Alsedà–Misiurewicz systems with place-dependent probabilities

Klaudiusz Czudek

Klaudiusz Czudek, Institute of Mathematics Polish Academy of Sciences, Śniadeckich 8, 00-656 Warszawa, Poland
E-mail: klaudiusz.czudek@gmail.com

Received 19 November 2019, revised 12 June 2020
Accepted for publication 26 June 2020
Published 8 October 2020

Abstract
We consider systems of two specific piecewise linear homeomorphisms of the unit interval, so called Alsedà–Misiurewicz systems, and investigate the basic properties of Markov chains which arise when these two transformations are applied randomly with probabilities depending on the point of the interval. Though this iterated function system is not contracting in average and known methods do not apply, stability and the strong law of large numbers are proven.

Keywords: iterated function systems, place-dependent probabilities, Markov chains, Alsedà–Misiurewicz systems

Mathematics Subject Classification numbers: primary 37E05, 60G10, 60J05.

(Some figures may appear in colour only in the online journal)

1. Introduction

1.1. Formulation of the problem

In this paper we are interested in Markov chains which in general setting may be defined as follows: we fix a Polish space $M$, a finite number of transformations $f_i : M \to M$, $i = 1, \ldots, r$, and real nonnegative functions $p_1, \ldots, p_r : M \to \mathbb{R}$, called probabilities, with $\sum_i p_i(x) = 1$ for every point $x \in M$. Choose a starting point $x \in [0, 1]$, and pick one of the transformations randomly according to the distribution $p_1(x), \ldots, p_r(x)$. If the result is $f_i$, go to the point $x_1 := f_i(x)$.
Pick a transformation again, this time with respect to the distribution \( p_1(x_1), \ldots, p_r(x_1) \), and if the result is \( f_j \), go to the point \( x_2 := f_j(x_1) \). One continue this procedure obtaining a random walk on \( M \), which we shall denote by \( (X_n^x)_{n \geq 0} \) in order to stress that the starting point was \( x \). The starting point may be chosen also randomly with respect to some distribution \( \mu \). In that case the resulting process will be denoted by \( (X_n^\mu)_{n \geq 0} \). Systems of functions of this type we call iterated function systems with probabilities.

If the initial distribution \( \mu_\ast \) is such that the process \( (X_n^\mu_\ast)_{n \geq 0} \) is stationary, then \( \mu_\ast \) is called stationary itself. Stationary distributions play the same role as in deterministic dynamical system, and are used to describe the statistical behaviour of a random walk. More precisely, if \( \mu_\ast \) is a stationary distribution for certain random walk, then for every function \( \varphi \) in \( L^1(\mu_\ast) \) there exists a function \( \psi \) in \( L^1(\mu_\ast) \) invariant for the random walk (which means that for every starting point \( x \) the value of \( \psi(X_n^x) \) is constant and independent of \( n \)) such that

\[
\frac{1}{n} \left( \varphi(X_1^x) + \cdots + \varphi(X_n^x) \right) \to \psi(x) \text{ a.s.} \tag{1}
\]

for \( \mu_\ast \)-a.e. point \( x \) in \( M \). Moreover, we have \( \int \varphi \mathrm{d}\mu_\ast = \int \psi \mathrm{d}\mu_\ast \).

Let \( \mu_\ast \) be a stationary measure. Let us assume that it has the property: if \( A \) is an arbitrary Borel set such that \( P(X_1^x \in A) = 1 \) for every \( x \in A \), then the \( \mu_\ast \) measure of \( A \) is either 0 or 1. In that case \( \mu_\ast \) is called ergodic. Similarly to deterministic dynamical systems, every stationary measure \( \mu_\ast \) may be decomposed as a convex combination of ergodic stationary measures. For justification of all these statements see chapter I.3 in [8].

When the stationary measure exists, it is reasonable to ask if it is unique (hence whether the statistical behaviour of the Markov chain depends on a starting point). If \( \mu_\ast \) is the only stationary measure, and, moreover, the law of the process \( (X_n^\mu_\ast)_{n \geq 0} \) is convergent in the weak-* topology to \( \mu_\ast \) whatever the initial distribution is, then the Markov chain is said to be asymptotically stable. In the present paper we are interested in the existence of stationary distribution, its uniqueness and asymptotic stability of the processes defined above for certain particular transformations of the interval.

### 1.2. Historical remarks

The first paper concerning aforementioned chains was published in 1935 by Onicescu and Mihoc [36], where the authors define and investigate a slightly different type of process, but still possible to be performed in terms of iterated function systems. Two years later Doeblin and Fortet published paper [17], where, again in a slightly different language, properties of a system of two affine contractions of the interval were explored. In 1953 Karlin [28] stated a theorem that if \( M \) is the interval, \( f_1, f_2 \) are affine contractions of the interval and \( p_1, p_2 \) are arbitrary continuous functions, then the corresponding process is asymptotically stable. However, a mistake was found in his proof, and almost 50 years later the counterexample to this statement was given by Stenflo [40].

In the second half of 20th century, when iterated function systems appeared to be connected with fractals, systems of this form gained more interest, see the papers by Barnsley, Demko and fundamental paper by Hutchinson [4, 25]. The next paper we would like to mention is [3] from 1988, in which Barnsley et al considered general systems on locally compact spaces \( M \) with transformations contractive in average, i.e. satisfying the condition

\[
\sum_{i=1}^{r} p_i(x) d(f_i(x), f_i(y))^{\alpha} \leq c d(x, y)^{\alpha}
\]

for some positive constants \( c, \alpha \).
where $d$ is the metric on $M$, $c < 1$, $\alpha > 0$ are independent of $x, y$ and $x, y$ are arbitrary points of the space $M$. Obviously one needs also to assume some regularity of $p_i$, since, as it was mentioned in the previous paragraph, even when $f_i$’s are contractions it is possible to construct continuous probabilities such that the random walk has at least two distinct stationary distributions. Therefore another assumption made by the authors is the Dini continuity of $p_i$’s, which means that the modulus of continuity $\beta_i$ of $p_i$ must satisfy the property that for every $C \geq 1$ and $t < 1$ the series $\sum \beta_i(C^t)$ is convergent.

Let us note that the proof consisted of two parts. In the first it was showed that given $\varepsilon > 0$ one can find $\delta > 0$ such that if $x, y$ are points in $M$ with $d(x, y) < \delta$, then the distance between the distributions of $X^n_x$ and $X^n_y$ in the Fortet–Mourier metric is less than $\varepsilon$ for every $n$. Recall that the Fortet–Mourier metric between two measures $\mu$ and $\nu$ is defined by

$$d_{FM}(\mu, \nu) := \sup \left\{ \int \varphi \, d\mu - \int \varphi \, d\nu : \varphi \in F \right\}$$

where $F$ is the set of all Lipschitz continuous functions with Lipschitz constant less than 1 whose absolute value is less than one. In the second part of the proof, the authors fix $\varepsilon$, take $\delta$ given in the first part of the proof and couple two independent Markov chains starting from two arbitrary distributions $(X^n_x)$, $(X^n_y)$. By coupling of the processes $(X^n_x)$, $(X^n_y)$ (which are independent) we mean to show that almost surely there exists $k$ with $d(X^n_x, X^n_y) < \delta$. Given $n$, one can divide the set of all trajectories into two parts. The first part contains trajectories $\omega$ that are ‘coupled’ before $n$, i.e. such that there exists $k < n$ with $d(X^n_x(\omega), X^n_y(\omega)) < \delta$. The second part consists of the remaining trajectories. By what was already mentioned, the measure of the latter tends to zero as $n$ goes to $\infty$. By the first part of the proof and some conditioning argument, the Fortet–Mourier distance between the processes restricted to the set of ‘coupled’ trajectories is less than $\varepsilon$. A similar technique was exploited by Lasota and Yorke in 1994 [30], where the notion of ‘nonexpansiveness’ for this kind of systems was studied.

The listed papers provide basic tools in investigating iterated function systems. We would like to emphasize that several other results have been proven about systems contracting in average, see the survey by Stenflo [41]. As far as we know, the contractivity assumption is always made when the probabilities in the explored system are place-dependent. The only exception we know are systems on the circle studied by Sinai [37] and more recently by Dolgopyat et al [18], where the two transformations are rotations of angle $+\alpha$ and $-\alpha$, and the probabilities are nonvanishing and smooth. These kind of systems are examined by investigating the corresponding random walk in random environment on $\mathbb{Z}$. The papers by Kaloshin and Sinai are also related [26, 27]. For more information about iterated function systems with probabilities we would like to recommend the interesting survey by Diaconis and Freedman [13].

The case when the probability functions $p_i$’s are constant deserves more attention. On one hand, it is much simpler to study than the general case. On the other hand, it appears in dynamical systems very often as a skew-product. Iterated function systems with constant probabilities may be linked with important concepts in one-dimensional dynamics: Bernoulli convolutions [39] and Furstenberg boundary [22]. They are used in studying groups acting on the circle [10, 11, 32] and the real line [9], whose importance in mathematics comes, among others, from the Zimmer program [21]. The general overview of the theory of groups acting on the circle and real line may be found in book and article by Navas [34, 35]. Moreover, certain examples of skew-products with the Bernoulli shift in the base are investigated as models of partially hyperbolic dynamics (see [14–16, 29]).

6223
The previously described techniques based on contractivity may be also applied in the special case when probabilities are constant. However, here the analysis may simplify. For instance in [7] an elementary, very short proof of uniqueness of a stationary distribution for random walks on the interval may be found, assuming only that transformations are increasing. It is also worth to distinguish the papers by Dubins and Freedman [19], Bhattacharaya and Lee [5], where the authors deal with iterated function systems on the interval without any contraction assumption.

The publication motivating this article is by Alsedà and Misiurewicz [1] from 2014. The authors were looking for skew-products with Bernoulli shift in the base and the interval in the fibre with the property that there exists a function from the base to the fibre whose graph is an attractor for the skew-product transformation. As a byproduct, the authors showed uniqueness of the stationary measure, stability and the strong law of large numbers for the Markov chain arising from system of two very specific, piecewise linear homeomorphisms chosen with probabilities $1/2, 1/2$. Later their results were proved in the case of two $C^2$ increasing diffeomorphisms ([23]) or even arbitrary finite number of increasing homeomorphisms ([42]) satisfying mild additional assumptions. A significant hypothesis for both [23, 42] is that the points 0 and 1 should be repelling in average, which is necessary to ensure the existence of a stationary distribution. Precisely this assumption states that the average Lyapunov exponent at 0 and 1 (defined in the next subsection) should be positive.

All these results were proven only in the case of constant probabilities, so the natural question was about the case where probabilities depend on a point. To apply the elaborated techniques, one would need to construct a metric in which the system (with place-dependent probabilities) is contractive in average. Although we have tried to, we failed to construct such a metric even for the simplest system considered by Alsedà and Misiurewicz. Despite of that, straightforward analysis of this system gave the uniqueness of a stationary distribution and stability of the Markov chain, which are the main results in the present paper.

