x)$, $f_\omega^{\tau_k}(y) \in [a, 1-a]$ for $k \geq 1$.

To start the proof let us define $\phi(x) = \sum_{i=1}^m p_i \log f_i'(x)$ on $[0,1]$.

Lemma 3.2. *There exists k such that*

$$\frac{1}{k+1}\mathbb{E}\big(\phi(Z_k^x) + \cdots + \phi(Z_0^x)\big) < 3\Lambda/4, \tag{3.12}$$

for all $x \in [a, 1-a]$, and

$$\frac{1}{k+1}\mathbb{E}\left(\mathbb{1}_{(0,a)}(Z_k^a)\log\frac{\bar{C}_1}{h^{2+\alpha}} + \cdots + \mathbb{1}_{(0,a)}(Z_0^a)\log\frac{\bar{C}_1}{h^{2+\alpha}}\right) < |\Lambda|/8, \tag{3.13}$$

$$\frac{1}{k+1}\mathbb{E}\left(\mathbb{1}_{(1-a,1)}(Z_k^{1-a})\log\frac{\bar{C}_1}{h^{2+\alpha}} + \cdots + \mathbb{1}_{(1-a,1)}(Z_0^{1-a})\log\frac{\bar{C}_1}{h^{2+\alpha}}\right) < |\Lambda|/8, \tag{3.14}$$

$$\mathbb{P}(f_\omega^k(a), f_\omega^k(1-a) \in [a, 1-a]) > 0. \tag{3.15}$$

Proof. The proof follows the lines of the proof of its counterpart version for deterministic dynamical systems (Proposition 4.1.13 and Corollary 4.1.14 in [KH95]).

Assume contrary to our claim that there exist sequences of points $(x_k) \subseteq [a, 1-a]$ and positive integers $n_1 < n_2 < \ldots$ such that

$$\frac{1}{n_k}\mathbb{E}\big(\phi(Z_{n_{k-1}}^x) + \cdots + \phi(Z_0^x)\big) \geq 3\Lambda/4 > \Lambda \tag{3.16}$$

for every $k \geq 1$ (recall that $\Lambda < 0$). Put $\nu_k := \frac{1}{n_k}(P^{n_k-1}\delta_{x_k} + \cdots + \delta_{x_k})$ (recall that P is the Markov operator corresponding to the system). Condition (3.16) may be written as $\int \phi d\nu_k \geq 3\Lambda/4$.

By Proposition 3.1 it holds that $\nu_k \in \mathcal{P}_{M,\alpha}$ for $k \geq 1$. The class $\mathcal{P}_{M,\alpha}$ is convex and weak-$*$ compact, and thus (ν_k) possesses an accumulation point $\nu \in \mathcal{P}_{M,\alpha}$. This measure is stationary due to the standard Krilov-Bogoliubov argument: if ψ is an arbitrary continuous function, then

$$\int \psi dP\nu_k = \int \psi d\nu_k + \frac{1}{n_k}\left(\int \psi dP^{n_k}\delta_{x_k} - \int \psi d\delta_{x_k}\right).$$

The modulus of the second summand tends to zero, hence $\lim_{k\to\infty}\int \psi dP\nu_k = \lim_{k\to\infty}\int \psi d\nu_k$. On the other hand, P is a Feller operator, therefore $\lim_{k\to\infty}\int \psi dP\nu_k = \lim_{k\to\infty}\int U\psi d\nu_k = \int U\psi d\nu = \int \psi dP\nu$. This proves that $P\nu = \nu$. Thus $\nu = \mu$ the unique stationary distribution. But if (3.16) holds, then $\int \phi d\nu_k$ does not tend to $\int \phi d\mu$. This is a contradiction with the fact that ϕ is a continuous observable.

Observe we have proven that there exists k' such that (3.12) holds for $k \geq k'$. The inequalities (3.13), (3.14) are the consequence of the stability of the processes (Z_n^a) and (Z_n^{1-a}). Indeed, since the stationary measure μ is atomless, the Portmanteau theorem (Theorem 2.1 in [Bil99]) implies the probability that $Z_n^a \in (0,a)$ tends to $\mu((0,a))$ and the probability that $Z_n^{1-a} \in (1-a,1)$ tends to $\mu((1-a,1))$. The same holds for the Cesàro convergence, which proves (3.13) and (3.14).

To show (3.15) observe that the second assumption made on a implies that there exists a point $z_0 \in (a, 1-a)$ belonging to the topological support of μ. Let $\varepsilon > 0$ be such that $B(z_0, 2\varepsilon) \subseteq (a, 1-a)$. Since the process is stable we have

$$\mathbb{P}(f_\omega^n(a) \in B(z_0,\varepsilon)) > \mu(B(z_0,\varepsilon))/2 > 0$$

for n sufficiently large. By Corollary 2.1 we have

$$\mathbb{P}(|f_\omega^n(a) - f_\omega^n(1-a)| < \varepsilon) \to 1$$

as n goes to infinity. Combining that gives us

$$\mathbb{P}(f_\omega^n(a), f_\omega^n(1-a) \in [a, 1-a]) \geq \mathbb{P}(f_\omega^n(a), f_\omega^n(1-a) \in B(z_0, 2\varepsilon)) > 0$$

for n sufficiently large. This gives (3.15). $\square$

A. Two sufficiently close points from $[a, 1-a]$ are in average contracted after k steps in some metric of the form $|x-y|^\beta$.

Let k denote the number from Lemma 3.2. Let $r_0(\omega)$ denote the number of appearances of the sequence $Z_0^a(\omega), \ldots, Z_k^a(\omega)$ in the set $(0, a)$, and let $r_1(\omega)$ denote the number of appearances of the sequence $Z_0^{1-a}(\omega), \ldots, Z_k^{1-a}(\omega)$ in the set $(1-a, 1)$. The summation of both expressions in the statement of Lemma 3.2 combined with $\log(ab) = \log(a) + \log(b)$ yields

$$\frac{1}{k+1}\mathbb{E}\log\left((f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(x)\left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)}\right) < \Lambda/2 < 0$$

for all $x \in [a, 1-a]$. By the compactness of $[a, 1-a]$ the supremum of the value of

$$\left(\log\left((f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(x)\left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)}\right)\right)^2\left((f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(x)\left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)}\right)^{\beta'},$$

where $x \in [a, 1-a], \beta' \in [0, 1]$, is bounded by some number M (it is also necessary to observe that $(\bar{C}_1/h^{2+\alpha})^{r_0(\omega)+r_1(\omega)} \leq (\bar{C}_1/h^{2+\alpha})^{2(k+1)}$). Once again let us apply the Taylor formula to the function $\beta \to a^\beta$, where $a > 0$, and $\beta \in (0, 1)$ is close to 0. We have $a^\beta = 1 + \beta\log a + 1/2\beta^2(\log a)^2 a^{\beta'}$, where β' is some number between 0 and β, which implies that $\beta' \in [0, 1]$. By this and the definition of M,

$$\left((f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(x) \cdot \left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)}\right)^\beta$$

$$\leq 1 + \beta\log\left((f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(x) \cdot \left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)}\right) + \frac{1}{2}M\beta^2$$

for all $x \in [a, 1-a]$ and $\beta > 0$ sufficiently close to zero. The expectation of the above is

$$\mathbb{E}\left((f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(x) \cdot \left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)}\right)^\beta$$

$$\leq 1 + \beta\mathbb{E}\log\left((f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(x) \cdot \left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)}\right) + \frac{1}{2}M\beta^2$$

$$< 1 + \beta\Lambda/2 + M\beta^2/2$$

for all $x \in [a, 1-a]$ and $\beta \in [0, 1]$. Since $\Lambda < 0$, we conclude that for some β close to zero, $\eta > 0$, and for all $x \in [a, 1-a]$ it holds that

$$\mathbb{E}\left((f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(x) \cdot \left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)}\right)^\beta \leq 1 - \eta < 1. \tag{3.17}$$

Let us also assume that β is so small that $e^{-\gamma}(L')^\beta < 1$, where L' is the supremum of derivatives of f_i's and γ is the constant in Proposition 3.2.

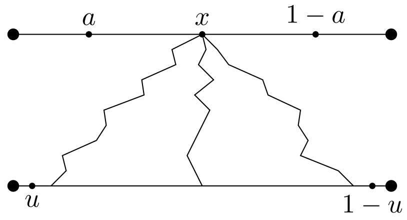

Figure 3.5: All points accessible from $[a, 1-a]$ in k steps are contained in $[u, 1-u]$.

Denote by u a positive number such that $f_\omega^k(x) \in [u, 1-u]$ for $x \in [a, 1-a]$ and $\omega \in \Omega$ (Figure 3.5). Fix n_0 so that

$$\sum_{n=n_0+1}^{\infty} \bar{C}_1 (a/u)^\alpha (e^{-\gamma}(L')^\beta)^n < \eta/2. \tag{3.18}$$

This is possible by the convergence of the series, which is implied by $e^{-\gamma}(L')^\beta < 1$ by the definition of β.

Take $a \le x < y \le 1-a$. By the mean value theorem, for every $\omega \in \Omega$ there exists $z_\omega \in [x, y]$ such that

$$|f_{\omega_k} \circ \cdots \circ f_{\omega_1}(x) - f_{\omega_k} \circ \cdots \circ f_{\omega_1}(y)|^\beta = \big((f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(z_\omega)\big)^\beta |x-y|^\beta. \tag{3.19}$$

But there is only finitely many functions $\omega \longmapsto f_{\omega_k} \circ \cdots \circ f_{\omega_1}$ hence using uniform equicontinuity we can take any $z \in [x, y]$ and write that

$$\Big(f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(z_\omega)\Big)^\beta < \Big(f_{\omega_k} \circ \cdots \circ f_{\omega_1})'(z)\Big)^\beta + \frac{\eta}{2} \cdot \frac{1}{(\bar{C}_1/h^{2+\alpha})^{2(k+1)}} \tag{3.20}$$

for every ω, provided x, y are sufficiently close to each other, say $|x-y| < \varepsilon$. Let us plug (3.20) into (3.19). Then take expectation of both sides, use (3.17) and the estimation

$$\left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)} \le \left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{2(k+1)}$$

to obtain

$$\mathbb{E}\left(|f_{\omega_k} \circ \cdots \circ f_{\omega_1}(x) - f_{\omega_k} \circ \cdots \circ f_{\omega_1}(y)| \left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0(\omega)+r_1(\omega)}\right)^\beta < (1-\eta/2)|x-y|^\beta \tag{3.21}$$

for all x, y with $|x-y| < \varepsilon$. We insist also that ε should satisfy

$$|f_\omega^n(x) - f_\omega^n(y)| \le a \quad \text{for } n = 1, 2, \ldots, k+n_0 \tag{3.22}$$

provided $|x-y| < \varepsilon$. This means in particular that after k iterates the random walk is locally (i.e. for $|x-y| < \varepsilon$) contracing in average in the metric $|x-y|^\beta$. Moreover, for each ω the distance $|f_{\omega_k} \circ \cdots \circ f_{\omega_1}(x) - f_{\omega_k} \circ \cdots \circ f_{\omega_1}(y)|^\beta$ can be multiplied by the number (greater than one) $(\bar{C}_1/h^{2+\alpha})^{\beta(r_0+r_1)}$, and the random walk is still contractive. Now, some part of trajectories is outside $[a, 1-a]$ at the moment k. The upcoming task is to compare how much the distance between trajectories at the moment of the first common return to $[a, 1-a]$ is greater than the distance at the moment k. It turns out that it may be estimated by $(\bar{C}_1/h^{2+\alpha})^{\beta(r_0+r_1)}$. This is the reason why we have proven the stronger version (3.21) of contractiveness.

B. Estimation of the average distance between trajectories of two sufficiently close points at the moment of the first return to $[a, 1-a]$.

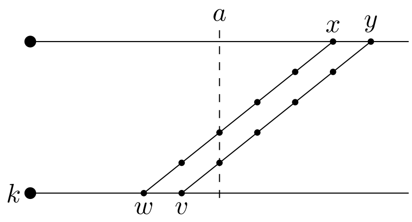

Figure 3.6

Fix x, y with $|x-y| < \varepsilon$. Put $\tau := \min\{n \geq k : f_\omega^n(x), f_\omega^n(y) \in [a, 1-a]\}$. We have

$$\mathbb{E}\frac{|f_\omega^\tau(x) - f_\omega^\tau(y)|^\beta}{|x-y|^\beta} = \mathbb{E}\bigg(\frac{|f_\omega^k(x) - f_\omega^k(y)|^\beta}{|x-y|^\beta}\mathbb{E}\bigg(\frac{|f_\omega^\tau(x) - f_\omega^\tau(y)|^\beta}{|f_\omega^k(x) - f_\omega^k(y)|^\beta}\bigg|\mathcal{F}_k\bigg)\bigg).$$

We are going to show that

$$\mathbb{E}\bigg(\frac{|f_\omega^\tau(x) - f_\omega^\tau(y)|^\beta}{|f_\omega^k(x) - f_\omega^k(y)|^\beta}\bigg|\mathcal{F}_k\bigg) \leq \bigg(\frac{\bar{C}_1}{h^{2+\alpha}}\bigg)^{\beta(r_0(\omega)+r_1(\omega))}(1+\eta/2) \quad \text{a.s.} \tag{3.23}$$

for every x, y with $|x-y| < \varepsilon$. By (3.21) this would give that

$$\mathbb{E}|f_\omega^\tau(x) - f_\omega^\tau(y)|^\beta \leq (1+\eta/2)\mathbb{E}\bigg(|f_\omega^k(x) - f_\omega^k(y)|\bigg(\frac{\bar{C}_1}{h^{2+\alpha}}\bigg)^{r_0(\omega)+r_1(\omega)}\bigg)^\beta$$

$$< (1-\eta/2)(1+\eta/2)|x-y|^\beta = (1-\eta^2/4)|x-y|^\beta. \tag{3.24}$$

Take ω with $\tau(\omega) > k$. Put $w := f_\omega^k(x)$, $v := f_\omega^k(y)$ (Figure 3.6) and observe that $w < a$ or $v > 1-a$ as $\tau > k$ (let us assume that $w < a$ to simplify the presentation). Then by the strong Markov property the value of

$$\mathbb{E}\bigg(\frac{|f_\omega^\tau(x) - f_\omega^\tau(y)|^\beta}{|f_\omega^k(x) - f_\omega^k(y)|^\beta}\bigg|\mathcal{F}_k\bigg)$$

on ω may be rewritten as

$$\mathbb{E}\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta},$$

where $T(\omega) := \min\{n \geq 1 : f_\omega^n(w), f_\omega^n(v) \in [a, 1-a]\}$.

We have

$$\mathbb{E}\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta} = \mathbb{E}\mathbb{1}_{\{T\leq n_0\}}\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta} + \mathbb{E}\mathbb{1}_{\{T>n_0\}}\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta},$$

where n_0 was defined to be such that $\sum_{n=n_0+1}^{\infty} \bar{C}_1(a/u)^\alpha (e^{-\gamma}(L')^\beta)^n < \eta/2$ (see (3.18)).

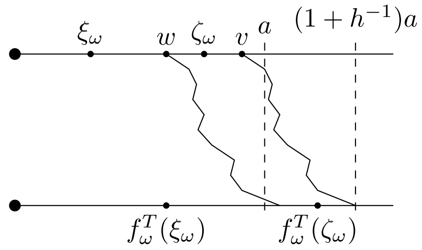

Figure 3.7

To estimate the first integral let us take $\omega \in \{T \le n_0\}$ and write by the mean value theorem

$$|f_\omega^T(w) - f_\omega^T(v)| = (f_\omega^T)'(\zeta_\omega)|w - v|$$

and

$$f_\omega^T(w) = f_\omega^T(w) - f_\omega^T(0) = (f_\omega^T)'(\xi_\omega) \cdot w$$

for some $\zeta_\omega \in [w, v]$ and $\xi_\omega \in [0, w]$. Since $T \le n_0$, we have $|f_\omega^n(w) - f_\omega^n(v)| \le a$ for $n \le T$ (by the choice of ε, see (3.22)). By the choice of a (see point 5.) it is impossible that $f_\omega^n(w) < a$ and $f_\omega^n(v) > 1 - a$ for some $n \le T$. Thus T may be regarded as the moment of the first visit of $f_\omega^n(w)$ in $[a, 1-a]$ (it is crucial we are restricted to the event $\{T \le n_0\}$). Since T is the moment of the first visit in $[a, 1-a]$ the value of $f_\omega^T(w)$ cannot be greater than $h^{-1}a$ (cf. the definition of h at the beginning of the section and point 4. in assumptions on a). Therefore $f_\omega^T(w)/w \le h^{-1}a/w$ and

$$\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w - v|^\beta} = \big((f_\omega^T)'(\zeta_\omega)\big)^\beta = \left(\frac{(f_\omega^T)'(\zeta_\omega)}{(f_\omega^T)'(\xi_\omega)}\right)^\beta \left(\frac{f_\omega^T(w)}{w}\right)^\beta$$

$$\le \left(\frac{h^{-1}a}{w}\right)^\beta \cdot \exp\Big(\beta\big(\log(f_\omega^T)'(\zeta_\omega) - \log(f_\omega^T)'(\xi_\omega)\big)\Big).$$

Using $T \le n_0$ again we have $|f_\omega^n(\xi_\omega) - f_\omega^n(\zeta_\omega)| < (1+h^{-1})a$ for $n \le T$. Indeed, we have just shown that $f_\omega^n(w) \le h^{-1}a$ for $n \le T$, hence this follows from (3.22) (see Figure 3.7). By the chain rule, the fact that L'' is a supremum of the derivative of $\log f_i'$ on $[0,1]$ for all $i = 1, \ldots, m$ we obtain

$$|\log(f_\omega^T)'(\zeta_\omega) - \log(f_\omega^T)'(\xi_\omega)| \le TL''(1 + h^{-1})a,$$

hence

$$\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w - v|^\beta} \le \left(\frac{h^{-1}a}{w}\right)^\beta \cdot \exp(TL''(1 + h^{-1})a\beta)$$

for ω such that $T \le n_0$.

If $T > n_0$, then

$$\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w - v|^\beta} \le (L')^{\beta T}$$

since L' is the supremum of derivatives of f_i for all $i = 1, \ldots, m$. Therefore

$$\mathbb{E}\mathbb{1}_{\{T>n_n\}} \frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w - v|^\beta} \le \sum_{n=n_0+1}^{\infty} \mathbb{P}(T \ge n)(L')^{\beta n} \le \sum_{n=n_0+1}^{\infty} \bar{C}_1 \left(\frac{a}{w}\right)^\alpha (L')^{\beta n} e^{-\gamma n},$$

where in the last inequality we used the Chebyshev inequality and the fact that $\mathbb{E}e^{\gamma T} \leq \bar{C}_1\left(\frac{a}{w}\right)^{\alpha}$ (which is the statement of Proposition 3.2). Now, recall that u was chosen so that $w, v \in [u, 1-u]$ whatever ω is, and thus $\frac{1}{w^{\alpha}} \leq \frac{1}{u^{\alpha}}$. Henceforth

$$\mathbb{E}\mathbb{1}_{\{T>n_n\}} \frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta} \leq \sum_{n=n_0+1}^{\infty} \bar{C}_1 \left(\frac{a}{w}\right)^{\alpha} (L')^{\beta n} e^{-\gamma n} \leq \eta/2$$

by (3.18). Combining these two estimations with the application of the Jensen inequality to the function $t \longmapsto t^\beta$ yields

$$\mathbb{E}\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta} = \mathbb{E}\mathbb{1}_{\{T\leq n_n\}} \frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta} + \mathbb{E}\mathbb{1}_{\{T>n_n\}} \frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta}$$

$$\leq \left(\frac{h^{-1}a}{w}\right)^{\beta} \mathbb{E}\exp(TL''(1+h^{-1})a\beta) + \eta/2 \leq \left(\frac{h^{-1}a}{w}\right)^{\beta} \left(\mathbb{E}\exp(TL''(1+h^{-1})a)\right)^{\beta} + \eta/2 \quad (3.25)$$

Recall that (point 5. in the assumptions on a) that $L''(1+h^{-1})a < \gamma$, where γ is the constant in Proposition 3.2. Using this proposition yields

$$\mathbb{E}\exp(TL''(1+h^{-1})a) \leq \mathbb{E}\exp(T\gamma) \leq \bar{C}_1\left(\frac{a}{w}\right)^{\alpha}.$$

Plugging that into (3.25) gives

$$\mathbb{E}\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta} \leq \left(\frac{h^{-1}a^{1+\alpha}}{w^{1+\alpha}}\bar{C}_1\right)^{\beta} + \eta/2.$$

Since the expression in parenthesis is greater than one (the choice of h easily implies $h < 1$) we can write

$$\mathbb{E}\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta} \leq \left(\left(\frac{a}{w}\right)^{1+\alpha}\frac{\bar{C}_1}{h}\right)^{\beta}(1+\eta/2).$$

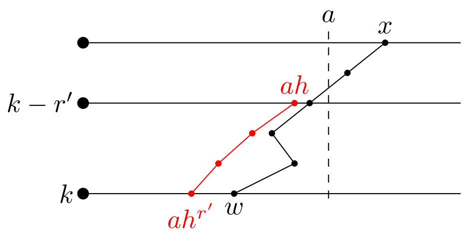

Figure 3.8

Going back to x and y, observe that the distance of w to zero may be estimated using $r_0(\omega)$. Namely $w = f_\omega^k(x) \geq f_\omega^k(a) \geq h^{r'}a$, where r' is the maximum integer n such that $f_\omega^k(a) < a$, $f_\omega^{k-1}(a) < a, \ldots, f^{k-n}(a) < a$ (it follows from the definition of h, see Figure 3.8). Hence

$$\frac{a}{w} \leq \frac{1}{h^{r'}} \leq \frac{1}{h^{r_0}}$$

and thus

$$\mathbb{E}\frac{|f_\omega^T(w) - f_\omega^T(v)|^\beta}{|w-v|^\beta} \leq \left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{r_0\beta}(1+\eta/2) \leq \left(\frac{\bar{C}_1}{h^{2+\alpha}}\right)^{(r_0+r_1)\beta}(1+\eta/2),$$

which is the desired assertion. Note that if ω was such that $f_\omega^k(y) > 1-a$, then r_1 would be used instead of r_0. Now (3.24) follows. In fact we proved more. Namely

$$\mathbb{E}|f_\omega^{\tau\wedge n}(x) - f_\omega^{\tau\wedge n}(y)|^\beta < (1-\eta^2/4)|x-y|^\beta \quad (3.26)$$

for an arbitrary integer $n \geq k$ and x, y with $|x-y| < \varepsilon$.

C. The definition of τ for all pairs of points in $[a, 1-a]$.

Fix $a \le x < y \le 1-a$ with $|x-y| < \varepsilon$. The random variable τ has already been defined as the minimum $n \ge k$ with $a \le f_\omega^n(x) < f_\omega^n(y) \le 1-a$. We are going to extend this definition to the case when $a \le x < y \le 1-a$ and $|x-y| \ge \varepsilon$. To this end recall that the system is synchronizing by Corollary 2.1, which implies that $|f_\omega^n(a) - f_\omega^n(1-a)| \to 0$ as $n \to \infty$ almost surely. Therefore there exists K such that

$$\mathbb{P}\bigg(\exists_{n\ge K}|f_\omega^n(a) - f_\omega^n(1-a)|^\beta \ge \frac{1}{2}(1-\eta^2/4)\varepsilon^\beta\bigg) \le \frac{1}{2}(1-\eta^2/4)\varepsilon^\beta. \tag{3.27}$$

Given $x, y \in [a, 1-a]$ with $|x-y| \ge \varepsilon$, set $\tau = \tau(x,y) = \min\{n \ge K : f_\omega^n(x), f_\omega^n(y) \in [a, 1-a]\}$. By (3.27) it holds that

$$\mathbb{E}|f_\omega^{\tau\wedge n}(x) - f_\omega^{\tau\wedge n}(y)|^\beta \le (1-\eta^2/4)|x-y|^\beta \tag{3.28}$$

and $f_\omega^\tau(x)$, $f_\omega^\tau(y) \in [a, 1-a]$ provided $n \ge K$ and $x, y \in [a, 1-a]$ satisfy $|x-y| \ge \varepsilon$. Moreover, using exactly the same argument as in the proof of (3.15) in Lemma 3.2 (i.e. combining stability with synchronization) we can assume that K is chosen so that

$$\mathbb{P}(f_\omega^K(a), f_\omega^K(1-a) \in [a, 1-a]) > 0. \tag{3.29}$$

D. Some exponential moment of τ is finite.

Lemma 3.3. *There exists $\bar{C}_4$ such that $\mathbb{E}e^{\gamma\tau(x,y)} \le \bar{C}_4$ for all x, y in $[a, 1-a]$, where γ is the constant in Proposition 3.2.*

Proof. There are two cases: if the distance between x and y is less than ε, then τ is equal to the moment of the first visit T of $(f_\omega^k(x), f_\omega^k(y))$ in $[a, 1-a] \times [a, 1-a]$ along the trajectory $\theta^k\omega$. Since we know that $u \le f_\omega^k(x) < f_\omega^k(y) \le 1-u$ for some $u > 0$, we can use Proposition 3.2 to show that

$$\mathbb{E}e^{\gamma\tau(x,y)} \le e^{\gamma k}\mathbb{E}e^{\gamma T} \le e^{\gamma k}\bar{C}_1\left(\frac{a}{u}\right)^\alpha,$$

which is independent of x and y (T has been defined in part B).

In the case when the distance between x and y is greater or equal to ε the proof is very similar to above. The only changes are that k is replaced by K and u is possibly smaller. $\square$

E. The definition of (τ_n)

Let us define the sequence $\tau_1, \tau_2 \dots$ of random moments inductively in the following way. Let $\tau_1 := \tau$. If $w = f_\omega^\tau(x)$ and $v = f_\omega^\tau(y)$, then $w, v \in [a, 1-a]$ hence we can put

$$\tau_2(x,y)(\omega) = \tau_1(x,y)(\omega) + \tau(w,v)(\theta^{\tau_1}\omega).$$

If $\tau_n(x,y)$ is already defined, then put $w = f_\omega^{\tau_n}(x)$, $v = f_\omega^{\tau_n}(y)$, and define

$$\tau_{n+1}(x,y)(\omega) = \tau_n(x,y)(\omega) + \tau(w,v)(\theta^{\tau_n}\omega).$$

Given $x, y \in [a, 1-a]$ and integer n define

$$A_j = \{\omega \in \Omega : |f_\omega^{\tau_j}(x) - f_\omega^{\tau_j}(y)| < \varepsilon \text{ and } n - \tau_j \ge k\}$$

$$\cup\{\omega \in \Omega : |f_\omega^{\tau_j}(x) - f_\omega^{\tau_j}(y)| \geq \varepsilon \text{ and } n - \tau_j \geq K\}$$

and set B_j to be the complement of A_j (note we do not include the dependence of n and x, y; it will be clear from the context). Combining (3.26) with (3.28) yields that for $\omega \in A_{j-1}$ we have

$$\mathbb{E}\left(|f_\omega^{\tau_j \wedge n}(x) - f_\omega^{\tau_j \wedge n}(y)|^\beta \middle| \mathcal{F}_{\tau_{j-1}}\right)(\omega) < (1 - \eta^2/4)|f_\omega^{\tau_{j-1}}(x) - f_\omega^{\tau_{j-1}}(y)|^\beta. \tag{3.30}$$

By the same reason

$$\mathbb{E}\left(|f_\omega^{\tau_j}(x) - f_\omega^{\tau_j}(y)|^\beta \middle| \mathcal{F}_{\tau_{j-1}}\right) < (1 - \eta^2/4)|f_\omega^{\tau_{j-1}}(x) - f_\omega^{\tau_{j-1}}(y)|^\beta \quad \text{a.s.} \tag{3.31}$$

Recall that $L' = \max_{i=1,\dots,m, z \in [0,1]} f_i'(z)$, and put $M := (L')^{\beta \max\{k,K\}}$, where k and K are the integers from the construction of τ. If follows from the definition of B_{j-1} that, given n, x, y, we have

$$|f_\omega^n(x) - f_\omega^n(y)|^\beta \leq M|f_\omega^{\tau_{j-1}}(x) - f_\omega^{\tau_{j-1}}(y)|^\beta \quad \text{for } \omega \in B_{j-1}. \tag{3.32}$$

F. The sequence τ_n is proportional to n up to an event of probability diminishing exponentially fast.

To finish the proof we need one more random variable. Let $\rho_n(\omega) := \max\{j \geq 0 : \tau_j \leq n\}$.

Lemma 3.4. *Let γ be the constant in Proposition 3.2, and let $\bar{C}_4$ be the constant given by Lemma 3.3. If $\lambda = \gamma/(2\log(\bar{C}_4))$, then $\mathbb{P}(\rho_n < \lambda n)$ decays exponentially fast as n goes to infinity.*

Proof. Let us observe that Lemma 3.3 gives that $\mathbb{E}e^{\gamma\tau_n} \leq \bar{C}_4{}^n$ for every n (this kind of proofs has already appeared several times). Indeed, we have

$$\mathbb{E}e^{\gamma\tau_n} = \mathbb{E}\big(e^{\gamma\tau_n}\big|\mathcal{F}_{\tau_{n-1}}\big) = \mathbb{E}e^{\gamma\tau_{n-1}}\mathbb{E}\big(e^{\gamma\tau(f_\omega^{\tau_{n-1}}(x), f_\omega^{\tau_{n-1}}(y)(\theta^{\tau_{n-1}}\omega)}\big|\mathcal{F}_{\tau_{n-1}}\big).$$

Lemma 3.3 and the strong Markov property imply that the conditional expectation above does not exceed $\bar{C}_4$. Proceeding in this manner gives that $\mathbb{E}e^{\gamma\tau_n} \leq \bar{C}_4{}^n$ for every n. Therefore the Chebyshev inequality yields

$$\mathbb{P}(\rho_n < \lambda n) \leq \mathbb{P}(\tau_{\lfloor \lambda n \rfloor} > n) \leq e^{-\gamma n}\mathbb{E}e^{\gamma\tau_{\lfloor \lambda n \rfloor}} \leq e^{\lambda n \log \bar{C}_4 - \gamma n} = e^{-\gamma/2n},$$

which decays exponentially fast. □

G. The end of the proof.

To simplify the notation put $x_n = x_n(\omega) := f_\omega^n(x)$ and $y_n = y_n(\omega) := f_\omega^n(y)$. Fix n, and put $l := \max\{\rho_n : \tau_{\rho_n} \leq n\}$. Recall that M has been defined just after (3.31) and is equal to $(L')^{\beta \max\{k,K\}}$.

Lemma 3.5. *It holds that*

$$|x_n - y_n|^\beta < M|x_{\tau_l} - y_{\tau_l}|^\beta \quad \textit{a.s. on } \{\tau_l \leq n\}.$$

Proof. Let us observe that

$$\mathbb{P}\Big(x_{\tau_l+k}, y_{\tau_l+k} \in [a, 1-a]|\mathcal{F}_{\tau_l}\Big)$$

is by (3.15) positive a.s. on $\{\omega \in \Omega : |x_{\tau_l}(\omega) - y_{\tau_l}(\omega)| < \varepsilon\}$, hence $\tau_{l+1}(\omega) = \tau_l(\omega) + k$ with positive conditional probability on $\{\omega \in \Omega : |x_{\tau_l}(\omega) - y_{\tau_l}(\omega)| < \varepsilon\}$. If $n - \tau_l(\omega) \geq k$, then $\tau_{l+1}(\omega) \leq n$ with positive conditional probability on $\{\omega \in \Omega : |x_{\tau_l}(\omega) - y_{\tau_l}(\omega)| < \varepsilon\}$, which is a contradiction with the definition of l provided the event $\{\omega \in \Omega : |x_{\tau_l}(\omega) - y_{\tau_l}(\omega)| < \varepsilon\}$ is of positive probability. The analogous statement may be proved in the case when $|x_{\tau_l}(\omega) - y_{\tau_l}(\omega)| \geq \varepsilon$ (then k is replaced by K and (3.29) is used instead of (3.15)). Since at least one of the events $\{\omega \in \Omega : |x_{\tau_l}(\omega) - y_{\tau_l}(\omega)| < \varepsilon\}$ and $\{\omega \in \Omega : |x_{\tau_l}(\omega) - y_{\tau_l}(\omega)| \geq \varepsilon\}$ has positive probability, this completes the proof. $\square$

Now we proceed with the final calculations.

$$\mathbb{E}|x_n - y_n|^\beta = \int_{\{\tau_l \leq n\}} |x_n - y_n|^\beta \mathrm{d}\mathbb{P} + \int_{\{\tau_{l-1} \leq n\} \cap B_{l-1} \cap \{\tau_l > n\}} |x_n - y_n|^\beta \mathrm{d}\mathbb{P}$$
$$+ \int_{\{\tau_{l-1} \leq n\} \cap A_{l-1} \cap \{\tau_l > n\}} |x_n - y_n|^\beta \mathrm{d}\mathbb{P} + \int_{\{\tau_{l-1} > n\}} |x_n - y_n|^\beta \mathrm{d}\mathbb{P}.$$

The application of Lemma 3.5 to the first term on the right-hand side, (3.32) to the second, and the multiplication of the third by $M \geq 1$ gives

$$\mathbb{E}|x_n - y_n|^\beta \leq M \int_{\{\tau_l \leq n\}} |x_{\tau_l} - y_{\tau_l}|^\beta \mathrm{d}\mathbb{P} + M \int_{\{\tau_{l-1} \leq n\} \cap B_{l-1} \cap \{\tau_l > n\}} |x_{\tau_{l-1}} - y_{\tau_{l-1}}|^\beta \mathrm{d}\mathbb{P}$$
$$+ M \int_{\{\tau_{l-1} \leq n\} \cap A_{l-1} \cap \{\tau_l > n\}} |x_n - y_n|^\beta \mathrm{d}\mathbb{P} + \int_{\{\tau_{l-1} > n\}} |x_n - y_n|^\beta \mathrm{d}\mathbb{P}.$$

The first and third term on the right-hand side may be replaced by one integral as follows:

$$\mathbb{E}|x_n - y_n|^\beta \leq M \int_{\{\tau_{l-1} \leq n\} \cap A_{l-1}} |x_{\tau_l \wedge n} - y_{\tau_l \wedge n}|^\beta \mathrm{d}\mathbb{P}$$
$$+ M \int_{\{\tau_{l-1} \leq n\} \cap B_{l-1} \cap \{\tau_l > n\}} |x_{\tau_{l-1}} - y_{\tau_{l-1}}|^\beta \mathrm{d}\mathbb{P} + \int_{\{\tau_{l-1} > n\}} |x_n - y_n|^\beta \mathrm{d}\mathbb{P}.$$

Since $\{\tau_{l-1} \leq n\} \cap A_{l-1}$ is $\mathcal{F}_{\tau_{l-1}}$-measurable we can apply (3.30) to it and obtain

$$\mathbb{E}|x_n - y_n|^\beta \leq M \int_{\{\tau_{l-1} \leq n\} \cap A_{l-1}} |x_{\tau_{l-1}} - y_{\tau_{l-1}}|^\beta \mathrm{d}\mathbb{P}$$
$$+ M \int_{\{\tau_{l-1} \leq n\} \cap B_{l-1} \cap \{\tau_l > n\}} |x_{\tau_{l-1}} - y_{\tau_{l-1}}|^\beta \mathrm{d}\mathbb{P} + \int_{\{\tau_{l-1} > n\}} |x_n - y_n|^\beta \mathrm{d}\mathbb{P}$$
$$\leq M \int_{\{\tau_{l-1} \leq n\}} |x_{\tau_{l-1}} - y_{\tau_{l-1}}|^\beta \mathrm{d}\mathbb{P} + \int_{\{\tau_{l-1} > n\}} |x_n - y_n|^\beta \mathrm{d}\mathbb{P}.$$

We can repeat this reasoning all over again but this time omitting the first step, in which Lemma 3.5 has been used. This leads us finally to the inequality

$$\mathbb{E}|x_n - y_n|^\beta \leq M \int_{\{\tau_{\lfloor \lambda n \rfloor} \leq n\}} |x_{\tau_{\lfloor \lambda n \rfloor}} - y_{\tau_{\lfloor \lambda n \rfloor}}|^\beta \mathrm{d}\mathbb{P} + M\mathbb{P}(\tau_{\lfloor \lambda n \rfloor} > n).$$

The second summand, by Lemma 3.4, decays exponentially fast. For the first summand we have, by (3.31),

$$\int_{\{\tau_{\lfloor \lambda n \rfloor} \leq n\}} |x_{\tau_{\lfloor \lambda n \rfloor}} - y_{\tau_{\lfloor \lambda n \rfloor}}|^\beta \mathrm{d}\mathbb{P} \leq \int_\Omega |x_{\tau_{\lfloor \lambda n \rfloor}} - y_{\tau_{\lfloor \lambda n \rfloor}}|^\beta \mathrm{d}\mathbb{P} \leq (1 - \eta^2/2)^{\lfloor \lambda n \rfloor} |x - y|^\beta.$$

The assertion follows.

3.5 The proof of Theorem 3.1

We have

$$\|U^n\varphi\|^2_{L^2(\mu)} = \int_{[0,1]} |U^n\varphi(x)|^2\mu(dx) = \int_{[0,1]} \left|U^n\varphi(x) - \int_{[0,1]} \varphi(y)\mu(dy)\right|^2 \mu(dx)$$

$$= \int_{[0,1]} \left|U^n\varphi(x) - \int_{[0,1]} \varphi(y)P^n\mu(dy)\right|^2 \mu(dx) = \int_{[0,1]} \left|U^n\varphi(x) - \int_{[0,1]} U^n\varphi(y)\mu(dy)\right|^2 \mu(dx)$$

$$= \int_{[0,1]} \left|\int_{[0,1]} \big(U^n\varphi(x) - U^n\varphi(y)\big)\mu(dy)\right|^2 \mu(dx) \le \int_{[0,1]}\int_{[0,1]} \big|U^n\varphi(x) - U^n\varphi(y)\big|^2\mu(dy)\mu(dx),$$

where the last inequality is the Jensen inequality.

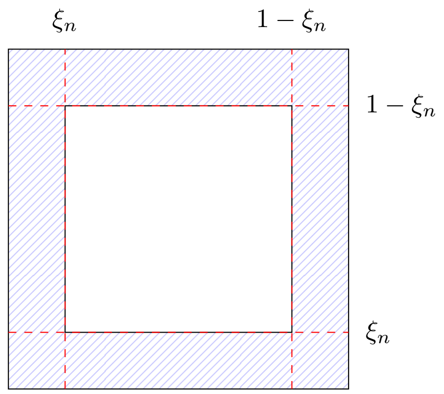

Figure 3.9: The set R_n. It may be covered by the sets $(0, \xi_n) \cup (1-\xi_n, 1) \times [0,1]$ and $[0,1] \times (0, \xi_n) \cup (1-\xi_n, 1)$. Each of these sets has measure at most $2M\xi_n^\alpha$ as $\mu \in \mathcal{P}_{M,\alpha}$ (see Corollary 3.1).

To continue the estimation we take the sequence $\xi_n := e^{-\gamma n/2}$, where γ is given in Proposition 3.2, and decompose $[0,1]\times[0,1]$ as the sum of $K_n = [\xi_n, 1-\xi_n]\times[\xi_n, 1-\xi_n]$ and the remainder R_n. By Corollary 3.1 it is easy to see (cf. Figure 3.9) that $\mu\otimes\mu(R_n) \le 2\mu((0,\xi_n)\cup(1-\xi_n,1)) \le 4M\xi_n^\alpha$, since $\mu \in \mathcal{P}_{M,\alpha}$. Then

$$\|U^n\varphi\|^2_{L^2(\mu)} \le \int_{[0,1]}\int_{[0,1]} \big|U^n\varphi(x) - U^n\varphi(y)\big|^2\mu(dy)\mu(dx)$$

$$\le \int\int_{K_n} \Big(\mathrm{Lip}(\varphi)\mathbb{E}\big|Z_n^x - Z_n^y\big|\Big)^2 \mu(dy)\mu(dx) + \int\int_{R_n} \big|U^n\varphi(x) - U^n\varphi(y)\big|^2\mu(dy)\mu(dx)$$

$$\le \Big(\mathrm{Lip}(\varphi)\mathbb{E}\big|Z_n^{\xi_n} - Z_n^{1-\xi_n}\big|\Big)^2 + \int\int_{R_n} 4\|\varphi\|^2_\infty\mu(dy)\mu(dx)$$

$$\le \Big(\mathrm{Lip}(\varphi)\mathbb{E}\big|Z_n^{\xi_n} - Z_n^{1-\xi_n}\big|\Big)^2 + 16M\xi_n^\alpha\|\varphi\|^2_\infty,$$

where in the one but last inequality we used the fact that $\|U^n\varphi\|_\infty \le \|\varphi\|_\infty$ and the triangle inequality.

Since $\xi_n^\alpha = e^{-\alpha\gamma n/2}$ shrinks exponentially fast, we are left to estimate the first summand. Let us fix n and write for short $T = T^a(\xi_n, 1-\xi_n)$, $w = Z_T^{\xi_n}$ and $v = Z_T^{1-\xi_n}$ (the definition of T^a is in the statement of Proposition 3.2). We have

$$\begin{aligned}\mathbb{E}\big|Z_n^{\xi_n} - Z_n^{1-\xi_n}\big| &= \mathbb{E}\mathbb{1}_{\{T\le n/2\}}\big|Z_n^{\xi_n} - Z_n^{1-\xi_n}\big| + \mathbb{E}\mathbb{1}_{\{T>n/2\}}\big|Z_n^{\xi_n} - Z_n^{1-\xi_n}\big| \\ &\le \mathbb{E}\mathbb{1}_{\{T\le n/2\}}\mathbb{E}\Big(\big|Z_n^{\xi_n} - Z_n^{1-\xi_n}\big| \;\Big|\; \mathcal{F}_T\Big) + \mathbb{P}(T > n/2) \\ &= \mathbb{E}\mathbb{1}_{\{T\le n/2\}}\mathbb{E}\big|Z_{n-T}^w - Z_{n-T}^v\big| + \mathbb{P}(T > n/2) \\ &\le \mathbb{E}\mathbb{1}_{\{T\le n/2\}}\mathbb{E}\big|Z_{n-T}^a - Z_{n-T}^{1-a}\big| + \mathbb{P}(T > n/2).\end{aligned}$$

Now, $n - T > n/2$ on $\{T \le n/2\}$ and $\mathbb{P}(T > n/2) = \mathbb{P}(e^{\gamma T} > e^{\gamma n/2}) \le e^{-\gamma n/2}\mathbb{E}e^{\gamma T}$ by the Chebyshev inequality. Combining this with Proposition 3.2 and Theorem 3.2 yields

$$\begin{aligned}\mathbb{E}\big|Z_{n-T}^{\xi_n} - Z_{n-T}^{1-\xi_n}\big| &\le \mathbb{E}\big|Z_{n/2}^a - Z_{n/2}^{1-a}\big| + e^{-\gamma n/2}\bar{C}_1(a/\xi_n)^\alpha \le \bar{C}_3\bar{q}_3{}^{n/2} + \bar{C}_1 a^\alpha e^{-\gamma n/2}\xi_n^{-\alpha} \\ &\le \bar{C}_3\bar{q}_3{}^{n/2} + \bar{C}_1 a^\alpha e^{-\gamma n/2} e^{\alpha\gamma n/2} = \bar{C}_3\bar{q}_3{}^{n/2} + \bar{C}_1 a^\alpha e^{\gamma n/2(\alpha-1)}.\end{aligned}$$

This completes the proof as $\alpha - 1 < 0$.

3.6 The central limit theorem

Limit theorems for additive functionals of Markov processes are usually proven by decomposing a process into the sum of a martingale and some rest R_n, which divided by the square root of n tends to zero almost surely. This method has been invented by Gordin and Lifšic [GL78]. Let $f_1, \ldots, f_m$ be C^2 orientation preserving interval diffeomorphisms satisfying (A1) and (A2). Let $p_1, \ldots, p_m$ be a probability vector such that Λ_0 and Λ_1 are positive. Finally, let φ be a Lipschitz function with $\int \varphi d\mu = 0$, where μ is the unique stationary distribution. The equation

$$\varphi = U\psi - \psi$$

is called the Poisson equation. Here it has a solution in $L^2(\mu)$ due to Theorem 3.1. Indeed, let

$$\psi := \sum_{n=0}^{\infty} U^n\varphi,$$

where the convergence is in $L^2(\mu)$ norm. Then

$$U\psi = \sum_{n=0}^{\infty} U^{n+1}\varphi = \sum_{n=0}^{\infty} U^n\varphi - \varphi = \psi - \varphi,$$

thus ψ is a solution of the Poisson equation. Observe that $U\psi(X_n) - \psi(X_{n+1})$ are stationary, ergodic increments of a martingale. Two first properties are trivial (since (X_n) has these properties), while the third one is due to the simple observation $\mathbb{E}(U\psi(X_n) - \psi(X_{n+1})|X_n) = U\psi(X_n) - U\psi(X_n) = 0$.

Using the Poisson equation $\varphi = U\psi - \psi$ we can decompose

$$\varphi(X_1) + \cdots + \varphi(X_n) = \big(U\psi(X_1) - \psi(X_2) + \cdots + U\psi(X_{n-1}) - \psi(X_n)\big) + U\psi(X_n) - \psi(X_1),$$

where the part in the parenthesis, denoted by M_n, is the sum of square integrable, ergodic, stationary martingale increments with respect to the filtration generated by (X_n). The rest, denoted by R_n, has the property that $R_n/\sqrt{n} \to 0$ almost surely.

The martingale (M_n) satisfies the central limit theorem for martingales proved in [Bro71] (actually a version of the central limit theorem for martingales sufficient for our needs has been proven previously by Billingsley [Bil61]).

Corollary 3.2. *Let $f_1, \ldots, f_m$ be a system of C^2 diffeomorphisms satisfying (A1) and (A2). Let $(p_1, \ldots, p_m)$ be such that $\Lambda_0, \Lambda_1 > 0$. Let (X_n) be the unique stationary Markov process corresponding to the system. If φ is a Lipschitz real function on $(0,1)$ with $\int \varphi d\mu = 0$, then the additive functional $(\varphi(X_1) + \cdots + \varphi(X_n))$ satisfies the central limit theorem, i.e.*

$$\mathbb{P}\left(\frac{\varphi(X_1) + \cdots + \varphi(X_n)}{\sqrt{n}} \in [a,b]\right) \to \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b e^{-\frac{x^2}{2\sigma^2}} dx$$

as $n \to \infty$, where $\sigma^2 = \|\psi\|^2_{L^2(\mu)} - \|U\psi\|^2_{L^2(\mu)}$.

The martingale satisfies also the law of the iterated logarithm (see Theorem 4.7 in [HH80]).

Corollary 3.3. *Let $f_1, \ldots, f_m$ be a system of C^2 diffeomorphisms satisfying (A1) and (A2). Let $(p_1, \ldots, p_m)$ be such that $\Lambda_0, \Lambda_1 > 0$. Let (X_n) be the unique stationary Markov process corresponding to the system. If φ is a Lipschitz real function on $(0,1)$ with $\int \varphi d\mu = 0$, then the additive functional $(\varphi(X_1) + \cdots + \varphi(X_n))$ satisfies the law of the iterated logarithm, i.e.*

$$\limsup_{n\to\infty} \frac{\varphi(X_1) + \cdots + \varphi(X_n)}{\sqrt{2n\sigma^2 \log\log n}} = 1 \quad a.s.,$$

where $\sigma^2 = \|\psi\|^2_{L(\mu)} - \|U\psi\|^2_{L(\mu)}$.

Note that $-\varphi$ is mean zero and Lipschitz as well, hence in the above theorem we can change the superior limit to the inferior and 1 to -1. Actually, much more may be deduced but it requires introducing more definitions.