We would like to close the subsection with a short remark on the strong law of large numbers, which we also prove here. Observe that in (1) the convergence holds for $\mu_*$ almost every point $x \in M$. When the stationary measure is unique, it is natural to ask whether the convergence holds for every point of the interval. It was proved in [6] in the case of Markov chains on compact spaces, non necessarily arising from iterated function systems. Later it was proved in [12] in the case of iterated systems of contractions in $\mathbb{R}^n$ applied with constant probabilities and in [20] for systems contracting in average with place dependent probabilities on locally compact spaces. Note that importance of this theorem follows from connection to the fractals [12]. In the present article the proof of the strong law of large numbers for examined system is also provided.

### 1.3. The main results

Let $f_0$ be an interval homeomorphism such that its graph consists of two straight lines, the first one connecting $(0, 0)$ with some point $(x_0, y_0) \in (1/2, 1) \times (1/2, 1)$ under diagonal, and the second one connecting $(x_0, y_0)$ with $(1, 1)$ (see figure 1). Next, let $f_1$ be the interval homeomorphism defined by $f_1(x) = 1 - f_0(1 - x)$, $x \in [0, 1]$ (see figure 1). Setting $a_0 = \frac{x_0}{y_0}$ and $a_1 = \frac{1 - y_0}{1 - x_0}$, we can write

$$f_0(x) := \begin{cases} a_0 x & \text{if } x \leq x_0 \\ a_1 (x - 1) + 1 & \text{if } x > x_0 \end{cases} \quad \text{and} \quad f_1(x) := 1 - f_0(1 - x).$$
After [2], we call these systems the Alsedà–Misiurewicz systems (in [2] the only restriction for \((x_0, y_0) \in (0, 1) \times (0, 1)\) is that it should be under diagonal). Further, fix two positive real functions \(p_0, p_1\) on \([0, 1]\) with 
\[
p_0(x) + p_1(x) = 1 \quad \text{for every } x \in [0, 1].
\]
It defines an iterated function system with probabilities. We keep the notation from the first subsection and denote the corresponding Markov chain with the initial distribution \(\mu\) by \((X_\mu^n)_{n=0}^\infty\). In order to simplify the notation, the Markov chain starting from a point \(x \in (0, 1)\) we denote by \((X^n_x)_{n=0}^\infty\) instead of \((X_\delta_x^n)_{n=0}^\infty\).

In the previous subsection the notion of ‘average Lyapunov exponents’ appeared. In our setting we define these by the formulae

\[
\Lambda_0 := p_0(0) \log(a_0) + p_1(0) \log(a_1),
\]
\[
\Lambda_1 := p_0(1) \log(a_1) + p_1(1) \log(a_0).
\]

If the system consisted of two homeomorphism differentiable at 0 and 1, the coefficients \(a_0, a_1\) would be replaced by derivatives at 0 and 1, respectively.

The goal of our paper is to provide proofs of the existence and uniqueness of stationary measure, stability and strong law of large numbers for defined Markov chains. To this end, we introduce the following assumptions:

(A1) \(\frac{1}{2} < x_0 < 1\) and \(\frac{1}{2} \leq y_0 < x_0\),
(A2) \(p_0, p_1\) are Dini continuous,
(A3) \(0 < p_i(x) < 1\) for \(x \in [0, 1]\) and \(i = 0, 1\),
(A4) \(\Lambda_0, \Lambda_1 > 0\).

We recall that the functions \(p_0, p_1\) are Dini continuous if for every \(C \geq 0\) and \(t < 1\) we have 
\[
\sum_n \beta(C t^n) < \infty \quad \text{where } \beta \text{ denotes the modulus of continuity of } p_0, p_1, \text{ i.e.}
\]
\[
\beta(t) := \max_{i=0,1} \sup_{x \in [0,1], |h| \leq t} |p_i(x) - p_i(x + h)|.
\]
We do not need any further assumptions about contractivity of the system. All of these will be motivated in the subsection ‘1.4 sketch of the proof’. Our main results are the following theorems.

**Theorem 1.** If \((A1)-(A4)\) hold then there exists a unique Borel probability measure \(\mu_* \in \mathcal{M}\) such that the Markov chain \((X_n^\mu)\) is stationary.

**Theorem 2.** If \((A1)-(A4)\) hold, and \(\nu\) is any Borel probability measure then the Markov chain \((X_n^\nu)\) is asymptotically stable.

**Theorem 3 (the strong law of large numbers).** If \((A1)-(A4)\) hold, \(x \in (0,1), \varphi \in C((0,1))\) then
\[
\frac{\varphi(X_1^x) + \cdots + \varphi(X_n^x)}{n} \to \int \varphi \, d\mu_* \text{ a.s.}
\]

2. Notation, sketch of the proof and open problems

2.1. Notation

The space of Borel probability measures on \((0,1)\) will be denoted by \(\mathcal{M}_1\) and the space of all positive Borel measures by \(\mathcal{M}\). The family of transition probabilities \(p(x, \cdot) \in \mathcal{M}_1, x \in (0,1)\) is defined by the formula
\[
p(x, \cdot):= p_0(x)\delta_{f_0(x)} + p_1(x)\delta_{f_1(x)} \text{ for } x \in (0,1).
\]

Let us choose an initial distribution \(\mu \in \mathcal{M}_1\). Together with the transition probabilities it defines the Markov chain \((X^\mu_n)\) on \((0,1)\). The canonical space for this Markov chain is constructed as follows. Put \(\Omega = (0,1)^\infty, \mathcal{G} = \mathcal{B}(0,1)^\infty\). Here \(\mathcal{B}(0,1)\) stands for the \(\sigma\)-algebra of Borel subsets of \((0,1)\). We define the family of measures \(\mathbb{P}_x^\infty\), \(x \in (0,1)\) on \((\Omega, \mathcal{G})\) by giving its values on cylinder sets, i.e.

\[
\mathbb{P}_x^\infty(A_1 \times \cdots \times A_k \times (0,1)^\infty) := \int_{A_1} p(x, \, dx_1) \int_{A_2} p(x_1, \, dx_2) \cdots \int_{A_k} p(x_{k-1}, \, dx_k)
\]

where \(A_1, \ldots, A_k \in \mathcal{B}(0,1), x \in (0,1)\). Existence of the unique extension to a measure on \(\mathcal{G}\) follows from the Kolmogorov extension theorem. Fix the initial distribution \(\nu \in \mathcal{M}_1\) and define the measures \(\mathbb{P}_x^\nu\) on cylinders by

\[
\mathbb{P}_x^\nu(A \times B) := \int_A \mathbb{P}_x^\infty(B) \nu(dx),
\]

for \(A \in \mathcal{B}(0,1), B \in \mathcal{G}\). This measure has the unique extension to \(\mathcal{G}\) by the Kolmogorov extension theorem. Now the sequence \((\pi_n)\) of projections defined on \((\Omega, \mathcal{G}, \mathbb{P}_x^\infty)\) by \(\pi_n(x_1, x_2, \ldots) := x_n, n \geq 1\), is the canonical realization of the Markov chain \((X_n^\mu)\).

The processes \((X_n^x), x \in (0,1)\), may be also realized on the space \(\Sigma = \{0,1\}^\mathbb{N}\) with the standard product \(\sigma\)-algebra \(\mathcal{F}\) and the probability measure \(\mathbb{P}_x\), defined on cylinders, \(C_{i_1, \ldots, i_k} = \{\omega \in \Sigma : \omega_1 = i_1, \ldots, \omega_k = i_k\}\), by

\[
\mathbb{P}_x(C_{i_1, \ldots, i_k}) := p_{i_1}(x)p_{i_2}(f_{i_1}(x)) \cdots p_{i_k}(f_{i_{k-1}} \circ \cdots \circ f_{i_1}(x)).
\]
Then it is clear that $f^n_\omega(x) := f_{\omega_n} \circ \cdots \circ f_{\omega_1}(x)$, where $\omega = (\omega_1, \omega_2, \ldots)$, is a realization of $(X^n_\alpha)$. Expectation with respect to $P_n$ is denoted by $E_n$. By $\theta_n$ we denote the shift $\theta_n : \Sigma \rightarrow \Sigma$, $\theta_n(\omega) := (\omega_{n+1}, \omega_{n+2}, \ldots)$, where $\omega = (\omega_1, \omega_2, \ldots)$. For $n \geq 1$ and $\omega \in \Sigma$ put

$$a^n_\omega := a_{\omega_n} \cdots a_{\omega_1}.\]

In order to describe the evolution of $(X^n_\alpha)$ we introduce the Markov–Feller operator $P : \mathcal{M} \rightarrow \mathcal{M}$ by

$$P\mu(A) := \int f_{\omega}^{-1}(A) p_0(x)\mu(dx) + \int f_{\omega}^{-1}(A) p_1(x)\mu(dx),$$

for $A \in \mathcal{B}(0,1), \mu \in \mathcal{M}$. Its predual operator $U : C(0,1) \rightarrow C(0,1)$ is given by

$$U\varphi(x) := p_0(x)\varphi(f_0(x)) + p_1(x)\varphi(f_1(x)),$$

for $\varphi \in C(0,1)$ and $x \in (0,1)$. By ‘predual’ we mean that

$$\int_{(0,1)} \varphi dP\mu = \int_{(0,1)} U\varphi d\mu$$

for every $\mu \in \mathcal{M}$ and $\varphi \in C(0,1)$. The operator $P$ is linear, i.e. $P(\lambda_1 \mu_1 + \lambda_2 \mu_2) = \lambda_1 P\mu_1 + \lambda_2 P\mu_2$ for $\lambda_1, \lambda_2 \geq 0, \mu_1, \mu_2 \in \mathcal{M}$. It also preserves the total mass of a measure, i.e. $P\mu((0,1)) = \mu((0,1))$ for $\mu \in \mathcal{M}$. We say that a measure $\mu_n \in \mathcal{M}$ is invariant for the operator $P$ if $P\mu_n = \mu_n$. In that case we say that the operator $P$ is asymptotically stable if $P^\nu \mu \rightharpoonup \mu_n$ weakly for every $\nu \in M_1$.

The Markov–Feller operator $P$ has the property that the distribution of $X^n_\alpha$ is $P^n\mu_n$ for all $n \geq 0$ and $\mu \in \mathcal{M}$. Therefore one can choose an initial distribution $\mu \in M_1$ such that $(X^n_\alpha)$ is stationary if and only if $\mu$ is $P$-invariant and the Markov chain $(X^n_\alpha)$ is stable if and only if $P$ is asymptotically stable.

Following [23] we define

$$\mathcal{P}_{M,\alpha} := \{\mu \in M_1 : \mu((0,x)) \leq Mx^\alpha \text{ and } \mu((1-x,1)) \leq Mx^\alpha \text{ for all } x \in (0,1)\}.\]

By what we have just mentioned, theorem 1 is equivalent to the existence of a unique invariant probability measure for the Markov–Feller operator $P$.