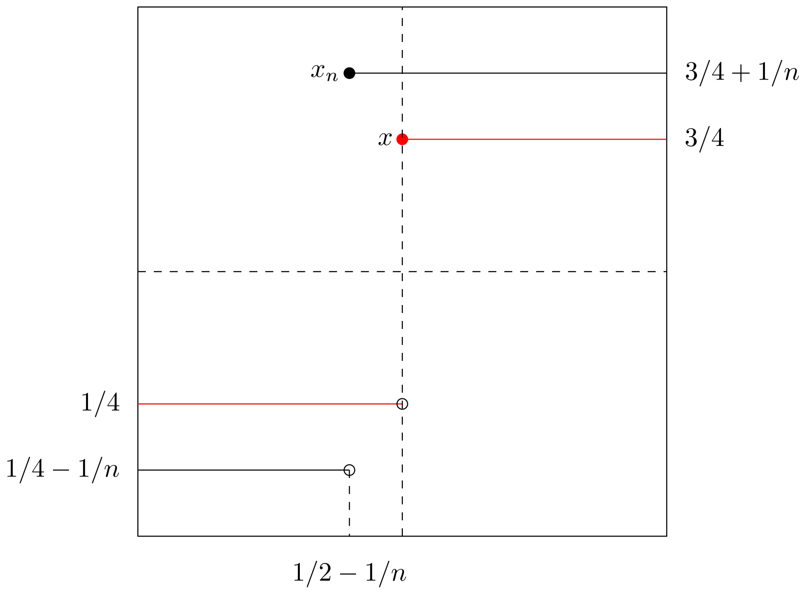

Figure 3.10: The sequence (x_n) as well as x are piecewise constant.

3.7 The space of trajectories. The Skorohod J_1 topology and invariance principles

Let us denote by $D[0,1]$ the space of functions $g : [0,1] \to \mathbb{R}$ that are right-continuous with left limits. It is somehow natural to think of equipping $D[0,1]$ with the topology induced by the supremum norm, however, this topology is too strong. Indeed, the sequence (x_n) depicted in Figure 3.10 is not convergent to x in sense of this topology.

The idea is to slightly modify this topology. For this purpose let us define a time change λ as an increasing bijection of $[0,1]$ onto itself (this in particular implies the continuity, $\lambda(0) = 0$ and $\lambda(1) = 1$). Let us define a topology in $D[0,1]$ with the property that $d(x_n, x) \to 0$ if and only if there exists a sequence of time-changes (λ_n) such that

$$\|\lambda_n(s) - s\|_\infty + \|x_n \circ \lambda_n - x\|_\infty \to 0$$

for $n \to \infty$. The proof of the existence of this topology, called Skorohod J_1 topology, may be found in section 3.5 in [EK86]. Moreover, in this topology the Borel σ-algebra is equivalent to the σ-field generated by the projections π_t, where t is from some arbitrary dense subset T of $[0,1]$.

Now, the Wiener process W is a random variable with values in $D[0,1]$. Similarly the process $t \longmapsto Y_n(t) := \sum_{k \leq nt} \varphi(X_k)$, $t \in [0,1]$. Since the space $D[0,1]$ is equipped with a topology, the weak-$*$ convergence and the convergence in distribution may be defined on it. The invariance principle or the functional central limit theorem states that $Y_n/\sqrt{n}$ converge in distribution to W.

The functional central limit theorem for the sums of i.i.d. random variables Y_n has been proven by Donsker [Don51] and by McLeish [McL74] for martingales with stationary ergodic square integrable increments (version for continuous paths had been proven by Brown [Bro71]). Since our process may be decomposed

$$Y_n(t)/\sqrt{n} = M_n(t)/\sqrt{n} + R_n(t)/\sqrt{n}$$

and $R_n(t)/\sqrt{n}$ tends to zero almost surely (in the space $D[0,1]$), we conclude the invariance principle for additive functional of Markov processes under consideration.

Corollary 3.4. *Let $f_1, \ldots, f_m$ be a system of C^2 diffeomorphisms satisfying (A1) and (A2). Let $(p_1, \ldots, p_m)$ be such that $\Lambda_0, \Lambda_1 > 0$. Let (X_n) be the unique stationary Markov process corresponding to the system. If φ is a Lipschitz real function on $(0,1)$ with $\int \varphi d\mu = 0$, then the sequence of random variables*

$$Y_n(t) := \sum_{k \leq nt} \varphi(X_k)$$

with values in $D[0,1]$ converge in distribution to $\sigma W(t)$ in $D[0,1]$ with Skorohod J_1-topology.

3.8 Comments

Using the functional central limit theorem one can show the arcsine law (see Theorem 12.11 in [Kal02]). The law of the iterated logarithm may be stated in much more general form of Strassen's invariance principle. To my best knowledge there are no techniques which can be applied to prove large deviations for the processes under consideration. Hopefully, Theorems 2.1 and 2.2 may turn out to be helpful in achieving this goal.

It is worth to mention that by Corollary 2.1 the law of the iterated logarithm holds for processes starting from an arbitrary distribution. The same holds for the central limit theorem but the proof requires estimating the difference between characteristic functions of suitable processes. This will be done in Section 4.4 under much weaker assumptions.

As I have mentioned in Section 1.4 all theorems here are towards proving chaotic properties of Kan's diffeomorphisms defined in Section 1.3. It is now interesting task to formulate and prove analogous version of Theorem 3.1 and 3.2 along with Theorems in Sections 3.6, 3.7. Mixing properties of porcupine-like horseshoes ([DG12]) for some special choice of invariant measure may be another direction of further research.

Chapter 4

Ergodic properties of systems of homeomorphisms

4.1 Introduction

We have proven already that when $f_1, \ldots, f_m$ are C^2 diffeomorphisms and the system has positive Lyapunov exponents, then $\|U^n\varphi\|_{L^2(\mu)}$ decays exponentially fast for any Lipschitz function φ with $\int \varphi d\mu = 0$. This implies the central limit theorem, the functional central limit theorem and the law of the iterated logarithm. It is unknown whether arbitrary system of homeomorphisms satisfying (A1), (A2) with positive Lyapunov exponent is exponentially mixing in that sense. Despite of that we are still able to prove the classical limit theorems.

A key ingredient of the proof of the exponential mixing was the Baxendale theorem (stated here as Theorem 2.4), which says that the volume Lyapunov exponents of the system are negative provided that $f_1, \ldots, f_m$ are C^2 diffeomorphisms. In 2014 Dominique Malicet published paper [Mal17], in which he proves some exponential contraction result without any smoothness assumption. It has already been invoked (Theorem 2.5), but since this is a stem of this chapter we state it here once again. To this end we define once again Ω to be $\{1, \ldots, m\}^{\mathbb{N}}$, $\mathcal{F}$ to be the standard product σ-algebra and $\mathbb{P}$ to be the product measure of the probability vector $(p_1, \ldots, p_m)$. Recall that $Z_n^x(\omega) = f_\omega^n(x) = f_{\omega_n} \circ \cdots \circ f_1(x)$ for $\omega = (\omega_1, \omega_2, \ldots)$. This notation is kept in the whole chapter.

Let us recall Corollary 2.1 from Section 2.7.

Corollary (cf. Corollary 2.13 in [Mal17])**.** *If $f_1, \ldots, f_m$ are interval homeomorphisms and $(p_1, \ldots, p_m)$ are such that*

- *there exists no nontrivial subinterval of $(0, 1)$ which is invariant by all f_i's, and*
- *there exists a measure μ with $\mu((0, 1)) = 1$ which is stationary for the random walk,*

then there exist $q < 1$ such that for every $x \in \mathbb{S}^1$ and for almost every $\omega \in \Omega$ there exits an open neighbourhood I of x such that

$$|f_\omega^n(I)| \leq q^n \text{ for every } n \geq 1.$$

Note that both assumptions are satisfied when Λ_0, Λ_1 are positive (the first condition is avtually a consequence of (A1)). This theorem allowed Tomasz Szarek and Anna Zdunik to prove the central limit theorem (see [SZ21]) and the law of the iterated logarithm ([SZ20]) for systems of homeomorphisms of the circle. Later the same has been proven in the case of interval systems by, respectively, myself and Tomasz Szarek [CS20b] and by myself, Hanna Wojewódka-Ściążko and Tomasz Szarek [CWSS20]. Although the general concept of the proof was the same as in the case of systems of the circle (i.e. to determine the rate of convergence of $|U^n\varphi(x) - U^n\varphi(y)|$ to zero using the Malicet theorem), in the case of interval systems it is necessary to deal with the lack of compactness, which is a substantial obstacle.

The main theorem of the present section is to show the central limit theorem and the law of the iterated logarithm for arbitrary system of homeomorphisms with (A1), (A2) and $\Lambda_0, \Lambda_1 > 0$. In the proof we shall use the Maxwell-Woodroofe criterion (Theorem 1 in [MW00]), which in our setting takes the form

$$\sum_{n=1}^{\infty} n^{-\frac{3}{2}} \Big\| \sum_{k=1}^{n} U^k\varphi \Big\|_{L_2(\mu)} < \infty, \tag{4.1}$$

where φ is a Lipschitz function with $\int \varphi d\mu = 0$. If this condition is satisfied and (X_n) is ergodic, then the additive functional $\varphi(X_1) + \cdots + \varphi(X_n)$ satisfies the central limit theorem.

Theorem 4.1. *Let $f_1, \ldots, f_m$ be a system of increasing interval homeomorphisms satisfying (A1) and (A2). Let $(p_1, \ldots, p_m)$ be such that $\Lambda_0, \Lambda_1 > 0$. If φ is a Lipschitz real function with $\int \varphi d\mu = 0$, then*

$$\Big\| \sum_{k=1}^{n} U^k\varphi \Big\|_{L^2(\mu)} \leq C n^{3/8}$$

for some constant $C > 0$.

Corollary 4.1. *Let $f_1, \ldots, f_m$ be a system of increasing interval homeomorphisms satisfying (A1) and (A2). Let $(p_1, \ldots, p_m)$ be such that $\Lambda_0, \Lambda_1 > 0$. Let (X_n) be the unique stationary process corresponding to the system. If φ is a Lipschitz real function with $\int \varphi d\mu = 0$, then the additive functional $(\varphi(X_1) + \cdots + \varphi(X_n))$ satisfies the central limit theorem, i.e. there exists $\sigma \geq 0$ such that*

$$\mathbb{P}\Big(\frac{\varphi(X_1) + \cdots + \varphi(X_n)}{\sqrt{n}} \in [a, b] \Big) \to \frac{1}{\sqrt{2\pi\sigma^2}} \int_a^b e^{-\frac{x^2}{2\sigma^2}} dx$$

as $n \to \infty$.

After proving the central limit theorem for additive functionals of Markov processes under the Maxwell-Woodroofe condition it was natural to ask about other limit theorems. The law of the iterated logarithm was proved in at least two papers roughly at the same time [MY08], [ZW08]. In the second one (Corollary 1 therein) the condition takes the form

$$\sum_{n=1}^{\infty} \Big(\frac{\log(n)}{n} \Big)^{-\frac{3}{2}} \Big\| \sum_{k=1}^{n} U^k\varphi \Big\|_{L_2(\mu)} < \infty.$$

Thus we can conclude the following corollary.

Corollary 4.2. *Let $f_1, \ldots, f_m$ be a system of increasing interval homeomorphisms satisfying (A1) and (A2). Let $(p_1, \ldots, p_m)$ be such that $\Lambda_0, \Lambda_1 > 0$. Let (X_n) be the unique stationary process corresponding to the system. If φ is a Lipschitz real function with $\int \varphi d\mu = 0$, then the additive functional $(\varphi(X_1) + \cdots + \varphi(X_n))$ satisfies the law of the iterated logarithm*

$$\limsup_{n\to\infty} \frac{\varphi(X_1) + \cdots + \varphi(X_n)}{\sqrt{2n\sigma^2 \log\log n}} = 1 \quad a.s.,$$

where $\sigma^2 = \lim_{n\to\infty} \mathbb{E}(\varphi(X_1) + \cdots + \varphi(X_n))^2$.

It has been proven in [CWSS20] that the process in the statement may be in fact taken with arbitrary initial distribution. Here it clearly follows from Corollary 2.1.

The invariance principle under the Maxwell-Woodroofe condition has been proven in [PU05]. Therefore we obtain.

Corollary 4.3. *Let $f_1, \dots, f_m$ be a system of homeomorphisms satisfying (A1) and (A2). Let $(p_1, \dots, p_m)$ be such that $\Lambda_0, \Lambda_1 > 0$. Let (X_n) be the unique stationary process corresponding to the system. If φ is a Lipschitz real function with $\int \varphi d\mu = 0$, then the sequence of random variables*

$$Y_n(t) := \sum_{k \le nt} \varphi(X_k)$$

with values in $D[0,1]$ converge in distribution to $\sigma W(t)$ in $D[0,1]$ with Skorohod J_1-topology, where $\sigma^2 = \lim_{n\to\infty} \mathbb{E}(\varphi(X_1) + \cdots + \varphi(X_n))^2$.

4.2 The proof of Theorem 4.1

The estimation proceeds in the same way as in the proof of Theorem 3.1. Recall here that Proposition 3.1 says that there exist $\alpha \in (0,1)$ and $c < 1$ such that for every $a > 0$ sufficiently small

$$\mathbb{P}\Big(\bigcap_{k=1}^{n} \{Z_k^x < a\}\Big) \le a^\alpha / x^\alpha c^n \quad \text{and} \quad \mathbb{P}\Big(\bigcap_{k=1}^{n} \{Z_k^{1-x} > 1-a\}\Big) \le a^\alpha / x^\alpha c^n \tag{4.2}$$

for $x < a$. Take $a > 0$ such that $\mu([a, 1-a]) > 4/5$ and such that transition from $(0,a)$ to $(1-a,1)$ as well as from $(1-a,1)$ to $(0,a)$ is impossible in one step. Let M, α be the constants given in Proposition 3.1. Define $\xi_n := c^{\lfloor \sqrt[4]{n} \rfloor /2}$, $K_n := [\xi_n, 1-\xi_n] \times [\xi_n, 1-\xi_n]$, $R_n := [0,1]\times[0,1] \setminus K_n$. By Corollary 3.1 if holds that $\mu \otimes \mu(R_n) \le 4M\xi_n^\alpha$.

With this constants we can reproduce the estimations from the proof of exponential decay of correlations:

$$\begin{aligned}
\Big\| \sum_{k=1}^{n} U^k \varphi \Big\|_{L^2(\mu)}^2 &= \int_{[0,1]} \Big| \sum_{k=1}^{n} U^k \varphi(x) \Big|^2 \mu(dx) = \int_{[0,1]} \Big| \sum_{k=1}^{n} \Big(U^k\varphi(x) - \int_{[0,1]} \varphi(y)\mu(dy) \Big) \Big|^2 \mu(dx) \\
&= \int_{[0,1]} \Big| \sum_{k=1}^{n} \Big(U^k\varphi(x) - \int_{[0,1]} \varphi(y) P^k \mu(dy) \Big) \Big|^2 \mu(dx) \\
&= \int_{[0,1]} \Big| \sum_{k=1}^{n} \Big(U^k\varphi(x) - \int_{[0,1]} U^k\varphi(y) \mu(dy) \Big) \Big|^2 \mu(dx) \\
&= \int_{[0,1]} \Big| \int_{[0,1]} \Big(\sum_{k=1}^{n} U^k\varphi(x) - U^k\varphi(y) \Big) \mu(dy) \Big|^2 \mu(dx) \\
&\le \int_{[0,1]} \int_{[0,1]} \Big(\sum_{k=1}^{n} \big| U^k\varphi(x) - U^k\varphi(y) \big| \Big)^2 \mu(dy)\mu(dx) \\
&\le \int \int_{K_n} \Big(\sum_{k=1}^{n} \mathrm{Lip}(\varphi) \mathbb{E}\big| Z_k^x - Z_k^y \big| \Big)^2 \mu(dy)\mu(dx)
\end{aligned}$$

$$+\int\int_{R_n}\left(\sum_{k=1}^n \left|U^k\varphi(x)-U^k\varphi(y)\right|\right)^2\mu(dy)\mu(dx)$$

$$\leq\left(\sum_{k=1}^n \mathrm{Lip}(\varphi)\mathbb{E}\left|Z_k^{\xi_n}-Z_k^{1-\xi_n}\right|\right)^2+\int\int_{R_n}4n^2\|\varphi\|_\infty^2\mu(dy)\mu(dx)$$

$$\leq\left(\sum_{k=1}^n \mathrm{Lip}(\varphi)\mathbb{E}\left|Z_k^{\xi_n}-Z_k^{1-\xi_n}\right|\right)^2+16M\xi_n^\alpha\|\varphi\|_\infty^2 n^2,$$

We are left to estimate

$$\mathbb{E}|Z_k^{\xi_n}-Z_k^{1-\xi_n}|=\int_\Omega\sum_{k=1}^n|f_\omega^k(\xi_n)-f_\omega^k(1-\xi_n)|\mathrm{d}\mathbb{P}.$$

The idea is to divide, for n fixed, Ω into three parts D_n, E_n, H_n. The probability of the events D_n and E_n shall diminish sufficiently fast as n increases, while on H_n some bound may be found for the value of $\sum_{k=1}^n|f_\omega^k(\xi_n)-f_\omega^k(1-\xi_n)|$.

Let

$$D_n:=\{\omega\in\Omega:\exists_{k\geq\lfloor\sqrt[4]{n}\rfloor}f_\omega^k(\xi_n)<\xi_n\quad\text{or}\quad f_\omega^k(1-\xi_n)>1-\xi_n\}.$$

Lemma 4.1. *There exists a constant C_1 such that $\mathbb{P}(D_n)\leq C_1nc^{\lfloor\sqrt[4]{n}\rfloor/2}$.*

Proof. Let τ be the moment of the first visit of $(Z_j^{\xi_n})_j$ in $[a,1-a]$. Recall that M was chosen so that $\nu\in\mathcal{P}_{M,\alpha}$ for every measure ν supported on $[a,1-a]$. This means that if $x\in[a,1-a]$, then $P^n\delta_x((0,\xi_n)\cup(1-\xi_n,1))\leq 2M\xi_n^\alpha$. Thus by the strong Markov property

$$\mathbb{P}\left(\{Z_k^{\xi_n}\notin[\xi_n,1-\xi_n]\}\cap\{\tau\leq\lfloor\sqrt[4]{n}\rfloor\}\;\middle|\;\mathcal{F}_\tau\right)\leq 2M\xi_n^\alpha\quad\text{a.s.}$$

for $k\geq\lfloor\sqrt[4]{n}\rfloor$. Therefore, by (4.2),

$$\mathbb{P}\left(\bigcup_{k=\lfloor\sqrt[4]{n}\rfloor}^n\{Z_k^{\xi_n}\notin[\xi_n,1-\xi_n]\}\right)\leq\sum_{k=\lfloor\sqrt[4]{n}\rfloor}^n\mathbb{E}\mathbb{P}\left(\{Z_k^{\xi_n}\notin[\xi_n,1-\xi_n]\}\cap\{\tau\leq\lfloor\sqrt[4]{n}\rfloor\}\;\middle|\;\mathcal{F}_\tau\right)$$

$$+\mathbb{P}(\tau>\lfloor\sqrt[4]{n}\rfloor)\leq 2nM\xi_n^\alpha+c^{\lfloor\sqrt[4]{n}\rfloor}a^\alpha/\xi_n^\alpha=2nMc^{\lfloor\sqrt[4]{n}\rfloor/2}+a^\alpha c^{\lfloor\sqrt[4]{n}\rfloor}c^{-\lfloor\sqrt[4]{n}\rfloor/2}\leq C_1nc^{\lfloor\sqrt[4]{n}\rfloor/2}$$

for some constant C_1. What is left is to proceed with the analogous computation for $1-\xi_n$ and possibly amend the choice of C_1. $\square$

The task of defining the sets E_n, H_n will be preceded with the construction of a measurable set $B\subseteq\Omega$ (dependent of n) of positive $\mathbb{P}$-measure (with a positive lower bound independent of n) and a constant C_2 (also independent of n) such that

$$\sum_{j=1}^n|f_j([\xi_n,1-\xi_n])|\leq\lfloor\sqrt[4]{n}\rfloor+C$$

for every $\omega\in B$ and n sufficiently large.

To this end we fix $x_0\in\mathrm{supp}(\mu)$, interval I containing x_0 and a measurable subset $B'''\subseteq\Omega$ such that $\mathbb{P}(B''')>1/2$ and $|f_\omega^n(I)|\leq q^n$ for every $n\geq 1$ and $\omega\in B'''$ (the existence follows from Corollary 2.1 recalled in Section 4.1). Now we shall briefly show that there exist r_1 and $\beta>0$ such that

$$\mathbb{P}(f_\omega^{r_1}([a,1-a])\subseteq I)\geq\beta.$$

Indeed, we know that the distribution of $\omega \to x_\omega$ is μ (the definition in Section 2.7). Since $x_0 \in \operatorname{supp}(\mu)$, we can find $\widetilde{I}$ whose closure is contained in I and $\mu(\widetilde{I}) > 0$. Taking $\varepsilon > 0$ sufficiently small we know that

$$\mathbb{P}(f_\omega^{r_1}([a, 1-a]) \subseteq I) \geq \mathbb{P}(x_\omega \in \widetilde{I} \quad \text{and} \quad |f_\omega^{r_1}([a, 1-a])| < \varepsilon) > \beta > 0.$$

Denote $B'' := \{\omega \in \{1, \ldots, m\}^{r_1} \ : \ f_\omega^{r_1}([a, 1-a]) \subseteq I\}$ (hence B'' is a set of finite sequences).

We have defined the set B''' of infinite sequences and the set B'' of finite sequences of length r_1 (independent of n). Now we are going to define a set of finite sequences of length depending on n, which we shall denote by B'. Take an integer r_2 such that

$$\mathbb{P}(f_\omega^i(x) \in [a, 1-a]) > \frac{4}{5} \tag{4.3}$$

for every $x \in [a, 1-a]$ and $i \geq r_2$. This is a consequence of the stability of the process (Theorem 2.2), Portmanteau theorem (Theorem 2.1 in [Bil99]) and the fact that $\mu([a, 1-a]) > 4/5$. By inequalities (4.2) we have

$$\mathbb{P}\Big(\bigcap_{k=1}^{\lfloor \sqrt[4]{n} \rfloor/2} \{f_\omega^k(\xi_n) < a\} \Big) \leq a^\alpha / \xi_n^\alpha c^{\lfloor \sqrt[4]{n} \rfloor/2} = a^\alpha \big(c^{(1-\alpha)/2}\big)^{\lfloor \sqrt[4]{n} \rfloor} \quad \text{and}$$

$$\mathbb{P}\Big(\bigcap_{k=1}^{\lfloor \sqrt[4]{n} \rfloor/2} \{f_\omega^k(1-\xi_n) > 1-a\} \Big) \leq a^\alpha \big(c^{(1-\alpha)/2}\big)^{\lfloor \sqrt[4]{n} \rfloor}$$

for every n. If n was sufficiently large, then

- the probability of each of the above sets is less than 1/4 and
- $\lfloor \sqrt[4]{n} \rfloor/2 > r_2$.

By this and (4.3) the probability of $\{f_\omega^{\lfloor \sqrt[4]{n} \rfloor}(\xi_n) \in [a, 1-a]\}$ is grater than $4/5 \cdot 3/4 = 3/5$. Similarly the probability of $\{f_\omega^{\lfloor \sqrt[4]{n} \rfloor}(1-\xi_n) \in [a, 1-a]\}$ is greater than 3/5, therefore the probability of $B' := \{f_\omega^{\lfloor \sqrt[4]{n} \rfloor}(\xi_n) \in [a, 1-a]$ and $f_\omega^{\lfloor \sqrt[4]{n} \rfloor}(1-\xi_n) \in [a, 1-a]\} \subseteq \{1, \ldots, m\}^{\lfloor \sqrt[4]{n} \rfloor}\}$ is greater than 1/5 (Figure 4.1). The set $B := B' \times B'' \times B''' \subseteq \Omega$ has measure $\mathbb{P}(B) = \mathbb{P}(B') \cdot \mathbb{P}(B'') \cdot \mathbb{P}(B''') \geq 1/5 \cdot \beta \cdot 1/2 = \beta/10 > 0$[1]. Moreover, for $\omega \in B$ we have

$$\sum_{j=0}^{n} |f_\omega^j([\xi_n, 1-\xi_n])| \leq \lfloor \sqrt[4]{n} \rfloor + r, \tag{4.4}$$

where r is independent of n.

Fix n sufficiently large. By abuse of notation the projection of a subset of Ω to $\Omega_n := \{1, \ldots, m\}^n$ will be denoted by the same letter as the subset (this rule will be applied especially to B). In the same way we shall denote by $\mathbb{P}$ the product measure of $(p_1, \ldots, p_m)$ on Ω_n. Define $A_0 := \Omega_n$ and $B_1 := B$. The set $A_1 := \Omega_n \setminus B_1$ is a sum of disjoint cylinders, hence there exists a set $F_1 \subseteq \Omega_* = \bigcup_{j=1}^\infty \Omega_j$ of finite sequences such that $A_1 = \bigcup_{(i_1, \ldots, i_k) \in F_1} C_{(i_1, \ldots, i_k)}$, where $C_{(i_1, \ldots, i_k)} = \{\omega \in \Omega : \omega_1 = i_1, \ldots, \omega_k = i_k\}$. This representation is generally not unique, but if we insist F_1 to have the smallest possible cardinality, then it becomes unique. Let

$$B_2 := \bigcup_{(i_1, \ldots, i_k) \in F_1} \{(i_1, \ldots, i_k)\} \times B \subseteq \Omega_n.$$

[1]By abuse of notation we write $\mathbb{P}$ to denote the product measure of $(p_1, \ldots, p_m)$ on a finite product of $\{1, \ldots, m\}$.

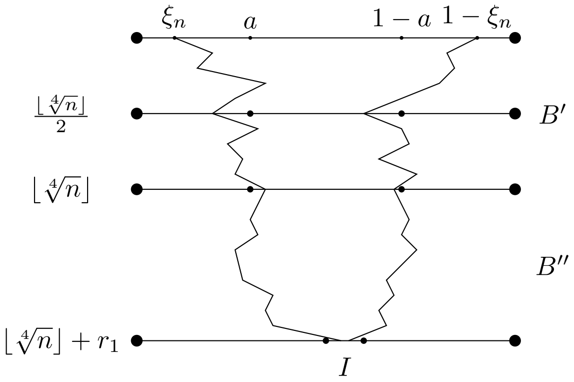

Figure 4.1

and $A_2 := \Omega_n \setminus (B_1 \cup B_2)$. Continue this procedure. For every l the set $A_l := \Omega \setminus (B_1 \cup \ldots \cup B_l)$ is a sum of some cylinders $C_{(i_1,\ldots,i_k)}$, $(i_1, \ldots, i_k) \in F_l$. Then define $B_{l+1} := \bigcup_{(i_1,\ldots,i_k) \in F_l} C_{(i_1,\ldots,i_k)} \times B$. Proceed until $l = \lfloor \sqrt[8]{n} \rfloor$.

We have already observed that

$$\sum_{j=0}^{n} |f_\omega^j([\xi_n, 1-\xi_n])| \leq \lfloor \sqrt[4]{n} \rfloor + r,$$

where r is independent of n and $\omega \in B_1$. We are going to show that

$$\sum_{j=0}^{n} |f_\omega^j([\xi_n, 1-\xi_n])| \leq l(\lfloor \sqrt[4]{n} \rfloor + r) + l,$$

where $\omega \in B_l \setminus D_n$ (recall that r is independent of n).

If $\omega \in B_2 \setminus D_n$, then $\omega \notin B_1$ and we can find $(i_1, \ldots, i_k) \in F_1$ such that $\omega \in C_{i_1,\ldots,i_k}$. This means that $\omega_1 = i_1, \ldots, \omega_k = i_k$ and $(i_1 \ldots, i_{k-1}) \in B$ (by which we mean that $(i_1, \ldots, i_{k-1})$ agrees with certain sequence from B on the first $k-1$ coordinates). The fact that $\omega \in B_2$ implies $(\omega_{k+1}, \ldots, \omega_n) \in B$ thus the application of (4.4) to $(\omega_1, \ldots, \omega_{k-1})$ and $(\omega_{k+1}, \omega_{k+2}, \ldots)$ gives

$$\sum_{j=1}^{n} |f_\omega^j([\xi_n, 1-\xi_n])| \leq \lfloor \sqrt[4]{n} \rfloor + r + 1 + \lfloor \sqrt[4]{n} \rfloor + r \leq 2(\lfloor \sqrt[4]{n} \rfloor + r + 2),$$

which completes the proof for $l = 2$.

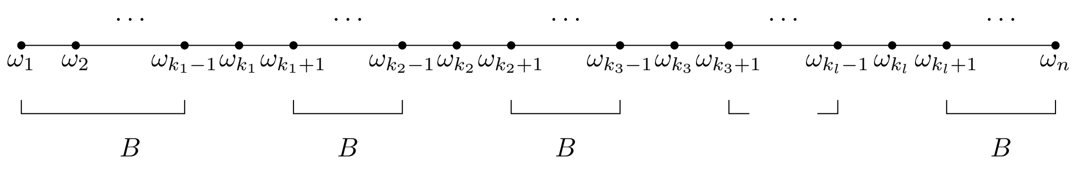

Figure 4.2: The decomposition of ω.

If $\omega \in B_3 \setminus D_n$, then $\omega \notin B_2$, and we can find $(i_1, \ldots, i_k) \in F_2$ such that $\omega \in C$. This means that $\omega_1 = i_1, \ldots, \omega_k = i_k$ and there exists k' with $(\omega_1, \ldots, \omega_{k'-1}) \in B$ and $(\omega_{k'+1}, \ldots, \omega_{k-1}) \in B$. Since $\omega \in B_3$, we know that $(\omega_{k+1}, \ldots, \omega_n) \in B$. Concluding,

$$\sum_{j=1}^{n} |f_\omega^j([\xi_n, 1-\xi_n])| \leq 3(\lfloor \sqrt[4]{n} \rfloor + r + 3),$$

which completes the proof for $l = 3$. We continue in this fashion. For every l a sequence $\omega \in B_l$ may be decomposed into l sequences (possibly one or more empty) from B with one step break between each neighbouring pair (see Figure 4.2). Therefore

$$\sum_{j=1}^{n} |f_\omega^j([\xi_n, 1-\xi_n])| \leq l(\lfloor \sqrt[4]{n} \rfloor + r + l),$$

for every $l = 1, \ldots, \lfloor \sqrt[8]{n} \rfloor$ (in fact one could replace the l in parenthesis by $l-1$). Eventually we have proved that

$$\sum_{j=1}^{n} |f_\omega^j([\xi_n, 1-\xi_n])| \leq \lfloor \sqrt[8]{n} \rfloor (\lfloor \sqrt[4]{n} \rfloor + r + \lfloor \sqrt[8]{n} \rfloor) \tag{4.5}$$

for $\omega \in B_1 \cup \cdots \cup B_{\lfloor \sqrt[8]{n} \rfloor} \setminus D_n$.

Put $H_n := B_1 \cup \cdots \cup B_{\lfloor \sqrt[8]{n} \rfloor} \setminus D_n$ and $E_n := A_{\lfloor \sqrt[8]{n} \rfloor}$.

Lemma 4.2. *We have*

$$\mathbb{P}(E_n) \leq (1 - \mathbb{P}(B))^{\lfloor \sqrt[8]{n} \rfloor}.$$

Proof. Recall we identify $\mathbb{P}$ with its projection on the first n coordinates. Therefore

$$\begin{aligned}\mathbb{P}\big(A_{\lfloor \sqrt[8]{n} \rfloor}\big) &= \mathbb{P}\big(A_{\lfloor \sqrt[8]{n} \rfloor - 1} \cap (\Omega \setminus B_{\lfloor \sqrt[8]{n} \rfloor})\big) = \mathbb{P}\big(\Omega \setminus B_{\lfloor \sqrt[8]{n} \rfloor} \bigm| A_{\lfloor \sqrt[8]{n} \rfloor - 1}\big) \mathbb{P}(A_{\lfloor \sqrt[8]{n} \rfloor - 1}) \\ &\leq (1 - \mathbb{P}(B)) \mathbb{P}(A_{\lfloor \sqrt[8]{n} \rfloor - 1}) \leq \ldots \leq (1 - \mathbb{P}(B))^{\lfloor \sqrt[8]{n} \rfloor}.\end{aligned}$$

□

We are in position to complete the proof. We have

$$\int_\Omega \sum_{k=1}^{n} |f_\omega^k(\xi_n) - f_\omega^k(1-\xi_n)| \mathrm{d}\mathbb{P} \leq \int_{D_n} \sum_{k=1}^{n} |f_\omega^k(\xi_n) - f_\omega^k(1-\xi_n)| \mathrm{d}\mathbb{P} + \int_{H_n} \sum_{k=1}^{n} |f_\omega^k(\xi_n) - f_\omega^k(1-\xi_n)| \mathrm{d}\mathbb{P}$$

$$+ \int_{E_n} \sum_{k=1}^{n} |f_\omega^k(\xi_n) - f_\omega^k(1-\xi_n)| \mathrm{d}\mathbb{P} \leq n\mathbb{P}(D_n) + n\mathbb{P}(E_n) + \lfloor \sqrt[8]{n} \rfloor (\lfloor \sqrt[4]{n} \rfloor + r + \lfloor \sqrt[8]{n} \rfloor).$$

By Lemmas 4.1 and 4.2 two first sequences are bounded. Therefore taking n sufficiently large, say $n \geq n_0$, yields

$$\int_\Omega \sum_{k=1}^{n} |f_\omega^k(\xi_n) - f_\omega^k(1-\xi_n)| \mathrm{d}\mathbb{P} \leq C n^{3/8} \tag{4.6}$$

for some constant C (independent of n). Finally

$$\Big\| \sum_{k=1}^{n} U^k \varphi \Big\|_{L^2(\mu)}^2 \leq \big(\mathrm{Lip}(\varphi) C n^{3/8}\big)^2 + 16 M \xi_n^\alpha \|\varphi\|_\infty^2 n^2$$

for $n \geq n_0$. The second term is again bounded, hence possibly changing n_0 and C we have

$$\left\| \sum_{k=1}^{n} U^k \varphi \right\|_{L^2(\mu)}^2 \leq \left(Cn^{3/8}\right)^2,$$

which proves the assertion (notice that the norm above is squared).

4.3 The value of σ

The approximation method by Maxwell and Woodroofe does not provide any information about the value of σ. In general it is known only that

$$\sigma^2 = \lim_{n\to\infty} \mathbb{E}(\varphi(X_1) + \cdots + \varphi(X_n))^2$$

(it is proven in [MW00] that the limit exists) and $\sigma < \infty$. The exact value is not so much important, however it should be determined whether it is positive or not. If not, then

$$\lim_{n\to\infty} \mathbb{E}(\varphi(X_1) + \cdots + \varphi(X_n))^2 = 0,$$

and in fact the central limit theorem does not hold. Here we are not able to show in general that σ is positive but there are situations in which the problem is possible to handle with.

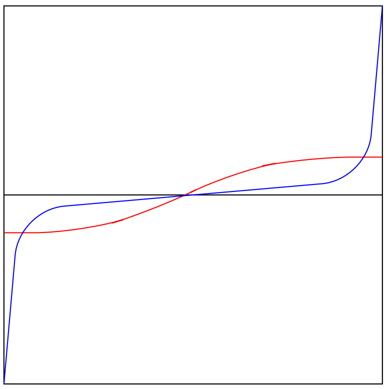

Figure 4.3: The plot of φ (red) and of $U^n\varphi$ (blue) for n large.

Let f_1, f_2 be a system satisfying assumptions of Theorem 2.5 with probability vector $(1/2, 1/2)$. Let us assume also that it is symmetric, thus $f_1(x) = 1 - f_2(1-x)$. Let φ be increasing, Lipschitz and symmetric with respect to the point $(1/2, 0)$ (see Figure 4.2) (this implies that the mean value is zero as for such system the stationary measure μ must be symmetric as well). Observe that for such observable $U\varphi$ is increasing too. Indeed, f_i's are increasing hence if $x_1 < x_2$, then $U\varphi(x_1) = \mathbb{E}\varphi(f_\omega(x_1)) \leq \mathbb{E}\varphi(f_\omega(x_2)) = U\varphi(x_2)$. Moreover, $U\varphi(1/2) = 0$ from the symmetry of the system. Both these facts combined yields that $U\varphi(x) < 0$ for $x < 1/2$ and $U\varphi(x) > 0$ for $x > 1/2$.

Clearly it holds also for all iterates $U^n\varphi$. If (X_n) is the stationary process corresponding to the system defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then

$$\frac{1}{n}\mathbb{E}(\varphi(X_1)+\cdots+\varphi(X_n))^2 = \frac{1}{n}\Big(\mathbb{E}\varphi(X_1)^2+\cdots+\mathbb{E}\varphi(X_n)^2\Big) + \frac{2}{n}\sum_{i=1}^{n-1}\mathbb{E}\varphi(X_i)\big(\varphi(X_{i+1})+\cdots+\varphi(X_n)\big)$$

$$= \int_{[0,1]} \varphi^2(x)\mu(dx) + \frac{2}{n}\sum_{i=1}^{n-1}\mathbb{E}\Big(\varphi(X_i)\mathbb{E}\big(\varphi(X_{i+1}) + \cdots + \varphi(X_n) \mid X_i\big)\Big)$$

$$= \int_{[0,1]} \varphi^2(x)\mu(dx) + \frac{2}{n}\sum_{i=1}^{n-1} \mathbb{E}\Big(\varphi(X_i)\big(U\varphi(X_i) + \cdots + U^{n-i}\varphi(X_i)\big)\Big).$$

By the considerations we have just made, the function under the integral in the second summand is positive, thus the same is true for the expectation. We conclude that

$$\frac{1}{n}\mathbb{E}(\varphi(X_1) + \cdots + \varphi(X_n))^2 \geq \int_{[0,1]} \varphi(x)^2\mu(dx) > 0$$

provided φ is not the constant function equal to zero.

4.4 Non-stationary processes

It is much more difficult to show that the central limit theorem holds for processes starting from a point. Although it has been proven that a condition slightly stronger than the Maxwell-Woodroofe condition implies the central limit theorem for the process (Z_n^x) starting from the point x for μ-almost every x (see [DL03]), our estimations give even that the central limit theorem holds for a process starting from completely arbitrary distribution.

Fix $x \in (0,1)$. The idea is to show that the characteristic functions of the random variables

$$\frac{1}{\sqrt{n}}\big(\varphi(X_1) + \cdots + \varphi(X_n)\big)$$

and

$$\frac{1}{\sqrt{n}}\big(\varphi(Z_1^x) + \cdots + \varphi(Z_n^x)\big)$$

converge pointwise to each other. Since the characteristic function of the first tends pointwise to $\exp(\frac{-t^2\sigma^2}{2})$ this would give the assertion[2].

We have

$$\begin{aligned}
\Bigg| \int_{[0,1]}\int_\Omega & \exp\left(it\frac{\varphi(y) + \cdots + \varphi(f_\omega^{n-1}(y))}{\sqrt{n}}\right)\mathbb{P}(d\omega)\mu(dy) \\
& - \int_\Omega \exp\left(it\frac{\varphi(x) + \cdots + \varphi(f_\omega^{n-1}(y))}{\sqrt{n}}\mathbb{P}(d\omega)\right)\Bigg| \\
& \leq \int_{[0,1]}\int_\Omega \Bigg|\exp\left(it\frac{\varphi(y) + \cdots + \varphi(f_\omega^{n-1}(y))}{\sqrt{n}}\right) \\
& - \exp\left(it\frac{\varphi(x) + \cdots + \varphi(f_\omega^{n-1}(y))}{\sqrt{n}}\right)\Bigg|\mathbb{P}(d\omega)\mu(dy) \\
\leq \frac{|t|}{\sqrt{n}} & \int_{[0,1]}\int_\Omega \Bigg|\sum_{i=0}^{n-1}\varphi(f_\omega^i(y)) - \varphi(f_\omega^i(x))\Bigg|\mathbb{P}(d\omega)\mu(dy)
\end{aligned}$$

by the inequality $|e^{itx_1} - e^{itx_2}| \leq |t||x_1 - x_2|$. It does not exceed

$$\frac{\mathrm{Lip}(\varphi)|t|}{\sqrt{n}}\int_{[0,1]}\int_\Omega \sum_{i=0}^{n-1}\Big|f_\omega^i(y) - f_\omega^i(x)\Big|\mathbb{P}(d\omega)\mu(dy).$$

[2] Recall a standard theorem that the sequence of measures is weak-$*$ convergent to some limit measure if and only if the characteristic functions of this measures converge pointwise to the characteristic function of the limit measure.

Similarly as in Section 4.2 the integral may be decomposed into two, one bounded from above by $\frac{\mathrm{Lip}(\varphi)|t|}{\sqrt{n}}\mu((0,\xi_n)\cup(1-\xi_n,1))\cdot n \le 2\mathrm{Lip}(\varphi)|t|\sqrt{n}M\xi_n^\alpha$, which decays to zero, and the second bounded by

$$\frac{\mathrm{Lip}(\varphi)|t|}{\sqrt{n}}\int_\Omega \sum_{k=1}^n |f_\omega^k(\xi_n) - f_\omega^k(1-\xi_n)|\mathrm{d}\mathbb{P}.$$

By (4.6) we know that this is growing not faster than $\frac{\mathrm{Lip}(\varphi)|t|}{\sqrt{n}}Cn^{3/8}$, which also tends to 0 as n goes to infinity. This completes the proof.

4.5 Comments

It is doubtful that Corollary 4.3 cannot be extended to processes with an arbitrary initial distribution. However, the proof would require two facts. The first that if (X_n) starts from some point $x \in (0,1)$, then $(Y_n(t))$ is a tight family in $D[0,1]$. The second is that the finite-dimensional distributions of Y_n tend to the corresponding finite dimensional distributions of the Wiener process. The second is an adaptation of our method in Section 4.4, but the first seems to be troublesome. Nevertheless I still conjecture the statement is true.

Apparently any consequence of the exponential decay of correlations in Chapter 3 has been proven for much general class of systems of homeomorphisms with the rate of convergence much weaker than exponential. The question arise whether it is worth to establish the exponential rate of convergence. I hope it may help to show some large deviations results, however to my best knowledge there is no literature and new methods should be elaborated. It seems interesting to find some large deviation results under the Maxwell-Woodroofe type conditions.

Chapter 5

Ergodic properties of systems with place-dependent probabilities

5.1 The main theorems and notation

Systems of homeomorphisms with place-dependent probabilities arise when some specific g-measures of Kan's diffeomorphisms are considered. Usually one wish to know whether g-measure is unique. In this chapter we give a partial answer to that question. Due to huge difficulties the uniqueness is proved for a very specific choice of transformations.

Let $f_1, \ldots, f_m$ be arbitrary increasing interval homeomorphisms, and let $p_1, \ldots, p_m$ be real positive functions on $[0,1]$ with $\sum_{i=1}^m p_i(x) = 1$ for $x \in [0,1]$ (one can think of continuous functions). One can define an analogous process to the processes defined in Chapter 2 with the only difference that the distribution according to which a homeomorphism is chosen depends now on the position in the interval and is given by $(p_1(x), \ldots, p_m(x))$ provided the current position is x.

The process is a Markov process with transition probabilities

$$p(x, \cdot) = \sum_{i=1}^{m} p_i(x)\delta_{f_i(x)}. \tag{5.1}$$

Its Markov and dual operators P and U are given by

$$P\mu(A) = \sum_{i=1}^{m} \int_{f_i^{-1}(A)} p_i(x)\mu(dx),$$

where $\mu \in \mathcal{M}$ (the space of Borel probability measures), A is a Borel set, and

$$U\varphi(x) = \sum_{i=1}^{m} p_i(x)\varphi(f_i(x)),$$

where φ is an arbitrary bounded Borel measurable real function. The operator P is Feller provided p_i's are continuous (we have assumed that f_i's are homeomorphisms). The notions of ergodicity and stability applies here as well.

Now let us turn to the very specific system. Fix $a < b < 1/2$. Let f_2 be an interval homeomorphism mapping $(0,a]$ linearly onto $(0,b]$ and $[a,1)$ onto $[b,1]$. Its graph consists of two straight lines, the first one connecting $(0,0)$ with the point (a,b) and the second one connecting (a,b) with

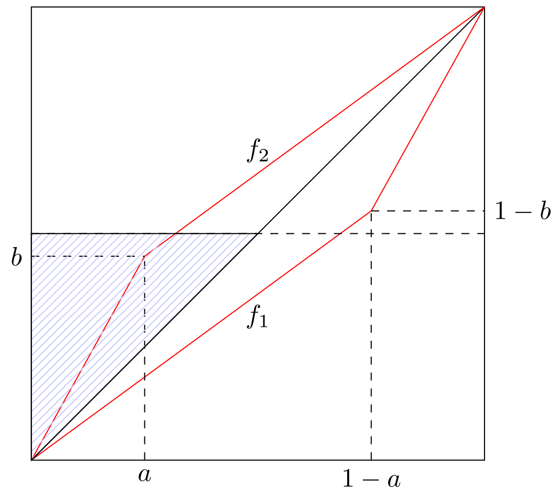

Figure 5.1: An example of Alsedà-Misiurewicz system. The interior of the hatched area is the set of points (a, b) satisfying assumption (A1).

$(1, 1)$. Next, let f_1 be the interval homeomorphism defined by $f_1(x) = 1 - f_2(1 - x)$, $x \in [0, 1]$ (see Figure 5.1). Setting $a_2 = \frac{b}{a}$ and $a_1 = \frac{1-b}{1-a}$, we can write

$$f_2(x) := \begin{cases} a_2 x & \text{if } x \leq a \\ a_1(x-1)+1 & \text{if } x > a \end{cases} \qquad \text{and} \qquad f_1(x) := 1 - f_2(1-x). \tag{5.2}$$

In [BS21] these systems were called the Alsedà-Misiurewicz systems (with the difference that the only restriction for $(a, b) \in (0, 1) \times (0, 1)$ is that it should be above diagonal). Further, fix two positive real functions p_1, p_2 on $[0, 1]$ with $p_1(x) + p_2(x) = 1$ for every $x \in [0, 1]$. It defines an iterated function system with probabilities and, together with some initial distribution μ, a Markov process.

In the previous chapters the notion of the average Lyapunov exponents appeared. In this setting it is defined by

$$\begin{aligned} \Lambda_0 &:= p_1(0)\log(a_1) + p_2(0)\log(a_2), \\ \Lambda_1 &:= p_1(1)\log(a_2) + p_2(1)\log(a_1). \end{aligned} \tag{5.3}$$

It is not hard to adapt the proofs in Sections 2.2, 2.3 to show an analogous results for the system above provided p_1, p_2 do not vanish. Actually, it may be proven for arbitrary systems with arbitrary finite number of homeomorphisms assuming only that the probabilities do not vanish and conditions (A1), (A2) are satisfied. However, dealing with the case $\Lambda_0, \Lambda_1 > 0$ is nontrivial. Even uniqueness of a stationary distribution of a system is hard to prove, therefore the considerations are restricted to the specific system above. Moreover, we need the following assumptions:

(B1) $0 < a < 1/2$ and $a < b < 1/2$, (see Figure 5.1)

(B2) p_1, p_2 are Dini continuous (see the definition below),

(B3) $0 < p_i(x) < 1$ for $x \in [0, 1]$ and $i = 1, 2$,

(B4) $\Lambda_0, \Lambda_1 > 0$.

The functions p_1, p_2 are Dini continuous if for every $C \geq 0$ and $t < 1$ we have $\sum_n \beta(Ct^n) < \infty$ where β denotes the modulus of continuity of p_1, p_2, i.e.

$$\beta(t) := \max_{i=1,2} \sup_{x\in(0,1),|h|\leq t} |p_i(x) - p_i(x+h)|.$$

Our main results is the following theorem.

Theorem 5.1. *Let f_1, f_2 be given by (5.2), and let p_1, p_2 be arbitrary positive continuous functions with $p_1(x) + p_2(x) = 1$, $x \in (0,1)$. If (B1)-(B4) hold, then there exists a unique Borel probability measure $\mu \in \mathcal{M}$ such that the Markov process (X_n^μ) with the family of transition probabilities (5.1) and the initial distribution μ is stationary. Moreover, the process (X_n^ν) starting from an arbitrary measure ν with the family of transition probabilities (5.1) is asymptotically stable (i.e. the law of X_n^ν converges in the weak-$*$ topology to μ).*

It may be proven also that if $x \in (0,1)$ and $\varphi \in C\big((0,1)\big)$, then

$$\frac{\varphi(X_1^x) + \cdots + \varphi(X_n^x)}{n} \to \int \varphi \mathrm{d}\mu \qquad \text{a.s.},$$

where (X_n^x) denotes the process starting from the point x. This holds under the assumptions of Theorem 5.1, and the proof may be found in [Czu20] (Theorem 3 therein).