### 2.2. Sketches of the proofs

The proof of existence of a stationary distribution is analogous to the proofs given in [23, 42]. The idea is to apply the Krylov–Bogoliubov technique: for a fixed $x \in (0,1)$, we define the sequence of measures $\frac{1}{P}(\delta_x + \cdots + P^{n-1}\delta_x)$, $n \in N$, and we show the existence of an accumulation point, which is the desired distribution. The obstacle is the measures in the above sequence are supported on the noncompact set $(0,1)$, therefore the Prokhorov lemma cannot be used. In order to deal with that problem, we use the assumption on Lyapunov exponents to find parameters $M, \alpha$ for which the class $\mathcal{P}_{M,\alpha}$ is $P$-invariant. As every such class is weakly-$*$ compact and convex, taking $\delta_x \in \mathcal{P}_{M,\alpha}$ shall complete the proof. Here assumption (A4) is essential.

The proof of uniqueness is much more complicated. The main goal is to show that taking two points $x, y$ close to each other the processes $X^n_\alpha, X^n_\beta$ have close distributions. Formally, we need to show that $(U^\alpha\varphi)_{\alpha \geq 1}$ is an equicontinuous family on $[1 - x_0, x_0]$ for any bounded Lipschitz function $\varphi$. Note that this is a similar approach as in [3, 30]. After proving that, the
Nonlinearity 33 (2020) 6221 K Czudek

assertion is a simple consequence of the ergodic theorem, as there exists a point \( c \) belonging to the support of every stationary measure. In order to show the equicontinuity of \((U^n \varphi)_{n \geq 1}\), the first aim is to justify that for any point \( x \in [1 - x_0, x_0] \) and any point \( y \) sufficiently close to \( x \) we have \( E_x | f^n(x) - f^n(y) | \leq Cq^n \) (proposition 2) for some \( q < 1 \), which is independent of \( x \). Then uncomplicated estimations using assumption (A2) yield the desired equicontinuity. It is interesting that, on one hand, assumption (A2) is exactly what we need and, on the other hand, it was also essential in papers [3, 30]. Thus, it seems to be optimal.

A crucial step in proving proposition 2 is proposition 1, which says that for any two points \( x, y \) sufficiently close to each other the distance \( | f^n(x) - f^n(y) | \) is not greater than \( a_1 | x - y | \) for any choice of \( \omega \) and \( n \). To conclude the first from the latter, we notice that the system is contracting on \([1 - x_0, x_0]\) whatever the probabilities are. Therefore, if the points \( f^n(x) \) and \( f^n(y) \) belong together to \([1 - x_0, x_0]\) for a sufficiently long time, the distance between them is contracted by a coefficient less than \( 1/2a_1 \). Because of proposition 1, one just needs to show that the probability of the low number of long stays in \([1 - x_0, x_0]\) up to \( n \) decays exponentially fast when \( n \) goes to infinity. For the trajectories with a sufficient amount of visits, the distance between \( f^n(x) \) and \( f^n(y) \) is contracted by a coefficient less that \( 1/2^{|n|} \). The equicontinuity of \((U^n \varphi)_{n \geq 1}\) may be concluded. The proof of stability is completed by an argument similar to this one used in [3, 30].

Showing the exponential convergence (proposition 2) is the key part of the proof of uniqueness. It strongly relies on proposition 1, which cannot be proven for any Alseda–Misiurewicz system beyond these satisfying (A1), possibly with some exceptions. The exceptions, as we suppose, are exactly the resonant cases described in the next subsection. Therefore (A1) is crucial in our technique, however, it is interesting question if uniqueness may be proven without it (see problem 1 in section 2.3).

To show the strong law of large numbers, we use idea of Elton [20] to prove that \( P_x, P_y \) are absolutely continuous to each other provided that \( x \) and \( y \) are sufficiently close. We pick a point \( z \) generic for the Birkhoff ergodic theorem. If \( x \) is sufficiently close to \( z \), then, by absolute continuity of \( P_x \) and \( P_z \) and by proposition 1, almost every trajectory of \((X^n)\) differs from the generic one by at most \( \varepsilon \). Showing that the process \((X^n)\) visits arbitrary neighbourhood of \( z \) infinitely many times (whatever \( x \) is) yields that for every \( \varepsilon > 0 \) almost every trajectory of \((X^n)\) differs from the generic one by at most \( \varepsilon \). Then simple and standard argument completes the proof of the strong law of large numbers.

At the end we would like to note that the assumption (A3) is very convenient and not restrictive at the same time. It appears all the time in the reasoning. When dropping this assumption, it is very easy to construct a system with at least two distinct stationary distributions (by constructing two intervals such that the probability of transition from one to another is equal to zero).

2.3. Open problems

As we explained, all assumptions are important in our reasoning. However, while (A2)–(A4) are essential, it is interesting if one can generalize the results to the case when assumption (A1) is dropped. The main problem is that we cannot prove proposition 2 without proposition 1, and proposition 1 is not true for general systems of homeomorphisms. Even restricting oneself to the Alseda–Misiurewicz systems, it is not true without (A1). The following question is open.

**Problem 1.** Is it possible to show proposition 2 for systems without assumption (A1)? If not, then it is possible to show the uniqueness of a stationary measure for such systems? What with more general systems, like of \( C^2 \) diffeomorphisms?
Figure 2. The order of the points $1 - x_0, 1 - y_0, y_0, x_0$.

Our method of proving theorem 2 does not provide any rate of convergence of $U^n \varphi$ to $\int \varphi \, d\mu$. In the case of systems contracting in average it was shown in [38] using coupling techniques that the rate of convergence is exponential provided that probabilities are Lipschitz continuous. This may be exploited to show the central limit theorem (see [24]). However, the central limit theorem holds also with the rate of convergence weaker than exponential [33]. The problem is if one can provide a sufficient rate of convergence to imply the central limit theorem.

Problem 2. Does the central limit theorem hold for the Markov chain $(X_\mu^n)$ provided that probabilities are Lipschitz continuous? Does it hold for the process $(X_\mu^n)$ for $\mu$-almost every starting point $x$?

The last question is connected with paper [2]. It was noticed in [23, 42] that the condition $\frac{\log f_0'(0)}{\log f_1'(0)} \notin \mathbb{Q}$ implies that the system is minimal. The authors of [2] call the systems for which $\frac{\log f_0'(0)}{\log f_1'(0)} \in \mathbb{Q}$ the systems with resonance and prove that if a system is with resonance, $y_0 < 1/2$ and some additional more technical assumptions are satisfied, then there exists an invariant Cantor sets for the iterated function system $(f_0, f_1)$ (also other interesting results are established there, like the value of the Hausdorff dimension of the invariant set). However, nothing is known in the case $y_0 > 1/2$ which is our case.

Problem 3. If the Alsedà–Misiurewicz system satisfies (A1) then is it necessarily minimal?

3. The proof of theorem 1

Proof of existence. The proof follows the lines of the proof from [23] with necessary changes. Namely, we shall show that there exist parameters $M \geq 1, \alpha \in (0, 1)$ such that the class $\mathcal{P}_{M, \alpha}$ is invariant under the operator $P$. It is sufficient, since in that case one can apply the standard Krylov–Bogoliubov technique, i.e. take any $\nu \in \mathcal{P}_{M, \alpha}$ and define $\nu_0 = \frac{1}{2}(\nu + \cdots + P^{n-1}\nu)$. By the $P$-invariance of $\mathcal{P}_{M, \alpha}$, all $\nu_j$'s are in $\mathcal{P}_{M, \alpha}$, and by weak-* compactness of $\mathcal{P}_{M, \alpha}$ there exists an accumulation point $\mu_\ast \in \mathcal{P}_{M, \alpha}$ of this sequence which is an invariant measure. Details are left to the reader. What remains to show is the existence of parameters $M, \alpha$ with the desired property.

By the continuity of $p_0, p_1$ at the boundary, (A4) and (2) one can find $0 < \varepsilon < 1 - x_0$ such that

$$\max_{t \leq \varepsilon} p_0(t) \log a_0 + \max_{t \leq \varepsilon} p_1(t) \log a_1 > \frac{\Lambda_0}{2},$$

$$\max_{t \leq \varepsilon} p_0(1 - t) \log a_1 + \max_{t \leq \varepsilon} p_1(1 - t) \log a_0 > \frac{\Lambda_1}{2}. \quad (3)$$
Proposition 1.

Writing the Taylor formula of the function \( a \mapsto a^{-\alpha} \) at 0 we obtain \( a^{-\alpha} = 1 - \alpha \log a + o(\alpha) \) where \( a \) is any fixed positive number. By this formula and (3) one can find \( \alpha \in (0, 1) \) and \( p \in (0, 1) \) with

\[
\begin{align*}
\max_{t \in \mathbb{C}} p_0(t)a_0^{-\alpha} + \max_{t \in \mathbb{C}} p_1(t)a_1^{-\alpha} &< p, \\
\max_{t \in \mathbb{C}} p_0(1-t)a_1^{-\alpha} + \max_{t \in \mathbb{C}} p_1(1-t)a_0^{-\alpha} &< p.
\end{align*}
\]

(4)

Eventually, put \( M \) to be any number greater or equal than \((a_0\varepsilon)^{-\alpha} > \varepsilon^{-\alpha} > 1\).

We are in position to show the invariance of \( \mathcal{P}_{M, \alpha} \) for \( M, \alpha \) chosen above. Take \( \mu \in \mathcal{P}_{M, \alpha} \) and \( x \in (0, 1) \). If \( x \geq a_0 \) then \( Mx^\alpha \geq M(a_0\varepsilon)^\alpha \geq 1 \), hence the condition \( P\mu((0, x)) \leq Mx^\alpha \) is trivially satisfied. If \( x < a_0 \) then also \( x < 1 - x_0 \) and

\[
P\mu((0, x)) = \int_{[0,a_0^{-1}x]} p_0(t)\mu(dt) + \int_{[a_0^{-1}x, 1]} p_1(t)\mu(dt)
\]

\[
\leq \max_{t \in \mathbb{C}} p_0(t)Ma_0^{-\alpha}x^\alpha + \max_{t \in \mathbb{C}} p_1(t)Ma_1^{-\alpha}x^\alpha < Mx^\alpha p < Mx^\alpha,
\]

where in the last line we used (4). Therefore \( P\mu((0, x)) \leq Mx^\alpha \). The proof that \( P\mu((1 - x, 1)) \leq Mx^\alpha \) is analogous. The invariance of \( \mathcal{P}_{M, \alpha} \) is established.

We are now going to make some use of (A1) (cf. figure 2). Take \( \eta_1 \) such that the following condition is satisfied

\[
a_0y - a_1(y - a_1\eta_1) < 0 \quad \text{for} \quad y \geq 1 - y_0.
\]

(5)

There exists such \( \eta_1 \). Indeed, since \( a_0 < 1 < a_1 \), the linear function \( y \mapsto a_0y - a_1(y - a_1\eta_1) \) is decreasing and, in consequence, it suffices to show that there exists such \( \eta_1 \) for \( y = 1 - y_0 \). But since

\[
a_0(1 - y_0) - a_1((1 - y_0) - a_1\eta_1) = (1 - y_0)(a_0 - a_1) + a_1^2\eta_1,
\]

it just follows from the inequality \( (1 - y_0)(a_0 - a_1) < 0 \). Let us also assume that \( \eta_1 \) is less than the length of the interval \( [1 - x_0, 1 - y_0] \) and satisfies

\[
f_1(y_0 + a_1\eta_1) < x_0.
\]

(6)

This is possible by the continuity of \( f_1 \) and

\[
f_1(y_0) < x_0.
\]

(7)

To show this, however, we compute \( f_1(y_0) = \frac{-y_0(1-y_0)}{y_0} + 1 \) and obtain that (7) is equivalent to the condition \( y_0(1 - y_0) > x_0(1 - x_0) \). By the assumptions made on \( x_0, y_0 \) we have \( 1/2 \leq y_0 < x_0 \), so our statement follows from the monotonicity of the function \( \psi(t) := t(1 - t) \) on \( [1/2, 1] \).