5.2 Auxiliary results and the existence of a stationary distribution

Recall that

$$\mathcal{P}_{M,\alpha} = \big\{\mu \in \mathcal{M}\big((0,1)\big) : \forall_{x\in(0,1)} \mu\big((0,x]\big) \leq Mx^\alpha \text{ and } \mu\big([1-x,1)\big) \leq Mx^\alpha \ \big\}.$$

We shall prove the existence of parameters M and α for which the class $\mathcal{P}_{M,\alpha}$ is P invariant provided Λ_0, Λ_1 are positive. This easily implies the existence of a stationary distribution as presented in Proposition 3.1. Recall the idea is to apply the Krylov-Bogoliubov technique, i.e. define $\nu_n = \frac{1}{n}(\delta_{1/2} + \cdots + P^{n-1}\delta_{1/2})$. Obviously $\delta_{1/2} \in \mathcal{P}_{M,\alpha}$, thus by the P-invariance of $\mathcal{P}_{M,\alpha}$ all ν_n's are in $\mathcal{P}_{M,\alpha}$, and by weak-$*$ compactness of $\mathcal{P}_{M,\alpha}$ there exists an accumulation point $\mu \in \mathcal{P}_{M,\alpha}$ of this sequence, which is a stationary measure. We omit the details.

Now we show the invariance of $\mathcal{P}_{M,\alpha}$ for suitably chosen parameters. By the continuity of p_1, p_2 and (B4) one can find $\xi \in (0,a)$ such that

$$\begin{aligned} \max_{t\leq a_1^{-1}\xi} p_1(t)\log a_1 + \max_{t\leq a_1^{-1}\xi} p_2(t)\log a_2 > \frac{\Lambda_0}{2}, \\ \max_{t\leq a_1^{-1}\xi} p_1(1-t)\log a_2 + \max_{t\leq a_1^{-1}\xi} p_2(1-t)\log a_1 > \frac{\Lambda_1}{2}. \end{aligned} \tag{5.4}$$

Writing the Taylor formula of the function $\alpha \longmapsto a^{-\alpha}$ at 0 we obtain $a^{-\alpha} = 1 - \alpha\log a + o(\alpha)$, where a is a fixed positive number. By this formula one can find $\alpha \in (0,1)$ and $c \in (0,1)$ with

$$\begin{aligned} \max_{t\leq a_1^{-1}\xi} p_1(t)a_1^{-\alpha} + \max_{t\leq a_1^{-1}\xi} p_2(t)a_2^{-\alpha} < c, \\ \max_{t\leq a_1^{-1}\xi} p_1(1-t)a_2^{-\alpha} + \max_{t\leq a_1^{-1}\xi} p_2(1-t)a_1^{-\alpha} < c. \end{aligned} \tag{5.5}$$

Eventually, take M so that $M\xi^\alpha = 1$ (this implies that $\nu \in \mathcal{P}_{M,\alpha}$ for ν supported on $[\xi, 1-\xi]$).

Take $\mu \in \mathcal{P}_{M,\alpha}$ and $x \in (0,1)$. If $x \geq \xi$, then $Mx^\alpha \geq M\xi^\alpha = 1$, hence the condition $P\mu((0,x]) \leq Mx^\alpha$ is trivially satisfied. If $x < \xi$, then also $x < a$ and (note that $a_2^{-1}x \leq x \leq a_1^{-1}x$)

$$P\mu((0,x)) = \int_{(0,a_1^{-1}x]} p_1(t)\mu(dt) + \int_{(0,a_2^{-1}x]} p_2(t)\mu(dt) \leq \max_{t \leq a_1^{-1}\xi} p_1(t)\mu((0,a_1^{-1}x])$$

$$+ \max_{t \leq a_2^{-1}\xi} p_2(t)\mu((0,a_2^{-1}x]) \leq \max_{t \leq a_1^{-1}\xi} p_1(t)Ma_1^{-\alpha}x^\alpha + \max_{t \leq a_1^{-1}\xi} p_2(t)Ma_2^{-\alpha}x^\alpha = Mx^\alpha c < Mx^\alpha,$$

where in the last line we used (5.5). Therefore $P\mu((0,x]) \leq Mx^\alpha$. The proof that $P\mu([1-x,1)) \leq Mx^\alpha$ is analogous, hence the invariance of $\mathcal{P}_{M,\alpha}$ is established.

Proposition 5.1. *Let f_1, f_2 be given by (5.2), and let p_1, p_2 be arbitrary positive continuous functions with $p_1(x) + p_2(x) = 1$, $x \in (0,1)$. If (B1)-(B4) hold, then there exists $\alpha \in (0,1)$ such that for every $\xi > 0$ sufficiently small there exist M such that $\mathcal{P}_{M,\alpha}$ is P invariant and every measure supported on $[\xi, 1-\xi]$ belongs to this class. If (X_n) is a Markov process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ starting from x and with transition probabilities (5.1) for the specific system above, then*

$$\mathbb{P}\Big(\bigcap_{i=0}^{n}\{X_i < \xi\}\Big) \leq (\xi/x)^\alpha c^n$$

for all $n \geq 0$ and $x < \xi$, and

$$\mathbb{P}\Big(\bigcap_{i=0}^{n}\{X_i > 1-\xi\}\Big) \leq \xi^\alpha/(1-x)^\alpha c^n$$

for all $n \geq 0$ and $x > 1-\xi$.

Proof. The first part has already been proven. We present the proof of the first inequality. For positive integer n put

$$C_n := \{X_0 < \xi\} \cap \cdots \cap \{X_n < \xi\}.$$

For every n and almost every $\omega \in C_n$ let us define a random variable A_n measurable with respect to $\sigma(X_1, \ldots, X_n)$ by $A_n = a_1$ if $X_n(\omega) = f_1(X_{n-1}(\omega))$ and $A_n = a_2$ if $X_n(\omega) = f_2(X_{n-1}(\omega))$. Clearly $(0,\xi] \subseteq (0,a]$, and on the latter interval the transformations f_1, f_2 are linear thus

$$C_{n-1} \subseteq \{X_n = A_n \cdots A_1 X_0\},$$

thus

$$C_n \subseteq \{\xi > A_n \cdots A_1 X_0\} \cap C_{n-1} = \{\xi > A_n \cdots A_1 x\} \cap C_{n-1} = \{(A_n \cdots A_1)^{-1} > (\xi/x)^{-1}\} \cap C_{n-1}.$$

By the Chebyshev inequality

$$\mathbb{P}(C_n) \leq \big(\xi/x\big)^\alpha \mathbb{E}\mathbb{1}_{C_{n-1}}(A_n \cdots A_1)^{-\alpha}.$$

It remains to estimate the last expression. The application of (5.5) gives

$$\int_{C_{n-1}} A_n^{-\alpha} \cdots A_1^{-\alpha} \mathrm{d}\mathbb{P} = \int_{C_{n-1}} \mathbb{E}\big(A_n^{-\alpha} | X_{n-1}, \ldots, X_0\big) A_{n-1}^{-\alpha} \cdots A_1^{-\alpha} \mathrm{d}\mathbb{P}$$

$$= \int_{C_{n-1}} \Big(p_1(X_{n-1})a_1^{-\alpha} + p_2(X_{n-1})a_2^{-\alpha}\Big) A_{n-1}^{-\alpha} \cdots A_1^{-\alpha} \mathrm{d}\mathbb{P} < c \int_{C_{n-2}} A_{n-1}^{-\alpha} \cdots A_1^{-\alpha} \mathrm{d}\mathbb{P}.$$

It was crucial that the set over which the integral was taken was C_{n-1} as otherwise we cannot use (5.5). Proceeding in this fashion completes the proof. □

5.3 The proof of uniqueness and stability

There are two fundamental papers dealing with the uniqueness and stability of Markov processes arising from a general class of iterated function systems with place-dependent probabilities. The first of these is [BDEG88]. For our purpose it is advantageous to know the proof there falls naturally into to parts, where in the first one it is shown that two Markov processes starting from two close points have similar distribution (this may be formally expressed as the equicontinuity of the family $(U^n\varphi)$ for an arbitrary continuous function φ; it is sometimes called the e-property), and in the second part it is shown that for two independent Markov processes (X_n) and (Y_n) starting from arbitrary two points the "coupling time" T (i.e. the minimum integer n for which (X_n) and (Y_n) are close to each other in the sense from the first part) is finite almost surely. Then, given n, the space may be decomposed into $\{T \le n\}$ and $\{T > n\}$. The probability of the second event tends to zero, and the distribution of X_n and Y_n provided $T \le n$ are close to each other[1]. It should be mentioned that the first part contained a mistake. The amendment was published two years later in erratum [BDEG89].

The second of the mentioned two papers is [LY94] in which the theorem from [BDEG88] was generalized to a wide class of Markov processes, going beyond these arising from iterated function system. The sketch is roughly the same. A reader more familiar with the literature probably find it interesting that the first part was replaced by the concept of nonexpansiveness, and the second by the lower bound technique. The sketch of our proof of Theorem 5.1 is the same as in [BDEG88], however each part is proven in a new way not being an adaptation of previous results or techniques.

Proposition 5.2. *Let f_1, f_2 be given by (5.2), and let p_1, p_2 be arbitrary positive continuous functions with $p_1(x) + p_2(x) = 1$, $x \in (0,1)$. Let us assume that (B1)-(B4) hold. If (X_n), (Y_n) are Markov processes on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ starting from x and y, respectively, and with transition probabilities (5.1), then for an arbitrary continuous function φ and $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\left|\mathbb{E}\varphi(X_n) - \mathbb{E}\varphi(Y_n)\right| < \varepsilon$$

provided x, $y \in [a, 1-a]$ and $|x - y| < \delta$.

Proposition 5.3. *Let f_1, f_2 be given by (5.2), and let p_1, p_2 be arbitrary positive continuous functions with $p_1(x) + p_2(x) = 1$, $x \in (0,1)$. Let us assume that (B1)-(B4) hold. If (X_n), (Y_n) are independent Markov processes on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ starting from x and $y > x$, respectively, and with transition probabilities (5.1), then for an arbitrary $\delta > 0$ "the coupling time"*

$$T = \min\{n \ge 0 : |X_n - Y_n| < \delta \text{ and } a \le X_n < Y_n \le 1 - a\}$$

is finite almost surely.

Now we show the proof of Theorem 5.1. The proofs of Propositions 5.2 and 5.3 are postponed to the next sections.

The proof of Theorem 5.1. Let (X_n) be a stationary Markov process from the hypothesis (existence has already been established). Let (Y_n) be a Markov process with the same transition probabilities and an arbitrary initial distribution. Take φ continuous and $\varepsilon > 0$, and let $\delta > 0$ be the constant given in Proposition 5.2. Set

$$T = \min\{n \ge 0 : |X_n - Y_n| < \delta \text{ and } X_n, Y_n \in [a, 1-a]\}.$$

[1] One can notice the resemblance to the proof of stability of Markov chains on a countable state space. In a moment the proof will be presented in our setting with all details.

By Proposition 5.3 the stopping time T is finite almost surely, thus $\mathbb{P}(T > n) \to 0$ as $n \to \infty$. Therefore

$$|\mathbb{E}\varphi(X_n) - \mathbb{E}\varphi(Y_n)| \le \Big|\mathbb{E}\mathbb{1}_{\{T\le n\}}\big(\varphi(X_n) - \varphi(Y_n)\big)\Big| + \Big|\mathbb{E}\mathbb{1}_{\{T>n\}}\big(\varphi(X_n) - \varphi(Y_n)\big)\Big|$$

$$\le \Big|\mathbb{E}\mathbb{1}_{\{T\le n\}}\mathbb{E}\big(\varphi(X_n) - \varphi(Y_n) \mid \mathcal{F}_T\big)\Big| + \mathbb{E}\mathbb{1}_{\{T>n\}}\big|\varphi(X_n) - \varphi(Y_n)\big|$$

By the triangle inequality the second summand does not exceed $2\|\varphi\|_\infty \mathbb{P}(T > n)$, and by the strong Markov property and Proposition 5.2 the conditional expectation

$$\mathbb{E}\big(\varphi(X_n) - \varphi(Y_n) \mid \mathcal{F}_T\big)$$

is less than ε almost surely on $\{T \le n\}$. Therefore

$$|\mathbb{E}\varphi(X_n) - \mathbb{E}\varphi(Y_n)| \le \varepsilon\mathbb{P}(T \le n) + 2\|\varphi\|_\infty \mathbb{P}(T > n) \to \varepsilon$$

as n goes to infinity. But $\mathbb{E}\varphi(X_n)$ is independent of n by the stationarity of the process, and is equal to $\int_{(0,1)} \varphi(x)\mu(dx)$. Since φ is an arbitrary Lipschitz function and every continuous real function on $[0, 1]$ may be approximated by a Lipschitz functions in the supremum norm, this proves the weak-$*$ convergence of the distribution of Y_n to μ. Since the initial distribution of (Y_n) was arbitrary, this completes the proof of stability. □

5.4 The proof of Proposition 5.2

Similarly to the proofs in Chapters 3 and 4 we find it useful to consider a specific model on which the random variables (X_n), (Y_n) are defined. Let $\Omega = \{1, 2\}^{\mathbb{N}}$, and let $\mathcal{F}$ be the standard product σ-algebra. The family of measures $\mathbb{P}_x$, $x \in (0, 1)$, is defined on Ω. The measure $\mathbb{P}_x$ on a cylinder $C_{i_1,\ldots,i_k}$, obtained by fixing k first coordinates to be $(i_1, \ldots, i_k)$, takes value

$$\mathbb{P}_x(C_{i_1,\ldots,i_k}) = p_{i_1}(x)p_{i_2}\big(f_{i_1}(x)\big)\cdots p_{i_k}\big(f_{i_{k-1}} \circ \cdots \circ f_{i_1}(x)\big).$$

It is a standard argument (see for example Theorem 3.1 in [Bil95]) that the measure defined on cylinders may be extended to the σ-algebra generated by cylinders, which here is identical to $\mathcal{F}$. For x fixed, the sequence of functions $\omega \longmapsto f_{\omega_n} \circ \cdots \circ f_{\omega_1}(x)$, $n \ge 0$ and $\omega = (\omega_1, \omega_2, \ldots)$, defined on $(\Omega, \mathcal{F}, \mathbb{P}_x)$ is the Markov process with transition probabilities (5.1) and starting from x. The expectation with respect to $\mathbb{P}_x$ is denoted by $\mathbb{E}_x$.

The first part is the following claim whose proof consists of several lemmas. To simplify the notation put $x_n = x_n(\omega) := f_{\omega_n} \circ \cdots \circ f_{\omega_1}(x)$, $y_n = y_n(\omega) := f_{\omega_n} \circ \cdots \circ f_{\omega_1}(y)$, $n \ge 0$ and $\omega = (\omega_1, \omega_2, \ldots)$, $x, y \in (0, 1)$.

Claim 1. *There exist $\eta > 0$, $C \ge 1$ and $q < 1$ such that $\mathbb{E}_x|x_n - y_n| \le Cq^n$ for $n \ge 1$, $x, y \in [a, 1-a]$, $|x - y| < \eta$.*

Fix $x, y \in [a, 1-a]$, $x < y$. If $a \le x_i < y_i \le 1-a$ for $i \le n$, then $|x_n - y_n| \le a_1^n|x - y|$ as both f_1 and f_2 restricted to $[a, 1-a]$ are contractions with the slope $a_1 < 1$. If it is not the case, then either (x_i) visits $(0, a)$ before (y_i) visits $(1-a, 1)$, or it is the other way around. In the first case denote the time of the first visit of (x_i) in $(0, a)$ by τ_0, and define the stopping times (see Figure 5.2)

$$\tau_1 := \min\{i \ge \tau_0 : x_i > 1/2\}$$

$$\tau_2 := \min\{i \geq \tau_1 : y_i < 1/2\}$$

and, generally,

$$\tau_{k+1} := \min\{i \geq \tau_k : x_i > 1/2\}$$

if k is even and

$$\tau_{k+1} := \min\{i \geq \tau_k : y_i < 1/2\}$$

if odd. It is clear how τ_i's should be defined in the second case.

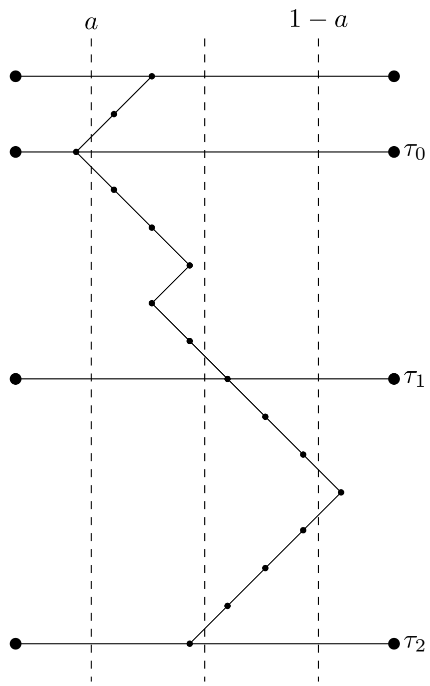

Figure 5.2: The definition of τ.

Lemma 5.1. *If $n \leq \tau_1 - 1$, then $|x_n - y_n| \leq b/a|x - y|$ for every $\omega \in \Omega$.*

Before starting the proof let us briefly present the idea behind it. Define $g_1(x) = a_1 x$ and $g_2(x) = a_2 x$ for $x > 0$. Take arbitrary positive u and v with $u < v$ and, using the same notation as for x and y, notice that the proportion of $|u_n - v_n|$ to $|u - v|$ is equal to the proportion of u_n to u (Figure 5.3; this is just a consequence of the linearity of g_1, g_2). For such system the length of $|u_n - v_n|$ may be controlled using just the information about the position of u_n with respect to the starting point (Figure 5.4). It will be helpful to keep that in mind in the sequel (note f_1 and f_2 are linear on $(0, a]$).

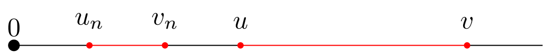

Figure 5.3: Here $u_n = 1/3u$, hence $|u_n - v_n| = 1/3|u - v|$

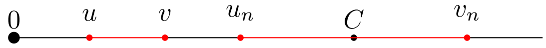

Figure 5.4: Here $u_n < C$, hence $u_n/u < C/u$ and $|u_n - v_n| \leq C/u|u - v|$

Proof. Fix $\omega \in \Omega$ and $n \leq \tau_1 - 1$. Assume for now that ω has the property that $x_i > a$ implies $x_{i+1} = f_1(x_i)$ for $i < n$. In other words ω is chosen in such a way that the trajectory of $(x_i(\omega))$ in the system f_1, f_2 corresponding to ω is the same as it would be in the system g_1, g_2. This implies in particular $x_i \leq b$ for $i \leq n$. Indeed, $x_i > b$ means that necessarily $x_j > a$ and $x_{j+1} = f_2(x_j)$ for some $j < i$ (Figure 5.5; recall that $f_2(a) = b$).

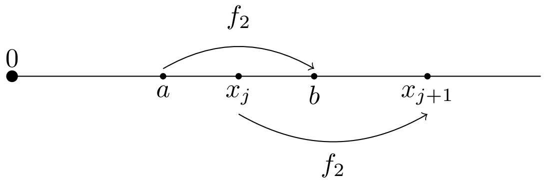

Figure 5.5: If $x_{j+1} > b$ and $x_j \leq b$, then $x_j \geq a$.

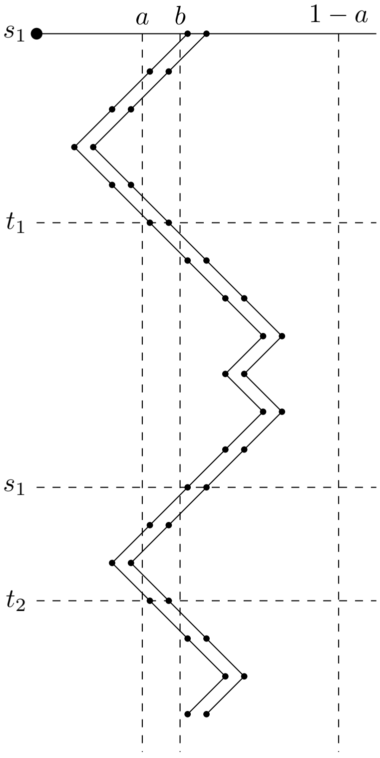

Figure 5.6: The definition of t_1, s_3, t_2.

Denote by $(x'_i(\omega))$, $(y'_i(\omega))$ the trajectories of x and y corresponding to ω in the system generated by g_1, g_2. By what has just been assumed, $(x_i(\omega))$ is equal to $(x'_i(\omega))$ for $i \in [0, n]$. Note the analogous fact for y is not true anymore. By the reasoning before the proof, $|x'_i - y'_i| \leq b/x|x - y|$ for $i \leq n$. Since $x \in [a, 1-a]$, clearly $|x'_i - y'_i| \leq \frac{b}{a}|x - y|$ for $i \leq n$. To deduce the information about the $|x_i - y_i|$ observe just that $f_2(u) \leq a_2 u$ for $u \in (0, 1)$. Therefore $[x_i, y_i] \subseteq [x'_i, y'_i]$ for $i \leq n$ and the conclusion follows.

To give the proof in the general case the trajectory must be decomposed into smaller pieces. The first piece is from 0 to the first moment s_1 such that $x_{s_1+1} < b$ and $x_{s_1} \geq b$ if $x \geq b$ (Figure 5.6) or $s_1 := 0$ if $x < b$. Both f_1 and f_2 are contracting on $[a, 1-a]$ hence $|x_i - y_i| \leq |x - y|$ for $i \leq s_1$.

Let $t_1 > s_1$ be the first index such that $x_{t_1+1} > b$. By the first part of the proof $|x_i - y_i| \leq b/x_{s_1}|x_{s_1} - y_{s_1}|$ for $i \in [s_1, t_1]$, and consequently $|x_i - y_i| \leq b/a|x - y|$ for $i \in [s_1, t_1]$ (note $a \leq x_{s_1}$. Obviously $|x_{t_1+1} - y_{t_1+1}| \leq |x_{t_1} - y_{t_1}|$ since both f_1 and f_2 are contractions on $[a, 1-a]$. Thus $|x_i - y_i| \leq b/a|x - y|$ for $i \in [0, t_1 + 1]$, and $x_{t_1+1} > b$.

After the moment $t_1 + 1$, the interval $[x_i, y_i]$ is contained in $[b, 1-a]$ for some time (let us recall here that we assumed $n \leq \tau_1 - 1$, hence $x_i < 1/2$ for $i \leq n$). So it happens till the first moment $s_2 \geq t_1 + 1$ when $x_{s_2} > b$ and $x_{s_2+1} \leq b$. It is worth to stress that possibly $s_2 = t_1 + 1$. Since both f_1, f_2 are contractions on $[a, 1-a]$, $|x_i - y_i| \leq |x_{t_1+1} - y_{t_1+1}|$ for $i \in [t_1 + 1, s_2]$ and therefore using the fact proven in the previous paragraph $|x_i - y_i| \leq b/a|x - y|$ for $i \in [0, s_2]$.

Then we repeat the whole procedure with respect to $x' := x_{s_2}$ and $y' := y_{s_2}$ and construct t_2

and s_3. The same argument as previously gives that $|x_i - y_i| \leq b/x_{s_2}|x_{s_2} - y_{s_2}|$ for $i \in [s_2, t_2]$. This time, however, the initial point x_{s_2} is greater or equal to b, hence $|x_i - y_i| \leq |x_{s_2} - y_{s_2}|$ for $i \in [s_2, t_2]$ and $|x_i - y_i| \leq |x_{s_2} - y_{s_2}| \leq b/a|x - y|$ for $i \in [0, t_2]$. Between t_2 and s_3 the interval $[x_i, y_i]$ is contained in $[b, 1-a]$ again, on which f_1 and f_2 are both contracting. Proceeding in this fashion completes the proof. □

Lemma 5.2. *There exist $\eta > 0$ and $q_1 < 1$ such that if $|x - y| < \eta$, then*

$$|x_n - y_n| \leq q_1 |x_{\tau_{k-1}-1} - y_{\tau_{k-1}-1}|$$

for some $n \in [\tau_k, \tau_{k+1} - 1]$.