**Proposition 1.** If \( x, y \in [1 - x_0, x_0] \) and \( |x - y| \leq \eta_1 \) then

\[
|f_n^\omega(x) - f_n^\omega(y)| \leq a_1|x - y|
\]

for every \( n \) and every \( \omega \in \Sigma \).
In order to simplify the reasoning, we assume that \( x < y \) and \( \omega \) is such that \( f^n(x) \) visits \((0, 1 - x_0)\) infinitely often and \( f^n(y) \) visits \((x_0, 1)\) infinitely often. At the end of the proof we will give a simple explanation that this assumption may be dropped.

**Lemma 1.** If \( 1 - x_0 \leq x < y \leq x_0, |x - y| < a_1 \eta_1, y > 1 - y_0 \) and \( u \) is such that \( f^n_u(y) \leq x_0 \) for all \( n \leq u \), then

\[
|f^n_u(x) - f^n_u(y)| \leq |x - y|
\]

for all \( n \leq u \).

**Proof of lemma 1.** Let \( t \) be the least integer for which \( f^n_0(x) < 1 - x_0 \) and let \( s < t \) be the maximal integer for which \( f^n_0(y) > 1 - y_0 \). Obviously \( |f^n_0(x) - f^n_0(y)| \leq |x - y| \) for \( n \leq s \), since both \( f_0, f_1 \) are contractions on \([1 - x_0, x_0] \). Moreover, \( f^n_0(x), f^n_0(y) \) again satisfy assumptions of the lemma, therefore we may assume \( s = 0 \). Next, define \( r \) to be the moment of the first visit of \( f^n_0(y) \) in \((1 - y_0, 1) \). If we will show the claim for \( n \leq r \) then the points \( f^n_u(x), f^n_u(y) \) again satisfy the assumptions of the lemma, therefore we may assume \( r = u \).

For this purpose observe that \( f^n_0(y) = a^n_0 y \) and \( f^n_0(x) = a^n_0 x \) for \( n \leq r - 1 \), i.e. application of \( f_0 \) and \( f_1 \) is actually a multiplication by \( a_0, a_1 \), respectively. Indeed, assume contrary to our claim that \( f^n_{u-1}(y) > 1 - x_0 \) and \( \omega_n = 1 \). Then \( f^n_{u-1}(y) = f_1(f^n_{u-1}(y)) > f_1(1 - x_0) = 1 - y_0 \), hence \( r = n \) which is a contradiction. Since \( f^n_0(y) = a^n_0 y \) and \( f^n_0(x) = a^n_0 x \) for \( n \leq r - 1 \), we have for these \( n \)’s

\[
|f^n_0(x) - f^n_0(y)| = a^n_0 |x - y|.
\]

But since \( f^n_0(y) \leq 1 - y_0 < y \) for \( n \leq r - 1 \), we have \( a^n_0 < 1 \) which completes the proof in the case \( n \leq r - 1 \).

The only point remaining now is to show that \( |f^n_0(x) - f^n_0(y)| \leq |x - y| \). If \( f^n_{u-1}(x) \geq 1 - x_0 \), then the statement is obviously true, since both \( f_0, f_1 \) are contractions on \([1 - x_0, x_0] \) and the statement is true for \( n = r - 1 \). We are reduced now to proving \( |f^n_0(x) - f^n_0(y)| \leq |x - y| \) provided that \( f^n_{u-1}(x) < 1 - x_0 < f^n_{u-1}(y) \). Let us consider the function \( k \mapsto |f_1(ky) - f_1(kx)| \) for \( k \in \left[ \frac{1 - x_0}{y}, \frac{1 - x_0}{x} \right] \) (this condition is equivalent to say that \( 1 - x_0 \in [kx, ky] \), thus the condition \( a^n_{u-1} \in \left[ \frac{1 - x_0}{y}, \frac{1 - x_0}{x} \right] \) is equivalent to our case now). We assert that this function is nonincreasing. Indeed,

\[
f_1(ky) - f_1(kx) = (f_1(ky) - f_1(1 - x_0)) + (f_1(1 - x_0) - f_1(kx))
\]

\[
= a_0 (ky - (1 - x_0)) + a_1 ((1 - x_0) - kx),
\]

hence the function is linear with the slope equal to \( a_0 y - a_1 x \) which is negative since \( |x - y| < a_1 \eta_1 \) and \( (5) \) holds for \( \eta_1 \).

We compute now \( |f_1(ky) - f_1(kx)| \) for \( k = k_0 := \frac{1 - x_0}{y} \). We have \( |f_1(k_0y) - f_1(k_0x)| = a_1(k_0y - k_0x) = a_1 k_0 (y - x) \) and \( a_1 k_0 y = f_1(k_0 y) = f_1(1 - x_0) = 1 - y_0 \leq y \) which implies that \( a_1 k_0 \leq 1 \). Combining that with the monotonicity of the considered function yields

\[
|f^n_0(x) - f^n_0(y)| = |f_1(a^n_{u-1} y) - f_1(a^n_{u-1} x)| \leq |f_1(k_0 y) - f_1(k_0 x)| \leq |x - y|
\]

which completes the proof of lemma 1. \( \Box \)
Lemma 2. If \( x, y \in [1 - x_0, x_0] \), \( |x - y| < \eta_1 \), and \( u \) is such that \( f_n(u) \leq x_0 \) for all \( n \leq u \), then \( |f_n(x) - f_n(y)| \leq a_1|x - y| \).

Proof of lemma 2. The proof is essentially the same as in the case of previous lemma. We define \( t \) and \( r \) in the same way and assume without loss of generality that \( t = 1, r = u \) (for \( n \geq r \) we can apply lemma 1). We again observe that \( f_n(y) = \alpha_n^x y, f_n(x) = \alpha_n^x x \), and \( f_n(y) \leq 1 - y_0 \) for \( n \leq r - 1 \). The difference is that \( y > 1 - y_0 \) is not true anymore. However, by the definition of \( a_1 \) we have \( 1 - y_0 = a_1(1 - x_0) \leq a_1 y \), thus \( \alpha_n^y \leq a_1 \) which proves the assertion for \( n \leq r - 1 \) (cf (8)).

If \( n = r \) then we have again two cases. If \( f_n^{-1}(x) > 1 - x_0 \), then the statement is obviously true, since both \( f_0, f_1 \) are contractions on \([1 - x_0, x_0]\) and the statement is true for \( n = r - 1 \). If \( f_n^{-1}(x) < 1 - x_0 \) then \( \alpha_n^x x = f_n(x) < 1 - y_0 = a_1(1 - x_0) \leq a_1 x \), so \( \alpha_n^y - \alpha_n^x y \leq a_1(y - x) \). Observation that \( f_n(y) < \alpha_n^y y \) yields the assertion. \( \square \)

Proof of proposition 1. We can define the following infinite sequences of natural numbers

\[
t_1 := \min \{ n \geq 1 : f_n(x) < 1 - x_0 \text{ or } f_n(y) > x_0 \},
\]

\[
t_{k+1} := \begin{cases} 
\min \{ n \geq t_k : f_n(y) > x_0 \} & \text{if } f_n(x) < 1 - x_0, \\
\min \{ n \geq t_k : f_n(x) < 1 - x_0 \} & \text{if } f_n(y) > x_0,
\end{cases} \quad k \geq 1,
\]

\[
u_k := \begin{cases} 
\max \{ n \leq t_{k+1} : f_n(x) < 1 - y_0 \} & \text{if } f_n(x) < 1 - x_0, \\
\max \{ n \leq t_{k+1} : f_n(y) > y_0 \} & \text{if } f_n(y) > x_0,
\end{cases} \quad k \geq 1.
\]

To finish the proof notice that the statement for \( n \leq u_1 \) follows from lemma 2 (or its symmetric version) with \( u = u_1 \). Hence, from the definition of \((u_k)\), the points \( f_n(u_k), f_n(y) \) satisfy assumptions of lemma 1 (or its symmetric version) with \( u = u_2 - u_1 \). We continue in this fashion: for every \( k \) the points \( f_n(u_k), f_n(y) \) satisfy assumptions of lemma 1 or its symmetric version with \( u = u_{k+1} - u_k \), and the conclusion follows.

To obtain the statement for any \( \omega \) observe that for some \( k \) we cannot define \( t_{k+1} \) and in this case either lemma 1 or 2 applies for \( f_n(x), f_n(y) \) with arbitrary large \( u \). \( \square \)

Proposition 2. There exists \( \eta_2 > 0 \) such that if \( x, y \in [1 - x_0, x_0] \) and \( |x - y| < \eta_2 \), then

\[
\mathbb{E}_x \{ |f_n(x) - f_n(y)| \} \leq Lq^n |x - y|
\]

for all natural \( n, L \geq 1, q < 1 \).

From now on, \( M, \alpha, \varepsilon \), and \( p \) always stand for the quantities chosen in the proof of existence of a stationary measure. Fix \( \varepsilon \in (0, 1) \) and define

\[
A_{x, \varepsilon} := \{ \omega \in \Sigma : f_n(x) < \varepsilon \}, \quad A_{y} := \{ \omega \in \Sigma : f_n(x) > 1 - \varepsilon \},
\]

\[
B_{x,n} := \bigcap_{j=1}^n A_{x,j}, \quad B_{x} := \bigcap_{j=1}^n A_{x,j}.
\]

Lemma 3. If \( x < \varepsilon \) then

\[
\mathbb{P}_x(B_{x,n}) \leq (\varepsilon / x)^n p^n
\]
for all \( n \geq 0 \). The same estimation holds for \( \mathbb{P}_{1-x}(B^{1-x,n}) \).

**Proof.** Fix \( x \in \varepsilon \) and recall that we write \( a^\omega = a_{\omega_n} \cdots a_{\omega_1} \) for \( n \geq 1 \) and \( \omega \in \Sigma \). We first observe that \( \mathbb{E}_x 1_{B_{x,\rho-1}}(a^\omega)^{-\alpha} \leq p^\alpha \). Indeed, by (3) we have

\[
\mathbb{E}_x(1_{B_{x,\rho-1}}(a^\omega)^{-\alpha}) = p_0(f_{\omega}^{n-1}(x))a_0^{-\alpha} + p_1(f_{\omega}^{n-1}(x))a_1^{-\alpha} < p
\]

provided that \( \omega \in B_{x,\rho-1} \). Here \( (F_n)_{n \geq 1} \) stands for the natural filtration in \( (\Sigma, \mathcal{F}) \). Therefore

\[
\mathbb{E}_x 1_{B_{x,\rho-1}}(a^\omega)^{-\alpha} = \mathbb{E}_x \left(1_{B_{x,\rho-1}}(a^\omega)^{-\alpha} \mathbb{E}_x \left(a^\omega^{-\alpha} | F_{n-1}\right)\right) < p\mathbb{E}_x 1_{B_{x,\rho-1}}(a^{n-1})^{-\alpha}.
\]

Proceeding by induction yields \( \mathbb{E}_x 1_{B_{x,\rho-1}}(a^\omega)^{-\alpha} \leq p^\alpha \).

Observe that for all \( \omega \in \Sigma \) with \( \omega \in B_{x,\rho} \) we have \( f_i(x) = a_i^\omega x \) and, in consequence,

\[
B_{x,n} = \{ \omega \in \Sigma : a^\omega_j x < \varepsilon \quad \text{for all} \ j \leq n \}.
\]

The Chebyshev inequality gives now

\[
\mathbb{P}_x(B_{x,n}) = \mathbb{P}_x(\{ \omega \in \Sigma : a^\omega_j x < \varepsilon \quad \text{for all} \ j \leq n \})
\]

\[
\leq \mathbb{P}_x(\{ \omega \in \Sigma : (a^\omega)^{-1} > x/\varepsilon \} \cap B_{x,\rho-1})
\]

\[
\leq (\varepsilon/x)\mu \mathbb{E}_x 1_{B_{x,\rho-1}}(a^\omega)^{-\alpha} \leq (\varepsilon/x)^\alpha p^\alpha
\]

which establishes our claim for \( \mathbb{P}_x(B_{x,n}) \). The same proof works for \( \mathbb{P}_{1-x}(B^{1-x,n}) \). \( \Box \)

**Lemma 4.** There exists a point \( c \in (1 - x_0, x_0) \) such that for every \( h > 0, \rho > 0 \) there exist a natural number \( n_1 \) and \( \delta > 0 \) such that

\[
\inf_{x \in [h, 1 - h]} \mathbb{P}_x(f_{\omega}^{n_1}(x) \in (c - \rho, c + \rho)) > 0
\]

for \( x \in [h, 1 - h] \).

**Proof.** First of all, by (6) and symmetry we have \( f_1([1 - x_0, y_0 + a_1\eta_1]) \subseteq [1 - y_0, x_0] \) and \( f_0([1 - y_0, x_0]) \subseteq [1 - x_0, y_0] \). Hence the composition \( f_0 \circ f_1 \) restricted to the interval \([1 - x_0, y_0 + a_1\eta_1]\) is a contraction and acts to the interval \([1 - x_0, y_0 + a_1\eta_1]\). Let \( c \) be the unique attractive fixed point for this composition on \([1 - x_0, y_0 + a_1\eta_1]\). For any point \( x \in [1 - x_0, y_0 + a_1\eta_1] \) and \( \rho > 0 \) there exists \( m' \) such that \( \mathbb{P}_x(f_{\omega}^{2m'}(x) \in (c - \rho, c + \rho)) > 0 \).

Choose \( h > 0 \). Now it is sufficient to show that for any \( x \in [h, 1 - h] \) there exists a number \( m'' \) such that \( \mathbb{P}_x(f_{\omega}^{2m''}(x) \in [1 - x_0, y_0 + a_1\eta_1]) > 0 \). Then \( n_1 = 2m + 2m' \) will be desired number, where \( m \) is the maximum of \( m'' \) for \( x \in [h, 1 - h] \). Indeed, the quantity

\[
\inf_{x \in [h, 1 - h]} \min_{(i_1, \ldots, i_m) \in \{0, 1\}^m} P_{x_1} (f_{i_{m-1}} \circ \cdots \circ f_{i_2} \circ f_{i_1}(x))
\]

is positive by (A3) and for any \( x \in [h, 1 - h] \) we can first take a sequence of length \( 2m'' \) with \( \mathbb{P}_x(f_{\omega}^{2m''}(x) \in [1 - x_0, y_0 + a_1\eta_1]) > 0 \) (which may be less than \( 2m \)) and then apply the composition \( f_0 \circ f_1 \) exactly \( m'' + (m - m'') \) many times.

We are left to show that for any \( x \in [h, 1 - h] \) there exists \( m'' \) such that \( \mathbb{P}_x(f_{\omega}^{2m''}(x) \in [1 - x_0, y_0 + a_1\eta_1]) > 0 \). It is readily seen that there exist \( m'' \) and a sequence \((i_1, \ldots, i_{m''}) \in \{0, 1\}^{m''} \) such that for all \( x \in [h, 1 - h] \), the composition \( f_{i_1} \circ \cdots \circ f_{i_{m''}} \) is a contraction and acts to the interval \([1 - x_0, y_0 + a_1\eta_1]\).
\{0,1\}^m\) such that } z_0 := f_{m'} \circ \ldots \circ f_1(x) \in [1 - x_0, y_0] + a_1 \eta_1 \). If } m'' \text{ is even then put } m''' = 2m''. \text{ If not then apply } f_0 \text{ to } z_0. \text{ If } f_0(z_0) \geq 1 - x_0 \text{ then } m''' + 1 = 2m'' \text{ is a desired number. If not then } f_0(z_0) < 1 - x_0, \text{ hence } z_1 := f_1 \circ f_0(z_0) \geq 1 - y_0. \text{ Note that } z_1 = a_1 a_0 z_0 > z_0. \text{ We can repeat this procedure and define } z_0, z_1, \ldots, z_n \text{ for some } n, \text{ which eventually become greater than } 1 - y_0 \text{ for some } n, \text{ which means that } f_{z_n}^{-1}(z_0) > 1 - x_0. \text{ A contradiction. Let } n \text{ be the minimal number with } f_0(z_n) \geq 1 - x_0. \text{ Then } 2m''' = m'''' + 2n + 1 \text{ has the desired property.} \]

\textbf{Proof of proposition 2.} Let } c \text{ be the point from lemma 4. Take } \rho > 0 \text{ to be any positive number less than distance from } c \text{ to the boundary points of } [1 - x_0, y_0]. \text{ Take } h = \varepsilon \text{ (recall that } M, \varepsilon, \alpha \text{ were the numbers given in the proof of existence of the stationary measure; see the comment under proposition 2). Take } n_1 \text{ to be the numbers given in lemma 4. By the continuity of } f_0, f_1 \text{ and the compactness of } [h, 1 - h], \text{ there exists } \eta_2 > 0 \text{ such that if } |x - y| < a_1 \eta_2, \text{ then}

\begin{equation}
\inf_{x \in [h, 1 - h]} \mathbb{P}_x(f_m^n(y) \in (c - \rho, c + \rho)) > 0.
\end{equation}

\text{ Let } n_2 \text{ be such that } (a_0)^{\frac{n_2}{2}} < 1 / (2a_1). \text{ Put } m := n_1 + n_2 \text{ and } \xi := f_0^{n_2}(c) \text{ (i.e. } \xi \text{ is such a number that } \mathbb{P}_x(f_m^n(x) \in [\xi, 1 - \xi]) = 1 \text{ for } x \in [\varepsilon, 1 - \varepsilon]). \text{ Eventually put}

\begin{equation}
\delta := \inf_{x \in [h, 1 - h]} \min_{(i_1, \ldots, i_m) \in [0,1]^m} p_m(f_{i_1} \circ \ldots \circ f_1(x)) \cdot \ldots \cdot p_1(x) > 0.
\end{equation}

\text{ Let us define the following optional times on } \Sigma \text{ for } x \in (0, 1) \text{ (recall that } \theta : \Sigma \to \Sigma \text{ is the left shift and } \theta_n = \theta^n)

\begin{align*}
T_1(x) & := \min\{n \geq 0 : \varepsilon \leq f_m^n(x) \leq 1 - \varepsilon\} + m, \\
T_{n+1}(x) & := T_n(x) + T_1(f_{T_{n}(x)}(x)) \circ \theta_{T_n(x)}, \\
S_1(x) & := \min \left\{ n \geq 1 : \forall_{|y - x| < \eta_1} |f_m^n(x) - f_m^n(y)| \leq \frac{1}{2a_1} |x - y|, \right. \\
& \left. \text{ and } f_m^n(x), f_m^n(y) \in [1 - x_0, x_0] \right\}, \\
S_{n+1}(x) & := S_n(x) + S_1(f_{T_{n+1}(x)}(x)) \circ \theta_{S_n(x)}, \\
\tau_n(x) & := \max\{k \geq 1 : T_k(x) \leq n\}, \\
\sigma_n(x) & := \max\{k \geq 1 : S_k(x) \leq n\},
\end{align*}

\text{ for } x \in (0, 1). \text{ From what has already been proved we conclude that}

\begin{equation}
\mathbb{P}_x(S_1(x) > T_1(x)) \leq 1 - \delta
\end{equation}

\text{ for all } \xi \leq x \leq 1 - \xi. \text{ By the strong Markov property}
\[ P_x(S_t(x) > T_{n+1}(x)) = E_x P_x \left( S_t(x) > T_{n+1}(x) | \mathcal{F}_n \right) \]
\[ = E_x \left( \mathbb{1}_{\{ S_t(x) > T_{n+1}(x) \}} P_{f_{T_{n+1}(x)}} \right) \]
\[ \times \left( S_t(f_{T_{n+1}(x)}) \circ \theta_{T_{n+1}(x)} > T_t(f_{T_{n+1}(x)}) \right) \]
\[ \leq (1 - \delta) P_x(S_t(x) > T_t(x)) \]

for all \( \xi \leq x \leq 1 - \xi \). By induction argument we get
\[ P_x(S_t(x) > T_t(x)) \leq (1 - \delta)^n \]

for such \( x \).