Proof. Let $d := 1/2 - b$ and $q_1 := \max\{\frac{1/2-d}{1/2-d/2}, \frac{1-b}{1-a}\}$. Put $\eta := ad/2b$ and take x, y with $|x - y| < \eta$. The situation is as follows: $x_{\tau_1 - 1} \leq 1/2$ by the definition of τ_1 (recall we have assumed the first "excursion" to be in the left part of the interval $(0,1)$), $|x_{\tau_1-1} - y_{\tau_1-1}| \leq b/a|x - y| < \eta b/a = d/2$, hence $y_{\tau_1-1} \leq 1/2 + d/2$ (since τ_1 is the first moment when $x_{\tau_1} > 1/2$). The presentation is more clear in the symmetric case, thus we change coordinates $x \longmapsto 1 - x$ and interchange x and y. After that, $y_{\tau_1-1} \geq 1/2$, $x_{\tau_1-1} \geq 1/2 - d/2$ (Figure 5.7).

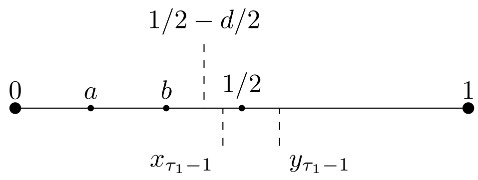

Figure 5.7: Points x_{τ_1-1}, y_{τ_1-1}.

Then either $x_{\tau_1} > b$ or $x_{\tau_1} \leq b$. In the first case $|x_{\tau_1} - y_{\tau_1}| \leq \frac{1-b}{1-a}|x_{\tau_1-1} - y_{\tau_1-1}|$ and the previous lemma (actually the reasoning in its proof) applied to $x' := x_{\tau_1}$ and $y' := y_{\tau_1}$ yields $|x_i - y_i| \leq b/x'|x' - y'|$ for $i \in [\tau_1, \tau_2 - 1]$. Since $x' \geq b$ it follows that $|x_i - y_i| \leq |x' - y'| \leq q_1|x_{\tau_1-1} - y_{\tau_1-1}|$ for $i \in [\tau_1, \tau_2 - 1]$.

In the second case let t_1 be the first moment after τ_1 with $x_{t_1} \leq b$ and $x_{t_1+1} > b$. Then $|x_{t_1+1} - y_{t_1+1}| \leq |x_{t_1} - y_{t_1}|$ since both f_1 and f_2 are contracting on $[a, 1-a]$. The trajectory (x_i) for $i \in [\tau_1 - 1, t_1]$ is the same as it would be in the linear system g_1, g_2. Therefore the proportion of $|x_i - y_i|$ to $|x_{\tau_1-1} - y_{\tau_1-1}|$ is the same as the proportion of x_i to x_{τ_1-1}, $i \in [\tau_1 - 1, t_1]$. Since $x_{t_1} \leq b$ and $x_{\tau_1-1} \geq 1/2 - d/2$ we have by this

$$|x_i - y_i| \leq b/(1/2 - d/2)|x_{\tau_1-1} - y_{\tau_1-1}| = (1/2 - d)/(1/2 - d/2)|x_{\tau_1-1} - y_{\tau_1-1}| \leq q_1|x_{\tau_1-1} - y_{\tau_1-1}|$$

for $i \in [\tau_1 - 1, t_1]$.

It remains to deal with the case $i \in [t_1 + 1, \tau_2 - 1]$, but since $x_{t_1+1} \geq b$ we can easily apply the reasoning in Lemma 5.1 to obtain the assertion. Obviously this reasoning is applicable to $i \in [\tau_2 - 1, \tau_3]$ since $|x_{\tau_2-1} - y_{\tau_2-1}| \leq |x_{\tau_1-1} - y_{\tau_1-1}| \leq \eta < d/2$. The induction argument completes the proof. □

Let $\rho_n(\omega)$ denote the maximal k with $\tau_k(\omega) \leq n$. Our goal now is to show that $\mathbb{P}_x(\rho_n < \lambda n)$ decays exponentially fast for some $\lambda \in (0,1)$. By the Chebyshev inequality

$$\mathbb{P}_x(\rho_n < \lambda n) = \mathbb{P}_x(\tau_{\lfloor \lambda n \rfloor} > n) \leq e^{-\gamma n}\mathbb{E}_x e^{\gamma \tau_{\lfloor \lambda n \rfloor}}$$

provided the expectation is finite for some $\gamma \in (0,1)$.

Lemma 5.3. *There exists $\gamma \in (0,1)$ and $C > 0$ such that $\mathbb{E}_x e^{\gamma\tau_1} \le C$ for all $x \in [a, 1-a]$.*

Proof. Let $\xi > 0$ be a number sufficiently small to satisfy Proposition 5.1. Let $\widehat{\tau_n}$ denote the length of the n-th visit of (x_n) in $(0,\xi)$. Proposition 5.1 implies the existence of constants $\widehat{C}$ and $\widehat{\gamma}$ such that $\mathbb{E}_x e^{\widehat{\gamma}\widehat{\tau_n}} \le \widehat{C}$ for all $x \in [a, 1-a]$ and all n's. Then

$$\mathbb{E}_x e^{\widehat{\gamma}(\widehat{\tau}_1 + \cdots + \widehat{\tau}_n)} = \mathbb{E}_x e^{\widehat{\gamma}(\widehat{\tau}_1 + \cdots + \widehat{\tau}_{n-1})} \mathbb{E}_x\big(e^{\widehat{\gamma}\widehat{\tau_n}} | \mathcal{F}_{\widehat{\tau}_{n-1}}\big)$$

The conditional expectation is, by the strong Markov property, bounded by $\widehat{C}$ thus induction yields

$$\mathbb{E}_x e^{\widehat{\gamma}(\widehat{\tau}_1 + \cdots + \widehat{\tau}_n)} \le \widehat{C}^n$$

for all natural n and $x \in [a, 1-a]$.

By the compactness of $[\xi, 1/2]$ and the assumption (B3) there exists $\beta > 0$ such that for every $z \in [\xi, 1/2]$ the probability that the Markov process starting from z visits $(1/2, 1)$ before the first visit in $(0,\xi)$ is grater than β. It is evident that $\mathbb{P}_x(\tau_1 \ge \widehat{\tau}_1 + \cdots + \widehat{\tau}_k) \le (1-\beta)^k$. Now

$$\mathbb{P}_x(\tau_1 > n) = \mathbb{P}_x(\{\tau_1 > n\} \cap \{\widehat{\tau}_1 + \cdots + \widehat{\tau}_k \le n\}) + \mathbb{P}_x(\{\tau_1 > n\} \cap \{\widehat{\tau}_1 + \cdots + \widehat{\tau}_k > n\})$$

$$\le \mathbb{P}_x(\tau_1 > \sigma_k) + \mathbb{P}_x(\widehat{\tau}_1 + \cdots + \widehat{\tau}_k > n) \le (1-\beta)^k + e^{-\widehat{\gamma} n}\widehat{C}^k = (1-\beta)^k + e^{-\widehat{\gamma} n + k \log \widehat{C}}.$$

Put $k = \lambda' n$ for some λ' such that $\lambda' \log \widehat{C} - \widehat{\gamma} < 0$. Then $\mathbb{P}_x(\tau_1 > n)$ decays exponentially fast, thus there exists γ such that $\mathbb{E}_x e^{\gamma\tau_1} \le C < \infty$. Note C and γ are independent of x. $\square$

Clearly $\mathbb{E}_x e^{\gamma\tau_n} \le C^n$ by the same argument the one as used for $\widehat{\tau}_1 + \cdots + \widehat{\tau}_n$. Hence, continuing the reasoning started before Lemma 5.3, $e^{-\gamma n}\mathbb{E}_x e^{\gamma\tau_{\lfloor \lambda n \rfloor}} \le e^{-\gamma n} C^{\lfloor \lambda n \rfloor} = \left(e^{-\gamma} C^{\lambda}\right)^n$. This decays exponentially fast if λ is small enough.

Now we are in position to complete the proof of Claim 1. We have

$$\mathbb{E}_x|x_n - y_n| = \mathbb{E}_x \mathbb{1}_{\{\rho_n < \lambda n\}}|x_n - y_n| + \mathbb{E}_x \mathbb{1}_{\{\rho_n \ge \lambda n\}}|x_n - y_n| \le \mathbb{P}_x(\rho_n < \lambda n) + b/a q_1^{\lambda n}|x - y|,$$

where two first lemmas were used. Third lemma implies that the first summand tends to 0 exponentially fast as described above. Hence the claim follows.

We are going to show that $(U^n\varphi)$ is equicontinuous at any point of $[a, 1-a]$, which is equivalent to the statement of Proposition 5.2. Let β denote[2] the common modulus of continuity of p_1 and p_2. Take $x \in [a, 1-a]$ and $\varepsilon > 0$. Take n_0 such that $\sum_{n=n_0}^{\infty} 2\beta(Cq^n) < \frac{\varepsilon}{6\|\varphi\|_\infty}$ (the convergence of the series comes from the Dini continuity of p_1 and p_2) and $Cq^n \le \frac{\varepsilon}{3\mathrm{Lip}(\varphi)}$ for $n \ge n_0$, where $\mathrm{Lip}(\varphi)$ denotes the Lipschitz constant of φ. By Theorem 8 on the page 45 in [Lor66] there exists a concave function β^* with $\beta(t) \le \beta^*(t) \le 2\beta(t)$. Thus we have $\sum_{n=n_0}^{\infty} \beta^*(Cq^n) < \frac{\varepsilon}{3\|\varphi\|_\infty}$.

Given a sequence $(i_1, \ldots, i_n)$ denote

$$p_{i_1,\ldots,i_n}(x) := p_{i_1}(x) p_{i_2}\big(f_{i_1}(x)\big) \cdots p_{i_n}\big(f_{i_{n-1}} \circ \cdots \circ f_{i_1}(x)\big).$$

Take y such that $|x - y| < \eta$ and such that

$$\sum \left| p_{i_1,\ldots,i_{n_0}}(x) - p_{i_1,\ldots,i_{n_0}}(y) \right| < \frac{\varepsilon}{3\|\varphi\|_\infty}, \tag{5.6}$$

[2]Sometimes β denotes a small constant and sometimes a modulus of continuity. However it is always clear from the context which of these holds.

where the summation is over all finite sequences $(i_1, \ldots, i_{n_0}) \in \{0,1\}^{n_0}$. It is satisfied provided that $|x-y|$ is less than, say, $\delta \in (0, \eta)$. Then for $n \geq n_0$ we have

$$|U^n\varphi(x) - U^n\varphi(y)|$$

$$\leq \sum p_{i_1,\ldots,i_n}(x)\Big|\varphi(f_{i_n} \circ \cdots \circ f_{i_1}(x)) - \varphi(f_{i_n} \circ \cdots \circ f_{i_1}(y))\Big|$$

$$+\Big|p_{i_1,\ldots,i_n}(x) - p_{i_1,\ldots,i_n}(y)\Big|\|\varphi\|_\infty,$$

where the summation is over all finite sequences $(i_1, \ldots, i_n) \in \{0,1\}^n$. The first term is bounded by $\mathrm{Lip}(\varphi)\mathbb{E}_x|f_\omega^n(x) - f_\omega^n(y)|$. To estimate the second, we have

$$\sum_{i_1,\ldots,i_n} \Big|p_{i_1,\ldots,i_n}(x) - p_{i_1,\ldots,i_n}(y)\Big|$$

$$= \sum_{i_1,\ldots,i_n} \Big|p_{i_n}(f_{i_{n-1}} \circ \cdots \circ f_{i_1}(x)) - p_{i_n}(f_{i_{n-1}} \circ \cdots \circ f_{i_1}(y))\Big| \cdot p_{i_1,\ldots,i_{n-1}}(x)$$

$$+ \sum_{i_1,\ldots,i_n} p_{i_n}(f_{i_{n-1}} \circ \cdots \circ f_{i_1}(y))\Big|p_{i_1,\ldots,i_{n-1}}(x) - p_{i_1,\ldots,i_{n-1}}(y)\Big|$$

$$\leq 2\mathbb{E}_x\beta^*(|f_\omega^n(x) - f_\omega^n(y)|) + \sum_{i_1,\ldots,i_{n-1}} \Big|p_{i_1,\ldots,i_{n-1}}(x) - p_{i_1,\ldots,i_{n-1}}(y)\Big|.$$

The modulus of continuity β^* is concave, therefore by the Jensen inequality we have

$$\sum_{i_1,\ldots,i_n} \Big|p_{i_1,\ldots,i_n}(x) - p_{i_1,\ldots,i_n}(y)\Big| \leq 2\beta^*(Cq^n) + \sum_{i_1,\ldots,i_{n-1}} \Big|p_{i_1,\ldots,i_{n-1}}(x) - p_{i_1,\ldots,i_{n-1}}(y)\Big|.$$

Continuing this procedure while $n > n_0$ and using (5.6) yields

$$\sum_{i_1,\ldots,i_n} \Big|p_{i_1,\ldots,i_n}(x) - p_{i_1,\ldots,i_n}(y)\Big|$$

$$\leq \sum_{i=n_0}^{n} 2\beta^*(Cq^i) + \sum_{i_1,\ldots,i_{n_0}} \Big|p_{i_1,\ldots,i_{n_0}}(x) - p_{i_1,\ldots,i_{n_0}}(y)\Big| < \frac{\varepsilon}{3\|\varphi\|_\infty} + \frac{\varepsilon}{3\|\varphi\|_\infty}.$$

Again by the definition of n_0 we have

$$|U^n\varphi(x) - U^n\varphi(y)| < \mathrm{Lip}(\varphi)\mathbb{E}_x|f_\omega^n(x) - f_\omega^n(y)|$$

$$+\|\varphi\|_\infty \frac{\varepsilon}{3\|\varphi\|_\infty} + \|\varphi\|_\infty \frac{\varepsilon}{3\|\varphi\|_\infty} < \varepsilon$$

for all n and y with $|x-y| < \delta$. Therefore $(U^n\varphi)$ is equicontinuous at any $x \in [a, 1-a]$.

5.5 The proof of Proposition 5.3

Claim 2. *Fix $\delta > 0$. There exists $\beta > 0$ such that $\mathbb{P}_x \otimes \mathbb{P}_y(T < \infty) \geq \beta$ for every $x, y \in (0,1)$, where $T = \min\{n \geq 0 : x_n, y_n \in [a, 1-a]$ and $|x_n - y_n| < \delta\}$.*

Let $\xi > 0$ be so small to satisfy Proposition 5.1. Let M, α be the constants given in Proposition 5.1, and let $\zeta \in (0, \xi)$ be such that $M\zeta^\alpha < \frac{1}{8}$. By Proposition 5.1, given two points x, y there exists k_1 such that

$$\mathbb{P}_x \otimes \mathbb{P}_y\left(\bigcap_{i=0}^{k_1}\{x_i \notin [\xi, 1-\xi]\} \cup \bigcap_{i=0}^{k_1}\{y_i \notin [\xi, 1-\xi]\}\right) < 1/2. \tag{5.7}$$

Since $\delta_y \in \mathcal{P}_{M,\alpha}$ for every $y \in [\xi, 1-\xi]$ and $M\zeta^\alpha < 1/8$ we easily conclude that

$$\mathbb{P}_x\left(x_{k_1} \notin [\zeta, 1-\zeta] \middle| \bigcup_{i=0}^{k_1}\{x_i \in [\xi, 1-\xi]\}\right) < 1/8 + 1/8 = 1/4$$

and

$$\mathbb{P}_y\left(y_{k_1} \notin [\zeta, 1-\zeta] \middle| \bigcup_{i=0}^{k_1}\{y_i \in [\xi, 1-\xi]\}\right) < 1/8 + 1/8 = 1/4,$$

thus

$$\mathbb{P}_x \otimes \mathbb{P}_y\left(x_{k_1}, y_{k_1} \in [\zeta, 1-\zeta] \middle| \bigcup_{i=0}^{k_1}\{x_i \in [\xi, 1-\xi]\} \cap \bigcup_{i=0}^{k_1}\{y_i \in [\xi, 1-\xi]\}\right) \geq 1/2.$$

By (5.7) $\mathbb{P}_x \otimes \mathbb{P}_y(x_{k_1}, y_{k_1} \in [\zeta, 1-\zeta]) \geq 1/4$.

Lemma 5.4. *There exists a point $z \in (a, 1-a)$ such that for every $\zeta > 0$ and $\delta > 0$ there exist a natural number k_2 and $\beta' > 0$ such that*

$$\mathbb{P}_x(x_{k_2} \in (z - \delta/2, z + \delta/2)) > \beta'$$

for $x \in [\zeta, 1-\zeta]$.

Proof. Observe that $f_1(b) = a_1 b = \frac{1-b}{1-a} b > a$. Indeed, it is equivalent to $(1-b)b > (1-a)a$, which is implied by $a < b < 1/2$ (see Figure 5.8).

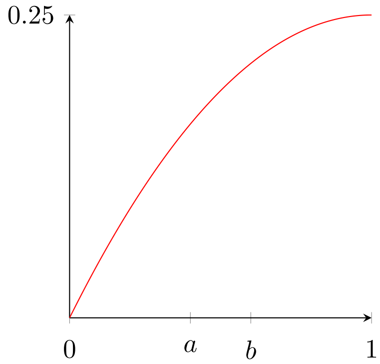

Figure 5.8: The plot of the function $t \longmapsto t(1-t)$ for $t \in (0, 1/2)$. The function is increasing there.

Hence $f_1([b,1-a]) \subseteq [a,1-b]$. By the symmetry of the system $f_2([a,1-b]) \subseteq [b,1-a]$. Hence the composition $f_1 \circ f_2$ restricted to the interval $[a,1-b]$ is a contraction and acts into the interval $[a,1/2]$. Let z be the unique attractive fixed point for this composition on $[a,1-b]$. For any point $x \in [a,1-b]$ and $\delta > 0$ there exists m' such that $\mathbb{P}_x(x_{2m'} \in (z-\delta/2, z+\delta/2)) > 0$.

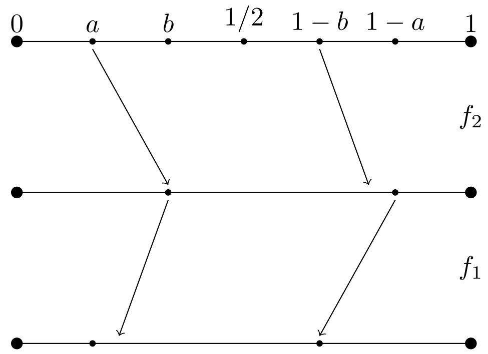

Figure 5.9: $f_1([b,1-a]) \subseteq [a,1-b]$ and $f_2([a,1-b]) \subseteq [b,1-a]$

Choose $\zeta > 0$. To complete the proof it is sufficient to show that for any $x \in [\zeta, 1-\zeta]$ there exists a number m'' and ω such that $x_{2m''}(\omega) \in [a,1-b]$ (it is problematic that the number should be necessarily even). Then $k_2 := 2m' + 2m'''$ will be desired number, where m''' is the maximum of m'' for $x \in [\zeta, 1-\zeta]$ (recall the probabilities p_1, p_2 are positive by (B3)).

It is readily seen that there exist m''' and a sequence ω such that $z_1 := x_{m'''} \in [a,1-b]$. If m''' is even, then put $m'' = m'''/2$. If not, then apply f_1 to z_1. If $f_1(z_1) \geq a$ then $m'' = (m'''+1)/2$ is the desired number. If not, then $f_1(z_1) < a$, hence $z_2 := f_2 \circ f_1(z_1 \geq b$. Note that $z_2 = a_2 a_1 z_1 = \frac{(1-b)b}{(1-a)a} z_1 > z_0$ by the same argument as in the beginning of the proof. We can repeat this procedure and define $z_{n+1} > z_n$ whenever $f_1(z_n) < a$. This procedure, however, must finish for some n, since $z_{n+1} = (a_2 a_1)^n z_1$ which eventually becomes greater than b for some n, which means that $f_1(z_n) = f_2^{-1}(z_{n+1}) > a$. If n is the minimal number with $f_1(z_n) \geq a$. Then $2m'' = m''' + 2n + 1$ has the desired property. ☐

Put $\beta := 1/2\big(\beta'\big)^2$. Lemma 5.4 clearly implies Claim 2.

We are ready to finish the proof of Proposition 5.2. Assume contrary that $\mathbb{P}_x \otimes \mathbb{P}_y$ measure of $F := \{T = \infty\}$ is positive for some points x and y. Note F is the complement of an open set and thus it is closed. By the regularity of $\mathbb{P}_x \otimes \mathbb{P}_y$ there exists an open G containing F with $\mathbb{P}_x \otimes \mathbb{P}_y(F|G) > 1-\beta$. The open set G is a sum of cylinders (G_i) and the sum may be assumed to be finite (by the compactness of F). Moreover,

$$\sum_i \frac{\mathbb{P}_x \otimes \mathbb{P}_y(G_i)}{\mathbb{P}_x \otimes \mathbb{P}_y(G)} \mathbb{P}_x \otimes \mathbb{P}_y(F|G_i) = \mathbb{P}_x \otimes \mathbb{P}_y(F|G) > 1-\beta$$

hence $\mathbb{P}_x \otimes \mathbb{P}_y(F|G_i) > 1-\beta$ for at least one of G_i's. The cylinder G_i is of the form $\{(\omega,\omega') \in \Omega \times \Omega : \omega_1 = i_1, \cdots, \omega_k = i_k \text{ and } \omega'_1 = \mathrm{j}_1, \cdots, \omega'_\mathrm{k} = \mathrm{j}_\mathrm{k}\}$ for some sequences $(i_1,\ldots,i_k)$ and $(j_1,\ldots,j_k)$. Put $u = f_{i_k} \circ \cdots \circ f_{i_1}(x)$ and $v = f_{j_k} \circ \cdots \circ f_{j_1}(y)$, and $T_{u,v}$ to be defined as T but with x and y replaced by u and v. Then

$$\mathbb{P}_u \otimes \mathbb{P}_v(T_{u,v} = \infty) = \mathbb{P}_x \otimes \mathbb{P}_y(F|G_i) > 1-\beta,$$

which contradicts Claim 2.

5.6 Comments

We use the expression "Alsedà-Misiurewicz systems" after [BS21]. The authors were interested in invariant Cantor sets and absolute continuity of stationary distributions of Alsedà-Misiurewicz systems with constant probabilities.

Assumptions (B2), (B3), (B4) are rather essential in the proof. It remains a question to what extent assumption (B1) may be relaxed. In [Czu20] the proof includes also the boundary case when $a = 1/2$. However, the reasoning is a bit more technical. The problem is that Lemmas 5.1 and 5.2 generally do not hold beyond systems (5.2) with (B1). The mentioned boundary case is an exception. One can formulate more general conditions implying Lemmas 5.1 and 5.2, but these appear to be rather artificial. It would be interesting thus to see a proof which does not rely on Lemmas 5.1 and 5.2. Then the expansion should be somehow controlled, which seems to be difficult (or even impossible) if no assumptions on p_i's are imposed.

It remains also an open problem to establish the rate of convergence and show the central limit theorem. Plausibly it would be helpful to prove that if V is small interval contained in $[a, 1-a]$ and T is defined to be the moment of the first common visit in V of two independent stationary processes X and Y, then $\mathbb{E}e^{\gamma T} < \infty$ for some $\gamma > 0$ sufficiently small. If V is sufficiently small, then $|f^n_\omega(x) - f^n_\omega(y)| \leq Cq^n|x - y|$ for some $q \in (0, 1)$ and ω from some set of positive $\mathbb{P}_x \otimes \mathbb{P}_y$-measure. Thus coupling techniques introduced in [Hai02] probably can be applied (see also [Ś11]).

Appendix A

The proof of the Baxendale theorem

Theorem (Lemma 4.1 in [GH17])**.** *Let $f_1, \ldots, f_m$ be C^2 orientation preserving diffeomorphisms of $[0,1]$ satisfying (A1) and (A2). If $(p_1, \ldots, p_m)$ is such that Λ_0, Λ_1 are positive, then the volume Lyapunov exponent (with respect to the unique stationary distribution μ)*

$$\sum_{i=1}^{m} p_i \int_{[0,1]} \log f_i'(x)\mu(dx)$$

is negative.

The proof is based on the notion of relative entropy. If ν_1, ν_2 are probability measures, then the relative entropy of ν_1 with respect to ν_2 is defined as

$$h(\nu_1|\nu_2) := \sup_{\psi \in C[0,1]} \ln \left(\int_{[0,1]} e^{\psi(x)} \nu_1(dx) \right) - \int_{[0,1]} \psi(x)\nu_2(dx).$$

In [DV75] (see Lemma 2.1 therein) it is shown that

- $0 \leq h(\nu_1|\nu_2) \leq \infty$,
- $h(\nu_1|\nu_2) = 0$ if and only if $\nu_1 = \nu_2$,
- $h(\nu_1|\nu_2)$ is finite if and only if ν_1 is absolutely continuous with respect to ν_2 and the density satisfies $\int_{[0,1]} \frac{d\nu_1}{d\nu_2} \log \frac{d\nu_1}{d\nu_2} d\nu_2 < \infty$. Moreover, in that case

$$h(\nu_1|\nu_2) = \int_{[0,1]} \frac{d\nu_1}{d\nu_2} \log \frac{d\nu_1}{d\nu_2} d\nu_2 = \int_{[0,1]} \log \frac{d\nu_1}{d\nu_2} d\nu_1.$$

When μ is absolutely continuous with respect to the Lebesgue measure and the density is positive and bounded, then one can check that $\sum_{i=1}^{m} p_i h((f_i)_*\mu|\mu) = -\Lambda$ (we shall do this later). Therefore the following lemma would give the assertion for μ absolutely continuous with respect to the Lebesgue measure having a bounded positive density.

Lemma A.1. *If $f_1, \ldots, f_m$ is a system satisfying (A1) and (A2), $p_1, \ldots, p_m$ is such that the Lyapunov exponents at* 0 *and* 1 *are positive, then $\sum_{i=1}^{m} p_i h((f_i)_*\mu|\mu) > 0$.*

Proof. Since (A1) holds, there exists i and $\xi > 0$ such that $f_i(x) < x$ for $x \leq \xi$. The same assumption implies $\mu((0,\xi)) =: r > 0$. If $h((f_i)_*\mu|\mu) = 0$, then $(f_i^n)_*\mu = \mu$ and thus $\mu((0, f_i^n(\xi))) = \mu((0,\xi)) = r > 0$ for every n, which implies that $\mu(\{\emptyset\}) \geq r$, which is a contradiction. Therefore $h((f_i)_*\mu|\mu) > 0$ and the average entropy is positive as well. $\square$

Unfortunately, it is hard to deduce whether a system has a stationary distribution absolutely continuous with respect to the Lebesgue measure. Actually the results in [CS20a] (Theorem 10), [BS21] (Theorem 2.16) and [BR21] (Theorem 5.1) say that one should expect the measure to be rather singular than absolutely continuous[1]. Therefore the presented sketch of reasoning is far from being sufficiently general. The idea is to perturb the system to obtain other Markov process with an absolutely continuous stationary distribution μ_ε with sufficiently regular density. For perturbed system one can show the relation $\mathbb{E}h((f_\omega)_*\mu_\varepsilon|\mu_\varepsilon) = -\Lambda_\varepsilon$. Relative entropy is upper semicontinuous (as the supremum of continuous functionals), thus $\limsup \mathbb{E}h((f_\omega)_*\mu_\varepsilon|\mu_\varepsilon) \geq \mathbb{E}h((f_\omega)_*\mu|\mu) > 0$ by Lemma A.1. The volume Lyapunov exponent is a continuous functional therefore the passage to the limit gives $\Lambda_\varepsilon \to \Lambda$. Therefore $-\Lambda \geq \mathbb{E}h((f_\omega)_*\mu|\mu) > 0$. Details are provided in the sequel.

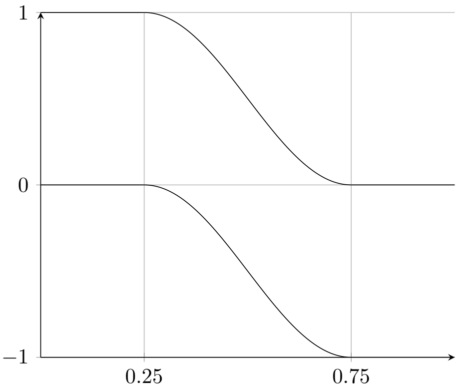

Figure A.1: The graph of ψ.

Let ψ be a nonincreasing smooth function such that $\psi(x) = 1$ for $x \leq 1/4$, $\psi(x) = 0$ for $x \geq 3/4$ (Figure A.1). Given $\varepsilon > 0$, $u \in [0, 1]$ and $i = 1, \ldots, m$ define $f_{i,u,\varepsilon}(x) := f_i(x) + \varepsilon(\psi(x) - 1) + \varepsilon u$, and let us consider a Markov process in which, at every step, a function $f_{i,u,\varepsilon}$ is chosen, where i is distributed on $\{1, \ldots, m\}$ according to the probability vector $(p_1, \ldots, p_m)$, u is uniformly distributed on $[0, 1]$ and i and u are independent. Observe that the graphs of $f_{i,u,\varepsilon}$, $u \in [0, 1]$, are parallel to each other (see Figure A.2). Note also the process is defined on $[0, 1]$, and the transition probabilities are of the form

$$p_\varepsilon(x, A) = \sum_{i=1}^{m} p_i \int_A k_{i,\varepsilon}(x, y)dy$$

for some positive real functions $k_{i,\varepsilon}$, $i = 1, \ldots, m$ (more exactly for x fixed these are characteristic functions of $[f_{i,0,\varepsilon}(x), f_{i,1,\varepsilon}(x)]$ normalized to be a density). Finally, the Markov and dual operators are given by formulae

$$P_\varepsilon\nu(A) = \sum_{i=1}^{m} p_i \int_{[0,1]} \int_{[0,1]} \mathbb{1}_A(y)k_{i,\varepsilon}(x, y)dy\nu(dx)$$

for a Borel set A and

$$U_\varepsilon\varphi(x) = \sum_{i=1}^{m} p_i \int_{[0,1]} k_{i,\varepsilon}(x, y)\varphi(y)dy$$

[1] It has been proven in [CS20a] a generic system of homeomorphisms with supremum norm has unique stationary measure singular with respect to the Lebesgue measure. It would be interesting to prove the same result for diffeomorphisms in C^k topology.

for $\varphi \in \mathcal{B}([0,1])$. Clearly, P_ε is a Markov-Feller operator.

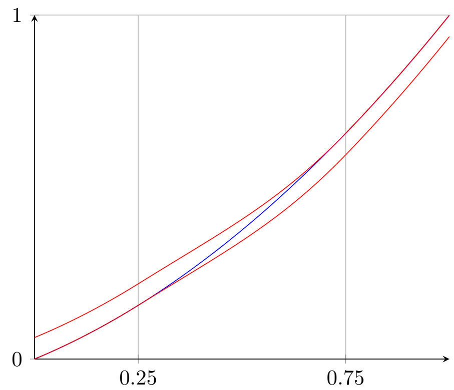

Figure A.2: The graph of f_i (blue), $f_{i,0,\varepsilon}$ and $f_{i,0,\varepsilon}$ (red). The graph of $f_{i,\varepsilon,u}$, $u \in [0,1]$, is parallel to the graphs of $f_{i,0,\varepsilon}$ and $f_{i,0,\varepsilon}$.

Lemma A.2. *Let $a > 0$ be such that the transition from $[1/4, 1)$ to $(0, a]$ and from $(0, 3/4]$ to $[1-a, 1)$ is impossible in one step. If M, α are the constants given by Proposition 3.1 suitable for a, then $\mathcal{P}_{M,\alpha}$ is P_ε- invariant for every $\varepsilon > 0$.*

Proof. Recall that M, α are such that every measure supported on $[a, 1-a]$ belongs to $\mathcal{P}_{M,\alpha}$. Moreover, $\mathcal{P}_{M,\alpha}$ is P-invariant. This implies that $\delta_a, \delta_{1-a} \in \mathcal{P}_{M,\alpha}$, which implies in turn that $1 \leq Ma^\alpha$ (this follows from the definition of $\mathcal{P}_{M,\alpha}$). This means that $\nu((0,x]) \leq Mx^\alpha$ for an arbitrary measure ν and $x \geq a$. Similarly, $\nu([1-x, 1)) \leq Mx^\alpha$ for an arbitrary measure ν and $x \geq a$. Therefore showing P_ε-invariance of $\mathcal{P}_{M,\alpha}$ requires only the proof that $P_\varepsilon\nu((0,x]) \leq Mx^\alpha$ and $P_\varepsilon\nu([1-x,1)) \leq Mx^\alpha$ for $x < a$. The proof will be carried out for the first of these inequalities while the second is just its simple adaptation.

Take φ nonincreasing and $t \leq 1/4$. The support of $k_{i,\varepsilon}(t, \cdot)$ is contained in $[f_i(t), 1)$ for $t \leq 1/4$. Thus the monotonicity of φ yields $\int_{[0,1]} k_{i,\varepsilon}(t,y)\varphi(y)dy \leq \varphi(f_i(t))$ for $i = 1, \ldots, m$ and $t \leq 1/4$. Plugging that into the definition of $U_\varepsilon\varphi$ gives $U_\varepsilon\varphi(t) \leq \sum p_i\varphi(f_i(t)) = U\varphi(t)$ for $t \leq 1/4$.

Now observe that the transition from $(1/4, 1)$ to $(0, x]$ is impossible for any $x \leq a$, both for the perturbed and the non-perturbed process. Indeed, for non-perturbed process it follows from the definition of a. For perturbed system one needs to use again the fact that the support of $k_{i,\varepsilon}(t,\cdot)$ lies in $[f_i(t), 1)$ for $t \leq 1/4$, hence the claim for perturbed system follows from the same claim for the non-perturbed system. In particular, if φ is zero on $[a, 1)$, then $U\varphi$ and $U_\varepsilon\varphi$ are equal to zero on $[1/4, 1)$.

Finally take $x < a$ and put φ to be the characteristic function of $(0, x]$. Using previous observations yields

$$P_\varepsilon\nu((0,x]) = \int_{[0,1]} U_\varepsilon\varphi(t)\nu(dt) = \int_{[0,1/4]} U_\varepsilon\varphi(t)\nu(dt)$$

$$\leq \int_{[0,1/4]} U\varphi(t)\nu(dt) = \int_{[0,1]} U\varphi(t)\nu(dt) = P\nu((0,x]) \leq Mx^\alpha,$$

provided $\nu \in \mathcal{P}_{M,\alpha}$. $\square$

Lemma A.3 ([HZ07])**.** *For every $\varepsilon > 0$ the Markov operator P_ε possess an absolutely continuous stationary measure μ_ε with a bounded, continuous and positive density φ_ε.*

Proof. Repeat the reasoning in the proof of the corollary to Proposition 3.1 to show that, given $\varepsilon > 0$, there exist $\mu_\varepsilon \in \mathcal{P}_{M,\alpha}$ invariant under P_ε (remember $\mathcal{P}_{M,\alpha}$ is P_ε- invariant). Observe that μ_ε is necessarily absolutely continuous with respect to the Lebesgue measure. It is a consequence of the fact that if A is of the Lebesgue measure zero, then $p(x, A) = 0$ whatever x is (the transition probabilities are absolutely continuous with respect to the Lebesgue measure).

The Perron- Frobenius operator L takes the form

$$L\varphi(x) = \sum_{i=1}^{m} p_i \int_{[0,1]} k_{i,\varepsilon}(y, x)\varphi(y)dy,$$

where φ is a non-negative Borel measurable real function. Indeed, $L\varphi \geq 0$ provided $\varphi \geq 0$, and L preserves integrals as the following calculation shows:

$$\int_{[0,1]} L\varphi(x)dx = \sum_{i=1}^{m} p_i \int_{[0,1]} \int_{[0,1]} k_{i,\varepsilon}(y, x)\varphi(y)dydx =$$

$$\sum_{i=1}^{m} p_i \int_{[0,1]} \left(\int_{[0,1]} k_{i,\varepsilon}(y, x)dx \right) \varphi(y)dy = \int_{[0,1]} \varphi(y)dy.$$

Therefore L preserves densities on $[0, 1]$ (note we could use the Fubini theorem as all functions are non-negative). Finally, if $\nu(dy) = \varphi(y)dy$, then

$$\int_A L\varphi(x)dx = \int_{[0,1]} \int_{[0,1]} \left(\sum_{i=1}^{m} p_i k_{i,\varepsilon}(y, x) \right) \mathbb{1}_A(x)\varphi(y)dydx$$

$$= \int_{[0,1]} \int_{[0,1]} \left(\sum_{i=1}^{m} p_i k_{i,\varepsilon}(y, x) \right) \mathbb{1}_A(x)\varphi(y)dxdy = \int_{[0,1]} U_\varepsilon \mathbb{1}_A(y)\varphi(y)dy = \int_{[0,1]} U_\varepsilon \mathbb{1}_A(y)\nu(dy)$$

$$= \int_{[0,1]} \mathbb{1}_A(y)P_\varepsilon\nu(dy) = P_\varepsilon\nu(A).$$

Since μ_ε is absolutely continuous with respect to the Lebesgue measure, L possess an invariant density φ_ε. The boundeness and continuity of invariant density is proven by showing that L transforms integrable functions into bounded continuous functions.

Boundeness is a consequence of the simple computation:

$$L\varphi(x) \leq \sum_{i=1}^{m} p_i \|k_{i,\varepsilon}\|_\infty \int_{[0,1]} \varphi(y)dy = \varepsilon^{-1} \int_{[0,1]} \varphi(y)dy,$$

as $\|k_{i,\varepsilon}\| = \varepsilon^{-1}$ (for x fixed it is a characteristic function of an interval of the length ε).

The density $L\varphi$ is continuous provided $\varphi \in L^1$. Indeed,

$$|L\varphi(x_1) - L\varphi(x_2)| = \left| \sum_{i=1}^{m} p_i \int_{[0,1]} k_{i,\varepsilon}(y, x_1)\varphi(y)dy - \sum_{i=1}^{m} p_i \int_{[0,1]} k_{i,\varepsilon}(y, x_1)\varphi(y)dy \right|$$

$$\leq \sum_{i=1}^{m} p_i \int_{[0,1]} \big|k_{i,\varepsilon}(y, x_1) - k_{i,\varepsilon}(y, x_2)\big|\varphi(y)dy.$$

Set $V_i = \{y \in [0,1] : k_{i,\varepsilon}(y, x_1) \neq k_{i,\varepsilon}(y, x_2)\}$, and observe that $|k_{i,\varepsilon}(y, x_1) - k_{i,\varepsilon}(y, x_2)| = 1/\varepsilon$ for y on V_i. Hence

$$|L\varphi(x_1) - L\varphi(x_2)| \leq 1/\varepsilon \sum_{i=1}^{m} p_i \int_{V_i} \varphi(y)dy.$$

Thus to show continuity we need to show that $\int_{V_i} \varphi(y)dy \to 0$ as $x_1 \to x_2$.

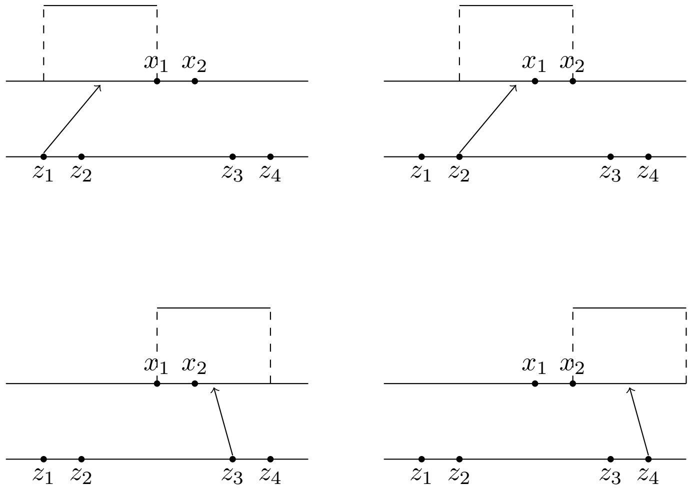

Figure A.3: The points z_1, z_2, z_3, z_4.

To this end let us assume, without loss of generality, that $x_1 < x_2$ and define four points (see Figure A.3):

- z_1 which is the unique number such that x_1 is the right endpoint of the support of $k_{i,\varepsilon}(z_1, \cdot)$,
- z_2 which is the unique number such that x_2 is the right endpoint of the support of $k_{i,\varepsilon}(z_2, \cdot)$,
- z_3 which is the unique number such that x_1 is the left endpoint of the support of $k_{i,\varepsilon}(z_3, \cdot)$,
- z_4 which is the unique number such that x_2 is the left endpoint of the support of $k_{i,\varepsilon}(z_4, \cdot)$.

Since the support of $k_{i,\varepsilon}(z, \cdot)$ has always length ε and we consider the situation when x_1 and x_2 are close to each other we can assume that $z_1 < z_2 < z_3 < z_4$ as depicted in Figure A.3. Note it may happen that z_1 or even z_2 is not well-defined when x_1, x_2 are close to 0. Similarly it may happen that z_4 or even z_3 are not well-defined when x_4 and x_3 are close to 1. In the former case put $z_1 = 0$ (and $z_2 = 0$ if it is also not well-defined) and in the latter put $z_4 = 1$ (and $z_3 = 1$ if it is not well-defined).

Observe that $V_i \subseteq [z_1, z_2] \cup [z_3, z_4]$. If all these points are well-defined then $z_1 = (f_{i,1,\varepsilon})^{-1}(x_1)$, $z_1 = (f_{i,1,\varepsilon})^{-1}(x_2)$, $z_1 = (f_{i,0,\varepsilon})^{-1}(x_1)$, $z_1 = (f_{i,0,\varepsilon})^{-1}(x_2)$. Since both $f_{i,0,\varepsilon}$ and $f_{i,0,\varepsilon}$ are diffeomorphisms onto their images this implies that z_1, z_2, z_3, z_4 depend continuously on x_1 and x_2. It is immediate to see that the continuous dependence remains even if some of z_1, z_2, z_3, z_4 appears to be 0 or 1. Therefore $\int_{[z_1,z_2]\cup[z_3,z_4]} \varphi(y)dy \to 0$ as $x_1 \to x_2$ by the integrability of φ.

The density φ_ε is positive on $(0, 1)$. To explain this observe the transition probabilities $P_\varepsilon \delta_x$, $x \in (0, 1)$,

(i) are absolutely continuous with respect to the Lebesgue measure,

(ii) with the densities which are piecewise continuous (moreover, there are finitely many of pieces of continuity),

(iii) depending locally continuously on $x \in (0,1)$ in the supremum norm (locally refers to a neighborhood when the density is continuous),

(iv) positive on some set of the form $(f_i(x) - \delta, f_i(x))$ or $(f_i(x), f_i(x) + \delta)$ for every $i = 1, \dots m$.

These are just consequences of the definition of the transition probabilities. All of this may be said about the family of transition probabilities in n steps $P_\varepsilon^n \delta_x$, $x \in (0,1)$.

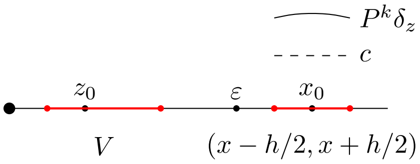

Figure A.4: The curve represents the density of $P^k\delta_z$.

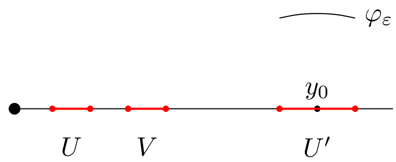

Figure A.5

Let us assume, contrary to the claim, that $\varphi_\varepsilon(x_0) = 0$ for some $x_0 \in (0,1)$. By assumption (A1) there exists $(i_1, \dots, i_k)$ and $z_0 \in (0, \varepsilon)$ with $f_{i_k} \circ \dots \circ f_{i_1}(z_0) = x_0$ (Figure A.4). By what has been stated previously, especially points (ii) and (iv), $P_\varepsilon^k \delta_{z_0}$ has the density continuous and positive at x_0. By (iii) the same is true for $P_\varepsilon^k \delta_z$, where z is from some open set V whose closure is contained in $(0, \varepsilon)$. Actually, the densities $P_\varepsilon^k \delta_z$, $z \in V$, are uniformly bounded from 0 on some neighborhood of x_0 by, say, $c > 0$.

Take $h > 0$ so small that $(x_0 - h/2, x_0 + h/2)$ is contained in the mentioned neighborhood, where the densities of $P_\varepsilon^k \delta_z$ are positive, $z \in V$. Then the stationarity of μ_ε yields

$$\mu_\varepsilon\big((x_0 - h/2, x_0 + h/2)\big) = \int_{[0,1]} P^k \delta_z\big((x_0 - h/2, x_0 + h/2)\big)\mu_\varepsilon(dz)$$
$$\geq \int_V P^k \delta_z\big((x_0 - h/2, x_0 + h/2)\big)\mu_\varepsilon(dz) \geq c \cdot h \cdot \mu_\varepsilon(V).$$

What remains to show is that V is of positive μ_ε measure, which follows just from the fact that the closure of V is contained in $(0, \varepsilon)$. Indeed, let us take y_0 with $\varphi_\varepsilon(y_0) > 0$. Again using (A1) there exists $(j_1, \dots, j_{k'})$ such that $f_{j_{k'}} \circ \dots \circ f_{j_1}(y_0) < \inf V$, therefore $P_\varepsilon^{k'} \delta_y$ is positive on some open nonempty set U whose closure is contained in $(0, \inf V)$, where y is from some neighborhood U' of y_0. Note that for every $z < \inf V$ the density of $P_\varepsilon \delta_z$ is positive on $[z, \varepsilon]$, hence is positive on V. Again the application of the stationarity of μ_ε and the Chapman-Kolmogorov equations (see Corollary 8.3 in [Kal02]) yields

$$\mu_\varepsilon(V) = \int_{[0,1]} P^{k'+1}\delta_y(V)\mu_\varepsilon(dy) = \int_{[0,1]} \int_{[0,1]} P\delta_z(V) P^{k'}\delta_y(dz)\mu_\varepsilon(dy)$$
$$\geq \int_{U'} \int_U P\delta_z(V) P^{k'}\delta_y(dz)\mu_\varepsilon(dy) > 0,$$

which is the conclusion. $\square$

Lemma A.4. *It holds that $\mu_\varepsilon \to \mu$ in the weak-$*$ topology.*

Proof. Note that $\mu_\varepsilon \in \mathcal{P}_{M,\alpha}$, which is weak-$*$ compact, hence it is sufficient to show that $\mu_{\varepsilon_k} \to \nu$ for some sequence (ε_k) converging to zero implies $\nu = \mu$.

We start with the observation that $\|U_\varepsilon \varphi - U\varphi\|_\infty \to 0$ as $\varepsilon \to 0$ for arbitrary continuous function φ on $[0,1]$. Indeed, φ is uniformly continuous then, hence, for every $\delta > 0$ there exists $\varepsilon > 0$ such that $|\varphi(f_i(x)) - \varphi(f_i(y))| \leq \delta$ for $i = 1, \dots, m$ provided $|x - y| < \varepsilon$. Hence for an arbitrary $\delta > 0$ and sufficiently small $\varepsilon >$ we have

$$
\begin{aligned}
|U_\varepsilon \varphi(x) - U\varphi(x)| &\leq \left| \sum_{i=1}^m p_i \int_{[0,1]} k_{i,\varepsilon}(x,y)\varphi(y)dy - \sum_{i=1}^m p_i \varphi(f_i(x)) \right| \\
&= \left| \sum_{i=1}^m p_i \frac{1}{\varepsilon} \int_{f_{i,0,\varepsilon}(x)}^{f_{i,1,\varepsilon}(x)} \varphi(y)dy - \sum_{i=1}^m p_i \frac{1}{\varepsilon} \int_{f_{i,0,\varepsilon}(x)}^{f_{i,1,\varepsilon}(x)} \varphi(f_i(x)))dy \right| \\
&\leq \frac{1}{\varepsilon} \sum_{i=1}^m p_i \int_{f_{i,0,\varepsilon}(x)}^{f_{i,1,\varepsilon}(x)} \big|\varphi(y) - \varphi(f_i(x))\big| dy.
\end{aligned}
$$

But for each i the diameter of the set on which the integral is defined is equal to ε, and $f_i(x)$ belongs to it. Hence $|\varphi(y) - \varphi(f_i(x))| < \delta$ for all y in this interval. This finally implies

$$
|U_\varepsilon \varphi(x) - U\varphi(x)| \leq \sum_{i=1}^m p_i \delta = \delta,
$$

for every $x \in [0,1]$ and ε sufficiently small, which implies the claim.