By lemma 3 there exists \( C_1 > 0 \) and \( \gamma \in (0, 1) \) such that \( E_x e^{\gamma T_t(x)} \leq C_1 \) for all \( x \in [\xi, 1 - \xi] \). Induction argument applied below yields
\[ E_x e^{T_t(x)} = E_x \left( e^{T_{n+1}(x)} \mathbb{E}_x \left( e^{T_{n+1}(x)} | \mathcal{F}_{n} \right) \right) \leq C_1 E_x e^{T_{n+1}(x)} \leq C_1^n \]

for \( x \in [\xi, 1 - \xi] \) and \( n \geq 1 \), since \( f_{T_{n+1}}(x) \in [\xi, 1 - \xi] \) for every \( n \). Fix \( \kappa \in (0, 1) \). We have again by the Chebyshev inequality
\[ P_x(S_t(x) < \kappa n) \leq P_x(T_{[\kappa n]}(x) > n) \leq C_1^{[\kappa n]+1} e^{-\gamma n} \]

for all \( x \in [\xi, 1 - \xi] \), thus
\[ P_x(S_t(x) > n) \leq P_x(S_t(x) > [\kappa n] \cap \{ \tau_n(x) \geq \kappa n \}) + P_x(\{ \tau_n(x) < \kappa n \}) \]
\[ \leq P_x(S_t(x) > T_{[\kappa n]}(x)) + C_1^{[\kappa n]+1} e^{-\gamma n} \leq (1 - \delta)^{[\kappa n]} + C_1(C_1 e^{-\gamma})^n. \]

Choose \( \kappa \) such that \( C_1 e^{-\gamma} < 1 \). By the above we have
\[ E_x e^{S_t(x)} \leq e^\gamma \sum_{n=0}^{\infty} e^{\rho n} P_x(S_t(x) > n) \leq C_2 < \infty, \]

for all \( x \in [\xi, 1 - \xi] \) provided that \( \rho \in (0, 1) \) was chosen sufficiently small. Again, conditioning argument yields
\[ E_x e^{\gamma S_t(x)} \leq C_2^n. \]

Eventually, using again the Chebyshev inequality we obtain for such \( x, y \) and any \( \lambda \in (0, 1) \),
\[ E_x |f_{T_{n+1}}(x) - f_{T_{n+1}}(y)| = E_x 1_{\{ \sigma_{T_t(x) < \lambda n} \}} |f_{T_{n+1}}(x) - f_{T_{n+1}}(y)| + E_x 1_{\{ \sigma_{T_t(x) \geq \lambda n} \}} |f_{T_{n+1}}(x) - f_{T_{n+1}}(y)| \]
\[ \leq a_1 |x - y| P_x(S_t(x) > n) + \frac{1}{2 \lambda} |x - y| \]
\[ \leq a_1 C_2^{[\lambda n]} e^{-\rho n} |x - y| + \frac{1}{2 \lambda} |x - y| \]
\[ \leq \left( a_1 C_2^2 e^{-\rho} + \frac{1}{2 \lambda} \right) |x - y|. \]
Take \( \lambda \) such that \( C_2 \lambda e^{-\rho} < 1 \) and put \( L = a_1 + 1, q = \max\{C_2 e^{-\rho}, \frac{1}{2\pi}\} < 1 \). Then by the above we have
\[
\mathbb{E}_x|f_n^\beta(x) - f_n^\beta(y)| \leq L q^n |x - y|
\]
for all natural \( n \) which is the desired conclusion.

**Proof of uniqueness.** Throughout the proof \( p_{1,...,\omega}(x) \) stands for
\[
p_n(f_{\omega_{n-1}} \circ \cdots \circ f_1(x)) \cdots \cdot p_1(x).
\]
First observe that for any \( x \in (0, 1) \) there exists a finite sequence \((i_1,...,i_l) \in \{0,1\}^l \) for some \( l \) such that \( f_{i_1} \circ \cdots \circ f_1(x) \in [1-x_0,y_0] \) which implies that the topological support of any \( P \)-invariant measure \( \mu \) must have nonempty intersection with \([1-x_0,y_0]\). Further, \( f_1 \circ f_0((1-x_0,y_0)) \subseteq (1-x_0,y_0) \) by (7) and \( f_1 \circ f_0 \) is a contraction on the interval \([1-x_0,y_0]\), hence this composition has exactly one attractive fixed point \( c \in (1-x_0,y_0) \). Combining these facts yields \( c \in \Gamma_{\mu} \) for all \( P \)-invariant measures \( \mu \), where \( \Gamma_{\mu} \) denotes the topological support of this measure. The proof is completed by showing that the family \( (U^n \varphi) \) is equicontinuous at \( c \) for any Lipschitz \( \varphi \). Indeed, if there exist at least two different ergodic invariant measures \( \mu_1, \mu_2 \), then there exists a Lipschitz function \( \varphi \) such that \( \| \varphi \|_\mu \leq 1 \) for some \( \delta > 0 \). We consider the averages \( \frac{1}{n}(\varphi(x) + U \varphi(x) + \cdots + U^{n-1} \varphi(c)) \) which must differ from \( \frac{1}{n}(\varphi(c) + U \varphi(c) + \cdots + U^{n-1} \varphi(c)) \) at most \( \delta/2 \), provided that \( x \) is sufficiently close to \( c \). On the other hand, \( c \in \Gamma_{\mu_1} \cap \Gamma_{\mu_2} \), therefore in any neighbourhood of \( c \) we can find points \( x_1, x_2 \) such that considered averages tend to \( \varphi \| \mu_1, \varphi \| \mu_2 \), respectively, which is a contradiction.

We are going to show that \( (U^n \varphi) \) is equicontinuous at any point of \((1-x_0,y_0)\). Take \( x \in (x_0, 1-x_0) \) and \( \delta > 0 \). Take \( n_0 \) such that \( \sum_{n=n_0}^{\infty} 2^\beta (L q^n) < \frac{\delta}{6 \| \varphi \|_\infty} \) (by (A2)) and \( L q^n \leq \frac{1}{6 \| \varphi \|_\infty} \) for \( n \geq n_0 \), where \( \operatorname{Lip}(\varphi) \) denotes the Lipschitz constant of \( \varphi \). By theorem 8 on the page 45 in [31] there exists a concave function \( \beta^* \) with \( \beta(t) \leq \beta^*(t) \leq 2 \beta(t) \). Thus we have \( \sum_{n=n_0}^{\infty} \beta^*(L q^n) < \frac{1}{6 \| \varphi \|_\infty} \).

Take \( y \) such that \( |x - y| < \eta_2 \) and
\[
\sum_{|i_1,...,i_l|} |p_{1,...,\omega}(x) - p_{1,...,\omega}(y)| < \frac{\delta}{2 \| \varphi \|_\infty}, \tag{9}
\]
where the summation is over all finite sequences \((i_1,...,i_n) \in \{0,1\}^n \). It is satisfied provided that \( |x - y| \) is less than, say, \( d > 0 \). Then for \( n \geq n_0 \) we have
\[
|U^n \varphi(x) - U^n \varphi(y)| \leq \sum_{|i_1,...,i_l|} p_{1,...,\omega}(x) \left| \varphi(f_{i_1} \circ \cdots \circ f_1(x)) - \varphi(f_{i_1} \circ \cdots \circ f_1(y)) \right| + |p_{1,...,\omega}(x) - p_{1,...,\omega}(y)| \| \varphi \|_\infty,
\]
where the summation is over all finite sequences \((i_1,...,i_n) \in \{0,1\}^n \). The first term is bounded by \( \operatorname{Lip}(\varphi) E_x|f_n^\beta(x) - f_n^\beta(y)| \). To estimate the second, we have
\[
\sum_{i_1,...,i_n} |p_{i_1,...,i_n}(x) - p_{i_1,...,i_n}(y)|
= \sum_{i_1,...,i_n} |p_n(f_{i_{n-1}} \circ \cdots \circ f_1(x)) - p_n(f_{i_{n-1}} \circ \cdots \circ f_1(y))| \cdot p_{1,...,i_{n-1}}(x)
\]

lemma 3 and the first part of the proof we have
\[ \sum_{i_1 \cdots i_n} p_{i_1 \cdots i_n} \left( f_{i_n} \circ \cdots \circ f_{i_1} \right) |p_{i_1 \cdots i_n-1}(x) - p_{i_1 \cdots i_n}(y)| \]

\[ \leq 2E_x \beta^x\left( f^x_0(x) - f^y_0(y) \right) + \sum_{i_1 \cdots i_n-1} |p_{i_1 \cdots i_n-1}(x) - p_{i_1 \cdots i_n}(y)| . \]

The modulus of continuity \( \beta^x \) is concave, therefore by the Jensen inequality we have
\[ \sum_{i_1 \cdots i_n} |p_{i_1 \cdots i_n}(x) - p_{i_1 \cdots i_n}(y)| \leq 2\beta^x(Lq^x) + \sum_{i_1 \cdots i_n-1} |p_{i_1 \cdots i_n-1}(x) - p_{i_1 \cdots i_n-1}(y)| . \]

Continuing this procedure while \( n > n_0 \) and using (9) yields
\[ \sum_{i_1 \cdots i_n} |p_{i_1 \cdots i_n}(x) - p_{i_1 \cdots i_n}(y)| \leq \sum_{i=0}^{n} 2\beta^x(Lq^x) + \sum_{i_1 \cdots i_n-1} |p_{i_1 \cdots i_n-1}(x) - p_{i_1 \cdots i_n-1}(y)| < \frac{\delta}{3\|\varphi\|_{\infty}} + \frac{\delta}{3\|\varphi\|_{\infty}} . \]

Again by the definition of \( n_0 \) we have
\[ |U^n\varphi(x) - U^n\varphi(y)| < \text{Lip}(\varphi)E_x|f^x_0(x) - f^y_0(y)| \]
\[ + \|\varphi\|_{\infty} \frac{\delta}{3\|\varphi\|_{\infty}} + \|\varphi\|_{\infty} \frac{\delta}{3\|\varphi\|_{\infty}} < \delta \]
for all \( n \) and \( y \) with \( |x - y| < d \). Therefore \( U^n\varphi \) is equicontinuous at any \( x \in \left[ 1 - x_0, x_0 \right] \) which proves the uniqueness of the invariant measure \( \mu_x \).

\[ \square \]

4. The proof of theorem 2

Lemma 5. There exists \( 0 < h < 1/2 \) such that for all \( 0 < \xi < 1/2 \) there exists \( n_0 \) such that \( P^n\delta_x([h, 1-h]) = P^n(f^x_0(x) \in [h, 1-h]) \geq 1/2 \) for all \( x \in [\xi, 1 - \xi] \) and \( n \geq n_0 \).

Proof. Recall that \( M, \alpha, \varepsilon \) are the quantities given in the proof of existence of a stationary measure. The class \( \mathcal{P}_{M,\alpha} \) is \( P \)-invariant. Take \( h > 0 \) such that \( Mh^\alpha < 1/8 \). The definition of \( M \) yields the relation \( Mx^\alpha \geq 1 \) which implies clearly \( P^n\delta_x \in \mathcal{P}_{M,\alpha} \) and thus \( P^n\delta_x([h, 1-h]) \geq 3/4 \) for every \( x \in [\varepsilon, 1-\varepsilon] \) and \( n \geq n_0 \).

Let \( n_0 \) be such that \( (\varepsilon/\xi)^\alpha p^\alpha < 1/8 \) for \( n \geq n_0 \). Take \( x \notin [\varepsilon, 1-\varepsilon] \) and \( x \in [\xi, 1-\xi] \). Denote by \( T \) the time of the first visit of \( x \) in \([\varepsilon, 1-\varepsilon] \). Then by the strong Markov property, lemma 3 and the first part of the proof we have
\[ \mathbb{P}_x \left( f^x_0(x) \notin [h, 1-h] \right) \]
\[ \leq \sum_{k=1}^{n} \mathbb{P}_x \left( f^x_0(x) \notin [h, 1-h] \middle| T = k \right) \mathbb{P}_x(T = k) + \mathbb{P}_x(T > n) < 1/2 , \]
for \( n \geq n_0 \), since \( \mathbb{P}_x(T > n) \leq (\varepsilon/\xi)^\alpha p^\alpha \leq (\varepsilon/\xi)^\alpha p^\alpha < 1/8 \) for \( x \in [\xi, 1-\xi] \). \[ \square \]

From now on, \( h \) denotes the quantity given in lemma 5.
Lemma 6. For every $\rho > 0$ there exists $\delta > 0$ such that for every $\xi > 0$ there exists a natural number $m$ such that for all $n \geq m$ we have

$$\inf_{x \in [\xi, 1 - \xi]} \mathbb{P}_x (f^n_\omega (x) \in (c - \rho, c + \rho)) \geq \delta.$$ 

Proof. Fix $\xi$. Let $n_0$ be the number given in lemma 5. Let $n_1, \zeta$ be the numbers given in lemma 4. Let $m := n_1 + n_0$ and $\delta := \zeta / 2$. Take $n \geq m$. Then $n - n_1 \geq n_0$, thus

$$\mathbb{P}_x (f^n_\omega (x) \in (c - \rho, c + \rho)) = \mathbb{P}_x \left( f^n_\omega (x) \in (c - \rho, c + \rho) \middle| f^{m-n_1}_\omega (x) \in [h, 1 - h] \right) \geq \frac{1}{2} \cdot \zeta > 0$$

for every $x \in [\xi, 1 - \xi]$ by lemma 4.

We are in position to show theorem 2. The idea is to apply the lower bound technique (cf theorem 4.1 in [30]).

Proof of theorem 2. Take $x < y$, $\lambda > 0$ and a Lipschitz function $\varphi$. By the equicontinuity of $(U^n \varphi)$ at $c$, there exists $\rho > 0$ such that

$$|U^n \varphi(u) - U^n \varphi(v)| < \lambda \quad \text{for } n \geq 1 \text{ and } u, v \in (c - \rho, c + \rho). \quad (10)$$

Define

$$A_n := \{ (\omega, \omega') \in \Sigma \times \Sigma : c - \rho < f^n_\omega (x) < f^n_\omega (y) < c + \rho \}.$$ 

By lemma 4 there exist $m_1, \delta > 0$ with $\mathbb{P}_x \otimes \mathbb{P}_y (A_{m_1}) \geq \delta^2$. Put $\xi_1 := \min \{ f_{m_1}^n (x), 1 - f_{m_1}^n (y) \}$. Then $\xi_1 \leq f_{m_1}^n (x) < f_{m_1}^n (y) \leq 1 - \xi_1$ for all $(\omega, \omega') \in \Sigma \times \Sigma$. Once again, by lemma 6 there exists $m_2$ such that $\mathbb{P}_x \otimes \mathbb{P}_y (A_{m_2} | F_{m_1}) \geq \delta^2$. Put $\xi_2 := \min \{ f_{m_1+m_2}^n (x), 1 - f_{m_1+m_2}^n (y) \}$. Obviously, $\xi_3 \leq f_{m_1+m_2}^n (x) < f_{m_1+m_2}^n (y) \leq 1 - \xi_2$ for all $(\omega, \omega') \in \Sigma \times \Sigma$.

We continue in this fashion to construct a sequence $m_1, m_2, \ldots$ such that $\mathbb{P}_x \otimes \mathbb{P}_y (A_{m_n} | F_{m_{n-1}}) \geq \delta^2$

for all $n$’s. It is easy to check that $\mathbb{P}_x \otimes \mathbb{P}_y (B_n) \leq (1 - \delta^2)^n \quad (11)$

where

$$B_n := \bigcap_{k=1}^n \Sigma \times \Sigma \setminus A_{m_k}.$$ 

Hence for $n \geq m_k$ we get

$$U^n \varphi (x) - U^n \varphi (y) = \int_{\Sigma \times \Sigma} \left( \varphi (f^n_\omega (x)) - \varphi (f^n_\omega (y)) \right) \mathbb{P}_x (d\omega) \otimes \mathbb{P}_y (d\omega')$$

$$= \sum_{j=1}^k \int_{A_{m_j}} \mathbb{E}_{\epsilon_j} \left( \varphi (f^n_\omega (x)) - \varphi (f^n_\omega (y)) \right) \mathbb{P}_x (d\omega) \otimes \mathbb{P}_y (d\omega')$$

$$+ \int_{B_k} \left( \varphi (f^n_\omega (x)) - \varphi (f^n_\omega (y)) \right) \mathbb{P}_x (d\omega) \otimes \mathbb{P}_y (d\omega')$$

6238
\[
= \sum_{j=1}^k \int \int_{A_{m_j}} \left( U^{m-m_j} \phi(f_j^{m_j}(x)) - U^{m-m_j} \phi(f_j^{m_j}(y)) \right) \mathbb{P}_x(\mathrm{d}\omega) \otimes \mathbb{P}_y(\mathrm{d}\omega') \\
+ \int \int_{B_k} (\phi(f_k^n(x)) - \phi(f_k^n(y))) \mathbb{P}_x(\mathrm{d}\omega) \otimes \mathbb{P}_y(\mathrm{d}\omega').
\]

By (10), (11) and the definition of \(A_{m_j}\)'s eventually we have
\[
|U^n \phi(x) - U^n \phi(y)| \leq \lambda + (1 - \delta^2)^k < 2\lambda
\]
provided that \(k\) was sufficiently large. Therefore
\[
\lim_{n \to \infty} |U^n \phi(x) - U^n \phi(y)| \to 0
\]
for every \(x, y \in (0, 1)\). If \(\mu_*\) is the stationary probability measure and \(\nu \in \mathcal{M}_1\), then for any Lipschitz function \(\phi\) we obtain
\[
\left| \int_{(0,1)} \phi(x) P^n \nu(\mathrm{d}x) - \int_{(0,1)} \phi(y) \mu_*(\mathrm{d}y) \right| \\
\leq \int_{(0,1) \times (0,1)} |U^n \phi(x) - U^n \phi(y)| \nu(\mathrm{d}x) \otimes \mu_*(\mathrm{d}y) \to 0
\]
by the Lebesgue convergence theorem. Thus \(P^n \nu \to \mu_*\) weakly-* for every \(\nu \in \mathcal{M}_1\) which is our assertion. \(\square\)

5. The proof of theorem 3

Let \(c\) be the number given in lemma 4. Recall that \(c\) is the unique attractive fixed point of the composition \(f_0 \circ f_1\) on \((1 - x_0, y_0)\). For any \(\rho > 0\) we will write \(S_\rho(x)\) for the time of the first visit of the process \((f_n \omega(x))\) in \((c - \rho, c + \rho)\).

**Lemma 7.** If \(\rho > 0, x \in (0, 1)\), then \(S_\rho(x)\) is finite \(P_x\text{-a.s.}\)

We omit the proof, since it is an easy consequence of lemmas 3 and 4.

**Lemma 8.** Let \(q < 1, L \geq 1, \eta_2 > 0\) be the quantities given in proposition 1. Let \(x, y \in [1 - x_0, x_0]\) be such that \(|x - y| < \eta_2\). If \(q < r < 1\) then for every \(\lambda > 0\) there exist a natural \(n_\lambda\) and a measurable set \(\Sigma \subseteq \Sigma\) such that \(\mathbb{P}_x(\Sigma) > 1 - \lambda\) and
\[
|f^n_\omega(x) - f^n_\omega(y)| < r^n
\]
for every \(\omega \in \Sigma\) and \(n \geq n_\lambda\).

**Proof.** This is an immediate consequence of the Chebyshev inequality and the Borel–Cantelli lemma. Indeed,
\[
\mathbb{P}_x(\{|f^n_\omega(x) - f^n_\omega(y)| \geq r^n\}) \leq E_x|f^n_\omega(x) - f^n_\omega(y)|^{-n} \leq L(q/r)^n,
\]
therefore \( \{|f^n_n(x) - f^n_n(y)| \geq r^k\} \) occurs only finitely many times \( \mathbb{P}_{\tau} \)-a.s. which completes the proof. \( \square \)

The following lemma is proven in [20], lemma 3. For the convenience of the reader we rewrite the proof here.

**Lemma 9.** There exists \( \rho > 0 \) such that for every \( x, y \in (c - \rho, c + \rho) \) the measures \( \mathbb{P}_x, \mathbb{P}_y \) on \( \Sigma \) are equivalent, i.e. the sets of \( \mathbb{P}_x \) and \( \mathbb{P}_y \) measure zero are the same.

**Proof.** Put \( \rho := \eta_j/2 \). We have \( \sum_{k=0}^{\infty} \beta(r^k) < \infty \), since \( p_0, p_1 \) are Dini continuous (let us recall that \( \beta \) stands for the modulus of continuity of \( p_0 \) and \( p_1 \)). Take \( \delta \) such that \( \delta < p_0(z) < 1 - \delta \) for every \( z \in (0, 1) \), by assumption (A3).

Fix \( x, y \in (c - \rho, c + \rho) \) and a measurable set \( E \) with \( \mathbb{P}_{\tau}(E) = 0 \). Take \( \lambda > 0 \) and \( q < r < 1 \), where \( q \) is given in lemma 2. We will show that \( \mathbb{P}_y(E) < 2\lambda \). Let \( \Sigma \) and \( n_1 \) be given in lemma 8. Let \( m \geq n_1 \) be such that \( \sum_{k=m+1}^{\infty} \beta(r^k) < \lambda/2 \).

Put \( \Sigma := \bigcup_{k=1}^{\infty} \{0, 1\}^k \) and let \( \Xi \subseteq \Sigma \) be a countable set such that \( E \subseteq \bigcup_{k=1}^{\infty} C_k \Xi \) and \( \mathbb{P}_\tau(\bigcup_{k=2}^{\infty} C_k) < (\lambda/2) \delta/(1 - \delta)^m \), where \( C_k \) denotes the cylinder set in \( \Sigma \) corresponding to the finite sequence \( i \in \Xi \). Moreover, we assume \( C_1 \) to be pairwise disjoint for \( i \in \Xi \). Let \( Q_n := \{(i_1, i_2, \ldots) \in \Sigma : |f^n_{\omega}(x) - f^n_{\omega}(y)| < r^k \} \) for \( n \geq m \) and \( Q_n := \Sigma \) for \( n < m \). \( Q := \bigcap_{n=1}^{\infty} Q_n \). By lemma 8 we have \( \mathbb{P}_\tau(\Sigma \setminus Q) \leq \mathbb{P}_\tau(\Sigma \setminus \Xi) < \lambda \). Take \( n \geq m \), \( (i_1, i_2, \ldots) \in Q_n \). We obtain the following estimation

\[
p_{i_1, \ldots, i_n}(y) \leq p_{i_1, \ldots, i_n}(x) \left( \frac{1 - \delta}{\delta} \right)^m \prod_{k=m+1}^{n} \left( 1 + \frac{|p_{i_1, \ldots, i_k}(y) - p_{i_1, \ldots, i_k}(x)|}{p_{i_1, \ldots, i_k}(x)} \right) \\
\leq p_{i_1, \ldots, i_n}(x) \left( \frac{1 - \delta}{\delta} \right)^m \prod_{k=m+1}^{n} \left( 1 + \frac{\beta(r^k)}{\delta} \right).
\]

One can show easily by induction the following claim: if \( r_1, r_2, \ldots \) are positive numbers such that \( \sum_{k=1}^{\infty} r_k < 1/2 \), then \((1 + r_1) \cdot \ldots \cdot (1 + r_k) \leq 1 + 2(r_1 + \cdots + r_k)\). Application of this claim yields

\[
\prod_{k=m+1}^{n} \left( 1 + \frac{\beta(r^k)}{\delta} \right) \leq 1 + 2 \sum_{k=m+1}^{\infty} \frac{\beta(r^k)}{\delta} \leq 2,
\]

and thus

\[
p_{i_1, \ldots, i_n}(y) \leq 2 \left( \frac{1 - \delta}{\delta} \right)^m p_{i_1, \ldots, i_n}(x)
\]

for \( n \geq m \). If \( n < m \) then this holds trivially for any \( \omega \in \Sigma \).