Let $\mu_{\varepsilon_k} \to \nu$ in the weak-$*$ topology for some sequence ε_k convergent to zero. Take φ continuous on $[0,1]$. Then

$$
\begin{aligned}
&\left| \int_{[0,1]} \varphi dP\nu - \int_{[0,1]} \varphi d\nu \right| = \left| \int_{[0,1]} U\varphi d\nu - \int_{[0,1]} \varphi d\nu \right| = \left| \lim_{k\to\infty} \left(\int_{[0,1]} U\varphi d\mu_{\varepsilon_k} - \int_{[0,1]} \varphi d\mu_{\varepsilon_k} \right) \right| \\
&= \left| \lim_{k\to\infty} \left(\int_{[0,1]} U\varphi d\mu_{\varepsilon_k} - \int_{[0,1]} U_{\varepsilon_k} \varphi d\mu_{\varepsilon_k} \right) \right| \leq \lim_{k\to\infty} \int_{[0,1]} \|U\varphi - U_{\varepsilon_k}\varphi\|_\infty d\mu_{\varepsilon_k},
\end{aligned}
$$

which tends to zero by the claim in the beginning of the proof. This means ν is stationary, thus $\nu = \mu$ by uniqueness. ☐

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which a random variable u distributed uniformly on $[0,1]$ is defined. Define Λ_ε to be the average volume Lyapunov exponent of the perturbed process, i.e.

$$
\Lambda_\varepsilon := \sum_{i=1}^m p_i \mathbb{E} \int_{[0,1]} \log f'_{i,u,\varepsilon}(x) \mu_\varepsilon(dx).
$$

By the definition of ψ and $f_{i,u,\varepsilon}$ it holds that

$$
\sup_{u\in[0,1]} \|f_{i,u,\varepsilon} - f_i\|_\infty \to 0 \quad \text{as } \varepsilon \to 0 \tag{A.1}
$$

and

$$
\sup_{u\in[0,1]} \|f'_{i,u,\varepsilon} - f'_i\|_\infty \to 0 \quad \text{as } \varepsilon \to 0 \tag{A.2}
$$

for every $i = 1, \dots, m$. Moreover $f_{i,u,\varepsilon}$, f_i, $f'_{i,u,\varepsilon}$, f'_i are uniformly bounded from 0 and infinity provided ε is sufficiently small.

Take φ continuous on $[0,1]$. Substitution yields

$$\int_{[0,1]} e^{\varphi(x)} (f_{i,u,\varepsilon})_* \mu_\varepsilon(dx) = \int_{[0,1]} e^{\varphi(f_{i,u,\varepsilon}(x))} \mu_\varepsilon(dx) \quad \text{and}$$

$$\int_{[0,1]} e^{\varphi(x)} (f_i)_* \mu(dx) = \int_{[0,1]} e^{\varphi(f_i(x))} \mu(dx).$$

Further,

$$\begin{aligned} &\Big| \int_{[0,1]} e^{\varphi(f_{i,u,\varepsilon}(x))} \mu_\varepsilon(dx) - \int_{[0,1]} e^{\varphi(f_i(x))} \mu(dx) \Big| \\ &\quad \le \int_{[0,1]} \big| e^{\varphi(f_{i,u,\varepsilon}(x))} - e^{\varphi(f_i(x))} \big| \mu_\varepsilon(dx) \\ &\quad + \Big| \int_{[0,1]} e^{\varphi(f_i(x))} \mu_\varepsilon(dx) - \int_{[0,1]} e^{\varphi(f_i(x))} \mu(dx) \Big| \end{aligned} \tag{A.3}$$

The first summand tends to 0 as $\varepsilon \to 0$ by (A.1) and the uniform continuity of φ. The second summand tends to 0 by the weak-$*$ convergence of μ_ε to μ. By the same reason the whole expression

$$\ln \int_{[0,1]} e^{\varphi(x)} (f_{i,u,\varepsilon})_* \mu_\varepsilon(dx) - \int_{[0,1]} \varphi(x) \mu_\varepsilon(dx) \tag{A.4}$$

tends to

$$\ln \int_{[0,1]} e^{\varphi(x)} (f_i)_* \mu(dx) - \int_{[0,1]} \varphi(x) \mu(dx) \tag{A.5}$$

as $\varepsilon \to 0$.

The entropies $h((f_{i,u,\varepsilon})_* \mu_\varepsilon | \mu_\varepsilon)$ and $h((f_i)_* \mu | \mu)$ are the supremum of (A.4) and (A.5), respectively, over all continuous functions φ on $[0,1]$. Take $\delta > 0$ and continuous φ with

$$h((f_i)_* \mu | \mu) - \delta < \ln \int_{[0,1]} e^{\varphi(x)} (f_{i,u,\varepsilon})_* \mu(dx) - \int_{[0,1]} \varphi(x) \mu(dx).$$

We have

$$\begin{aligned} \liminf_{\varepsilon \to 0} \mathbb{E} h((f_{i,u,\varepsilon})_* \mu_\varepsilon | \mu_\varepsilon) &\ge \liminf_{\varepsilon \to 0} \mathbb{E} \ln \int_{[0,1]} e^{\varphi(x)} (f_{i,u,\varepsilon})_* \mu_\varepsilon(dx) - \int_{[0,1]} \varphi(x) \mu_\varepsilon(dx) \\ &= \mathbb{E} \ln \int_{[0,1]} e^{\varphi(x)} (f_i)_* \mu(dx) - \int_{[0,1]} \varphi(x) \mu(dx) \ge h((f_i)_* \mu | \mu) - \delta, \end{aligned}$$

thus

$$\liminf_{\varepsilon \to 0} \mathbb{E} h((f_{i,u,\varepsilon})_* \mu_\varepsilon | \mu_\varepsilon) \ge h((f_i)_* \mu | \mu).$$

Much simpler reasoning[2] yields $\Lambda_\varepsilon \to \Lambda$.

We are in position to make the final computation, i.e. to show that $\sum_{i=1}^m p_i \mathbb{E} h(f_{i,u,\varepsilon})_* \mu_\varepsilon | \mu_\varepsilon) = -\Lambda_\varepsilon$ for every $\varepsilon > 0$. Let us recall that φ_ε denotes the density of μ_ε. As we have proven, φ_ε is bounded. For any $u \in [0,1]$ the measure $(f_{i,u,\varepsilon})_* \mu_\varepsilon$ is absolutely continuous with respect to the Lebesgue measure with the density

[2] We just need to use (A.2) and rewrite (A.3)

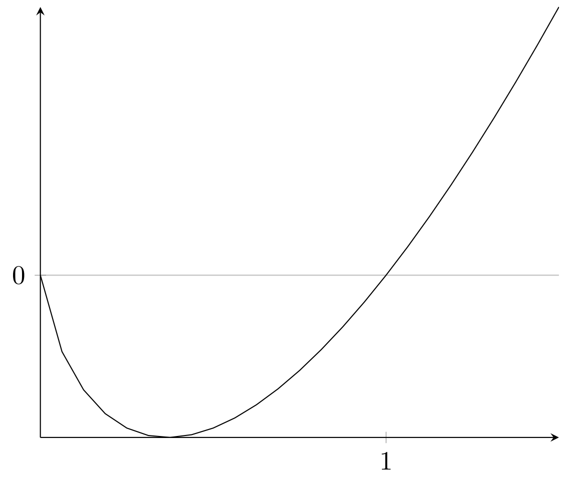

Figure A.6: The graph of $x \log x$.

$$\frac{\varphi_\varepsilon(f_{i,u,\varepsilon}^{-1}(x))}{f_{i,u,\varepsilon}'(f_{i,u,\varepsilon}^{-1}(x))}$$

(it is just integration by substitution). The measure above is supported on the image of $f_{i,u,\varepsilon}$, which is a compact subset of $(0,1)$, and has a bounded density (φ_ε is bounded). As we have proven, φ_ε is continuous and positive on $(0,1)$. These facts combined yield that $(f_{i,u,\varepsilon})_*\mu_\varepsilon$ is absolutely continuous with respect to μ_ε and the density

$$\frac{\varphi_\varepsilon(f_{i,u,\varepsilon}^{-1}(x))}{\varphi_\varepsilon(x) f_{i,u,\varepsilon}'(f_{i,u,\varepsilon}^{-1}(x))}$$

is bounded. Therefore

$$\frac{\varphi_\varepsilon(f_{i,u,\varepsilon}^{-1}(x))}{\varphi_\varepsilon(x) f_{i,u,\varepsilon}'(f_{i,u,\varepsilon}^{-1}(x))} \log\left(\frac{\varphi_\varepsilon(f_{i,u,\varepsilon}^{-1}(x))}{\varphi_\varepsilon(x) f_{i,u,\varepsilon}'(f_{i,u,\varepsilon}^{-1}(x))}\right)$$

is bounded (see the plot in Figure A.6) and, as a consequence, integrable with respect to μ_ε. By what was mentioned in the beginning of the section the relative entropy of $(f_{i,u,\varepsilon})_*\mu_\varepsilon$ with respect to μ_ε is given by

$$\begin{aligned} h((f_{i,u,\varepsilon})_*\mu_\varepsilon|\mu_\varepsilon) &= \int_{(0,1)} \log \frac{\varphi_\varepsilon(f_{i,u,\varepsilon}^{-1}(x))}{\varphi_\varepsilon(x) f_{i,u,\varepsilon}'(f_{i,u,\varepsilon}^{-1}(x))} (f_{i,u,\varepsilon})_*\mu_\varepsilon(dx) \\ &= \int_{(0,1)} \log \frac{\varphi_\varepsilon(x)}{\varphi_\varepsilon(f_{i,u,\varepsilon}(x)) f_{i,u,\varepsilon}'(x)} \mu_\varepsilon(dx) \\ &= \int_{(0,1)} \Big(\log\varphi_\varepsilon(x) - \log\varphi_\varepsilon(f_{i,u,\varepsilon}(x)) - \log f_{i,u,\varepsilon}'(x) \Big) \mu_\varepsilon(dx). \end{aligned}$$

The function is μ_ε integrable. Hence taking the expectation with respect to u and i we can use the Fubini theorem and write

$$\sum_{i=1}^m p_i \mathbb{E} h((f_{i,u,\varepsilon})_*\mu_\varepsilon|\mu_\varepsilon)$$

$$= \int_{(0,1)} \left(\sum_{i=1}^{m} p_i \mathbb{E} \log \varphi_\varepsilon(x) - \sum_{i=1}^{m} p_i \mathbb{E} \log \varphi_\varepsilon(f_{i,u,\varepsilon}(x)) - \sum_{i=1}^{m} p_i \mathbb{E} \log f'_{i,u,\varepsilon}(x) \right) \mu_\varepsilon(dx).$$

Now

$$\sum_{i=1}^{m} p_i \mathbb{E} \log \varphi_\varepsilon(x) = \log \varphi_\varepsilon(x),$$

$$\sum_{i=1}^{m} p_i \mathbb{E} \log \varphi_\varepsilon(f_{i,u,\varepsilon}(x)) = U_\varepsilon \log \varphi_\varepsilon(x) \text{ and}$$

$$\int_{(0,1)} \sum_{i=1}^{m} p_i \mathbb{E} \log f'_{i,u,\varepsilon}(x) \mu_\varepsilon(dx) = -\Lambda_\varepsilon.$$

By the stationarity of μ_ε

$$\sum_{i=1}^{m} p_i \mathbb{E} h((f_{i,u,\varepsilon})_* \mu_\varepsilon | \mu_\varepsilon) = \int_{(0,1)} \log \varphi_\varepsilon(x) \mu_\varepsilon(dx) - \int_{(0,1)} U_\varepsilon \log \varphi_\varepsilon(x) \mu_\varepsilon(dx) - \Lambda_\varepsilon = -\Lambda_\varepsilon,$$

which completes the proof.

Bibliography

[AM14] L. Alsedà and M. Misiurewicz. Random interval homeomorphisms. *Publ. Mat.*, 58(suppl.):15–36, 2014.

[AS70] R. Abraham and S. Smale. Nongenericity of Ω-stability. In *Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968)*, pages 5–8. Amer. Math. Soc., Providence, R.I., 1970.

[Bax89] P. H. Baxendale. Lyapunov exponents and relative entropy for a stochastic flow of diffeomorphisms. *Probab. Theory Related Fields*, 81(4):521–554, 1989.

[BBS20] S. Brofferio, D. Buraczewski, and T. Szarek. On uniqueness of invariant measures for random walks on Homeo($\mathbb{R}$), 2020. Accepted in *Ergodic theory and Dynamical Systems.*

[BDEG88] M. F. Barnsley, S. G. Demko, J. H. Elton, and J. S. Geronimo. Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities. *Ann. Inst. H. Poincaré Probab. Statist.*, 24(3):367–394, 1988.

[BDEG89] M. F. Barnsley, S. G. Demko, J. H. Elton, and J. S. Geronimo. Erratum: "Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities" [Ann. Inst. H. Poincaré Probab. Statist. **24** (1988), no. 3, 367–394; MR0971099 (89k:60088)]. *Ann. Inst. H. Poincaré Probab. Statist.*, 25(4):589–590, 1989.

[BFV19] S. Berghout, R. Fernández, and E. Verbitskiy. On the relation between Gibbs and g-measures. *Ergodic Theory Dynam. Systems*, 39(12):3224–3249, 2019.

[Bil61] P. Billingsley. The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.*, 12:788–792, 1961.

[Bil95] P. Billingsley. *Probability and measure.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[Bil99] P. Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[BM53] R. R. Bush and F. Mosteller. A stochastic model with applications to learning. *Ann. Math. Statistics*, 24:559–585, 1953.

[BM08] A. Bonifant and J. Milnor. Schwarzian derivatives and cylinder maps. In *Holomorphic dynamics and renormalization*, volume 53 of *Fields Inst. Commun.*, pages 1–21. Amer. Math. Soc., Providence, RI, 2008.

[BR21] J. Bradík and S. Roth. Typical behaviour of random interval homeomorphisms, 2021. https://arxiv.org/pdf/2102.02023.pdf.

[Bro71] B. M. Brown. Martingale central limit theorems. *Ann. Math. Statist.*, 42:59–66, 1971.

[BS21] K. Barański and A. Śpiewak. Singular stationary measures for random piecewise affine interval homeomorphisms. *J. Dynam. Differential Equations*, 33(1):345–393, 2021.

[CP15] S. Crovisier and R. Potrie. Introduction to partially hyperbolic dynamics. https://www.imo.universite-paris-saclay.fr/ crovisie/00-CP-Trieste-Version1.pdf, 2015.

[Cra90] H. Crauel. Extremal exponents of random dynamical systems do not vanish. *J. Dynam. Differential Equations*, 2(3):245–291, 1990.

[CS20a] W. Czernous and T. Szarek. Generic invariant measures for iterated systems of interval homeomorphisms. *Arch. Math. (Basel)*, 114(4):445–455, 2020.

[CS20b] K. Czudek and T. Szarek. Ergodicity and central limit theorem for random interval homeomorphisms. *Israel J. Math.*, 239(1):75–98, 2020.

[CWSS20] K. Czudek, H. Wojewódka-Ściążko, and T. Szarek. The law of the iterated logarithm for random interval homeomorphisms. To appear in *Israel Journal of Mathematics*, 2020.

[Czu20] K. Czudek. Alsedà-Misiurewicz systems with place-dependent probabilities. *Nonlinearity*, 33(11):6221–6243, 2020.

[Den89] M. Denker. The central limit theorem for dynamical systems. In *Dynamical systems and ergodic theory (Warsaw, 1986)*, volume 23 of *Banach Center Publ.*, pages 33–62. PWN, Warsaw, 1989.

[DF66] L. E. Dubins and D. A. Freedman. Invariant probabilities for certain Markov processes. *Ann. Math. Statist.*, 37:837–848, 1966.

[DF99] P. Diaconis and D. Freedman. Iterated random functions. *SIAM Rev.*, 41(1):45–76, 1999.

[DG12] L. J. Díaz and K. Gelfert. Porcupine-like horseshoes: transitivity, Lyapunov spectrum, and phase transitions. *Fund. Math.*, 216(1):55–100, 2012.

[DHN85] S. Demko, L. Hodges, and B. Naylor. Constructions of fractal objects with iterated function systems. *ACM SIGGRAPH Computer Graphics*, 19(3), 1985.

[DKN07] B. Deroin, V. Kleptsyn, and A. Navas. Sur la dynamique unidimensionnelle en régularité intermédiaire. *Acta Math.*, 199(2):199–262, 2007.

[DKNP13] B. Deroin, V. Kleptsyn, A. Navas, and K. Parwani. Symmetric random walks on $\mathrm{Homeo}^+(\mathbf{R})$. *Ann. Probab.*, 41(3B):2066–2089, 2013.

[DL03] Y. Derriennic and M. Lin. The central limit theorem for Markov chains started at a point. *Probab. Theory Related Fields*, 125(1):73–76, 2003.

[Dol08] D. Dolgopyat. Limit theorems for hyperbolic systems. https://www.math.umd.edu/ dolgop/LimLect.pdf, 2008.

[Don51] M. D. Donsker. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.*, 6:12, 1951.

[DR37] W. Doeblin and Fortet R. Sur des chaines à liaisons complètes. *Bull. Soc. Math. de France*, 65:132–148, 1937.

[DS86] P. Diaconis and M. Shahshahani. Products of random matrices and computer image generation. In *Random matrices and their applications (Brunswick, Maine, 1984)*, volume 50 of *Contemp. Math.*, pages 173–182. Amer. Math. Soc., Providence, RI, 1986.

[DSL15] J. De Simoi and C. Liverani. The martingale approach after Varadhan and Dolgopyat. In *Hyperbolic dynamics, fluctuations and large deviations*, volume 89 of *Proc. Sympos. Pure Math.*, pages 311–339. Amer. Math. Soc., Providence, RI, 2015.

[DV75] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I. II. *Comm. Pure Appl. Math.*, 28, 1975.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[GH17] M. Gharaei and A. J. Homburg. Random interval diffeomorphisms. *Discrete Contin. Dyn. Syst. Ser. S*, 10(2):241–272, 2017.

[GK20] A. Gorodetski and V. Kleptsyn. Parametric Furstenberg theorem on random products of $SL(2,\mathbb{R})$ matrices, 2020.

[GL78] M. I. Gordin and B. A. Lifšic. Central limit theorem for stationary Markov processes. *Dokl. Akad. Nauk SSSR*, 239(4):766–767, 1978.

[Gor15] A. Gordenko. Nonsymmetric random dynamics on the real line, 2015. Unpublished manuscript, in Russian.

[Gor20] A. Gordenko. Random dynamical systems on a real line, 2020. https://arxiv.org/pdf/2009.14686.pdf.

[Gou15] S. Gouëzel. Limit theorems in dynamical systems using the spectral method. In *Hyperbolic dynamics, fluctuations and large deviations*, volume 89 of *Proc. Sympos. Pure Math.*, pages 161–193. Amer. Math. Soc., Providence, RI, 2015.

[Hai02] M. Hairer. Exponential mixing properties of stochastic PDEs through asymptotic coupling. *Probab. Theory Related Fields*, 124(3):345–380, 2002.

[HH80] P. Hall and C. C. Heyde. *Martingale limit theory and its application.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980. Probability and Mathematical Statistics.

[Hut81] J. E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30(5):713–747, 1981.

[HZ07] A. J. Homburg and H. Zmarrou. Bifurcations of stationary measures of random diffeomorphisms. *Ergodic Theory Dynam. Systems*, 27(5):1651–1692, 2007.

[IKS08] Y. S. Ilyashenko, V. A. Kleptsyn, and P. Saltykov. Openness of the set of boundary preserving maps of an annulus with intermingled attracting basins. *J. Fixed Point Theory Appl.*, 3(2):449–463, 2008.

[Ily08] Yu. S. Ilyashenko. Diffeomorphisms with intermingled attracting basins. *Funktsional. Anal. i Prilozhen.*, 42(4):60–71, 112, 2008.

[ITM48] C. Ionescu Tulcea and G. Marinescu. Sur certaines chaînes à liaisons complètes. *C. R. Acad. Sci. Paris*, 227:667–669, 1948.

[ITM50] C. Ionescu Tulcea and G. Marinescu. Théorie ergodique pour des classes d'opérations non complètement continues. *Ann. of Math. (2)*, 52:140–147, 1950.

[Kal02] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[Kan94] I. Kan. Open sets of diffeomorphisms having two attractors, each with an everywhere dense basin. *Bull. Amer. Math. Soc. (N.S.)*, 31(1):68–74, 1994.

[Kar53] S. Karlin. Some random walks arising in learning models. I. *Pacific J. Math.*, 3:725–756, 1953.

[Kea72] M. Keane. Strongly mixing g-measures. *Invent. Math.*, 16:309–324, 1972.

[KH95] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.

[Led86] F. Ledrappier. Positivity of the exponent for stationary sequences of matrices. In *Lyapunov exponents (Bremen, 1984)*, volume 1186 of *Lecture Notes in Math.*, pages 56–73. Springer, Berlin, 1986.

[Liv96] 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.

[Lor63] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmospheric Sci.*, 20(2):130–141, 1963.

[Lor66] G. G. Lorentz. *Approximation of functions.* Holt, Rinehart and Winston, New York-Chicago, Ill.-Toronto, Ont., 1966.

[LP82] É. Le Page. Théorèmes limites pour les produits de matrices aléatoires. In *Probability measures on groups (Oberwolfach, 1981)*, volume 928 of *Lecture Notes in Math.*, pages 258–303. Springer, Berlin-New York, 1982.

[Łuc21] G. Łuczyńska. Unique ergodicity for function systems on the circle. *Statistics and Probability Letters*, 173:109084, 2021.

[LY94] A. Lasota and J. A. Yorke. Lower bound technique for Markov operators and iterated function systems. *Random Comput. Dynam.*, 2(1):41–77, 1994.

[Mal17] D. Malicet. Random walks on Homeo(S^1). *Comm. Math. Phys.*, 356(3):1083–1116, 2017.

[McL74] D. L. McLeish. Dependent central limit theorems and invariance principles. *Ann. Probability*, 2:620–628, 1974.

[MW00] M. Maxwell and M. Woodroofe. Central limit theorems for additive functionals of Markov chains. *Ann. Probab.*, 28(2):713–724, 2000.

[MW05] I. Melbourne and A. Windsor. A C^∞ diffeomorphism with infinitely many intermingled basins. *Ergodic Theory Dynam. Systems*, 25(6):1951–1959, 2005.

[MY08] Y. Miao and G. Yang. The law of the iterated logarithm for additive functionals of Markov chains. *Statist. Probab. Lett.*, 78(3):265–270, 2008.

[OM35] O. Onicescu and G. Mihoc. Sur les chaines de variables statistiques. *Bulletin des Sciences Mathématiques*, 59:174–192, 1935.

[Pet89] K. Petersen. *Ergodic theory*, volume 2 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1989. Corrected reprint of the 1983 original.

[PU05] M. Peligrad and S. Utev. A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.*, 33, 2005.

[RT71] D. Ruelle and F. Takens. On the nature of turbulence. *Comm. Math. Phys.*, 20:167–192, 1971.

[Ś11] M. Ślęczka. The rate of convergence for iterated function systems. *Studia Math.*, 205(3):201–214, 2011.

[Ste12] Ö. Stenflo. A survey of average contractive iterated function systems. *J. Difference Equ. Appl.*, 18(8):1355–1380, 2012.

[SZ16] T. Szarek and A. Zdunik. Attractors and invariant measures for random interval homeomorphisms. Unpublished manuscript, 2016.

[SZ20] T. Szarek and A. Zdunik. The rate of convergence for function systems on the circle. *Colloq. Math.*, 159(1):77–89, 2020.

[SZ21] T. Szarek and A. Zdunik. The central limit theorem for iterated function systems on the circle. *Mosc. Math. J.*, 21(1):175–190, 2021.

[ZW08] O. Zhao and M. Woodroofe. Law of the iterated logarithm for stationary processes. *Ann. Probab.*, 36(1):127–142, 2008.