Take \( i = (i_1, \ldots, i_n) \in \Xi \). If \( Q \) and \( C_1 \) are not disjoint then also \( Q_i \) and \( C_1 \) are not disjoint, hence we have the above estimation

\[
\mathbb{P}_\tau(Q \cap C_1) \leq \mathbb{P}_\tau(Q_i \cap C_1) = p_{i_1, \ldots, i_n}(y)
\leq 2 \left( \frac{1 - \delta}{\delta} \right)^m p_{i_1, \ldots, i_n}(x) = 2 \left( \frac{1 - \delta}{\delta} \right)^m \mathbb{P}_\tau(C_1).
\]

Moreover,

\[
\mathbb{P}_\tau \left( \Sigma \setminus Q \cap \bigcup_{i \in \Xi} C_i \right) < \lambda.
\]
Recall here that the cylinders $C_i, i \in \Xi$ are disjoint. Combining that with two above inequalities yields
\[
\mathbb{P}_y(E) \leq \mathbb{P}_y \left( \bigcup_{i \in \Xi} C_i \right) = \mathbb{P}_y \left( (\Sigma \setminus Q) \cap \bigcup_{i \in \Xi} C_i \right) + \mathbb{P}_y \left( Q \cap \bigcup_{i \in \Xi} C_i \right)
\]
\[
< \lambda + \sum_{i \in \Xi} \mathbb{P}_y(Q \cap C_i) \leq \lambda + \sum_{i \in \Xi} 2 \left( \frac{1 - \delta}{\delta} \right)^m \mathbb{P}_x(C_i)
\]
\[
= \lambda + 2 \left( \frac{1 - \delta}{\delta} \right)^m \mathbb{P}_x \left( \bigcup_{i \in \Xi} C_i \right) < 2 \lambda
\]
which is the desired assertion.

\[\square\]

**Proof of theorem 3.** Let $\varphi$ be any Lipschitz function. The statement for any continuous function follows from the density of the set of Lipschitz functions in $C((0,1))$ with the supremum norm. Let $\rho$ be given in lemma 9. There exists $y \in (c - \rho, c + \rho)$ such that
\[
\frac{\varphi(y) + \cdots + \varphi(f^{-1}_n(y))}{n} \xrightarrow{n \to \infty} \int \varphi \, d\mu_x
\]
for $\mathbb{P}_x$-a.e. $\omega \in \Sigma$ where $\mu_x$ is the unique $P$-invariant measure. It follows by the fact that $c$ is in the support of $\mu_x$ (see the beginning of the proof of uniqueness) and by the Birkhoff ergodic theorem. Take any $z \in (c - \rho, c + \rho)$. For $\mathbb{P}_x$-a.e. $\omega \in \Sigma$ there exists $n(\omega)$ such that $|f^n_c(z) - f^n_c(y)| \leq r^n$ for $n \geq n(\omega)$, where $q < r < 1$ by lemma 8. Therefore
\[
\left| \frac{\varphi(y) + \cdots + \varphi(f^{-1}_n(y))}{n} - \frac{\varphi(z) + \cdots + \varphi(f^{-1}_n(z))}{n} \right| \xrightarrow{n \to \infty} 0
\]
for $\mathbb{P}_x$-a.e. $\omega \in \Sigma$. By lemma 9 the measures $\mathbb{P}_z$, $\mathbb{P}_y$ on $\Sigma$ are absolutely continuous. Hence
\[
\frac{\varphi(z) + \cdots + \varphi(f^{-1}_n(z))}{n} \xrightarrow{n \to \infty} \int \varphi \, d\mu_x
\]
for $\omega \in D_x$, where $D_x \subseteq \Sigma$ is certain measurable set with $\mathbb{P}_x(D_x) = 1$.

To complete the proof fix any $x \in (0, 1)$ and observe that by lemma 7 one can find a set $\Xi \subseteq \Sigma = \bigcup_{i=1}^\infty \{0, 1\}^i$ such that $f_{i_1} \circ \cdots \circ f_{i_i}(x) \in (c - \rho, c + \rho)$ for $i = (i_1, \ldots, i_i) \in \Xi$. For $i \in \Xi$ we have $\mathbb{P}_x(C_i) = 1$ and
\[
\frac{\varphi(x) + \cdots + \varphi(f^{-1}_n(x))}{n} \xrightarrow{n \to \infty} \int \varphi \, d\mu_x
\]
for every $\omega \in C$. 

\[\square\]
Acknowledgments

We are grateful to Tomasz Szarek for his advice. We also would like to thank the anonymous referees whose suggestions helped to improve the paper.

ORCID iDs

Klaudiusz Czudek https://orcid.org/0000-0003-1977-165X

References

[1] Alsedé L and Misiurewicz M 2014 Random interval homeomorphisms Publ. Mat. 58 15–36
[2] Barański K and Świątek A 2019 Singular stationary measures for random piecewise affine interval homeomorphisms J. Dyn. Diff. Equa. https://doi.org/10.1007/s10884-019-09807-5
[3] Barnsley M, Demko S, Elton J and Geronimo J 1988 Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities Ann. Inst. Henri Poincaré B 24 367–94
Barnsley M, Demko S, Elton J and Geronimo J 1989 Erratum to: “Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities” Ann. Inst. Henri Poincaré B 25 589–590
[4] Barnsley M F and Demko S 1985 Iterated function systems and the global construction of fractals Proc. R. Soc. A 399 243–75
[5] Bhattacharya R N and Lee O 1988 Asymptotics of a class of Markov processes which are not in general irreducible Ann. Probab. 16 1333–47
[6] Breiman L 1960 The strong law of large numbers for a class of Markov chains Ann. Math. Stat. 31 801–3
[7] Czudek K and Szarek T 2020 Ergodicity and central limit theorem for random interval homeomorphisms (arXiv:2003.09490)
[8] Da Prato G and Zabczyk J 1996 Ergodicity for Infinite-Dimensional Systems (London Mathematical Society Lecture Note Series vol 229) (Cambridge: Cambridge University Press)
[9] Deroin B, Kleptsyn V, Navas A and Parwani K 2013 Symmetric random walks on Homeo+(R) Ann. Probab. 41 2066–89
[10] Bertrand D 2005 Hypersurfaces Levi-plates immergées dans les surfaces complexes de courbure positive Ann. Sci. École Norm. Sup. 38 57–75
[11] Bertrand D, Kleptsyn V and Navas A 2007 Sur la dynamique unidimensionnelle en régularité intermédiaire Acta Math. 199 199–262
[12] Diaconis P and Shahshahani M 1986 Products of random matrices and computer image generation Contemp. Math 50 173–82
[13] Diaconis P and Freedman D 1999 Iterated random functions SIAM Rev. 41 45–76
[14] Díaz J and Gelfert K 2013 Porcupine-like horseshoes: topological and ergodic aspects Progress and Challenges in Dynamical Systems (Springer Proceedings in Mathematics & Statistics vol 54) ed S Ibíñez, J Pérez del Río, A Pumariño and J Rodríguez (Heidelberg: Springer) pp 199–219
[15] Díaz L J, Gelfert K, Marcarini T and Rams M 2019 The structure of the space of ergodic measures of transitive partially hyperbolic sets Monatsh. Math. 190 441–79
[16] Díaz L J, Gelfert K and Rams M 2017 Nonhyperbolic step skew-products: ergodic approximation Ann. Inst. Henri Poincaré Anal. Non Linéaire 34 1561–98
[17] Doeblin W and Fortet R 1937 Sur des chaînes à liaisons complètes Bull. Soc. Math. de France 65 132–48
[18] Dolgopyat D, Fayad B and Saprykina M 2020 Erratic behaviour for one-dimensional random walks in generic quasi-periodic environment (arXiv:1901.10709)
[19] Dubins L E and Freedman D A 1966 Invariant probabilities for certain Markov processes Ann. Math. Stat. 37 837–48
[20] Elton J H 1987 An ergodic theorem for iterated maps Ergod. Theor. Dynam. Syst. 7 481–8
[21] Fisher D 2011 Groups acting on manifolds: around the Zimmer program Geometry, Rigidity, and Group Actions (Chicago Lectures in Mathematics) (Chicago, IL: University of Chicago Press) pp 72–157

[22] Furstenberg H 1971 Random walks and discrete subgroups of Lie groups Advances in Probability and Related Topics vol 1 (New York: Dekker) pp 1–63

[23] Gharaei M and Homburg A J 2017 Random interval diffeomorphisms Discrete Contin. Dyn. Syst. S 10 241–72

[24] Gordin M I and Lifšic B A 1978 The central limit theorem for stationary Markov processes Dokl. Akad. Nauk SSSR 239 766–67

[25] Hutchinson J E 1981 Fractals and self-similarity Indiana Univ. Math. J. 30 713–47

[26] Kaloshin V Y and Sinai Ya G 2000 Nonsymmetric simple random walks along orbits of ergodic automorphisms On Dobrushin’s Way. From Probability Theory to Statistical Physics (American Mathematical Society Translation Series 2 vol 198) (Providence, RI: American Mathematical Society) pp 109–15

[27] Kaloshin V Y and Sinai Ya G 2000 Simple random walks along orbits of Anosov diffeomorphisms Tr. Mat. Inst. Steklova 228 236–45 reprinted in 2000 Proc. Steklov Inst. Math. 228 224–33

[28] Karlin S 1953 Some random walks arising in learning models. I. Pac. J. Math. 3 725–56

[29] Kleptsyn V and Volk D 2014 Physical measures for nonlinear random walks on interval Mosc. Math. J. 14 339–65

[30] Lasota A and Yorke J A 1994 Lower bound technique for Markov operators and iterated function systems Random Comput. Dyn. 2 41–77

[31] Lorentz G G 1966 Approximation of Functions (New York: Holt, Rinehart and Winston)

[32] Malicet D 2017 Random walks on Homeo(S^1) Commun. Math. Phys. 356 1083–116

[33] Maxwell M and Woodroofe M 2000 Central limit theorems for additive functionals of Markov chains Ann. Probab. 28 713–24

[34] Navas A 2011 Groups of circle diffeomorphisms (translation of the 2007 Spanish edition) Chicago Lectures in Mathematics Spanish edn (Chicago, IL: University of Chicago Press)

[35] Navas A 2018 Group actions on 1-manifolds: a list of very concrete open questions Proc. of the Int. Congress of Mathematicians—Rio de Janeiro 2018 (Invited Lectures vol 3) (Hackensack, NJ: World Scientific Publishing) pp 2035–62

[36] Onicescu O and Mihoc G 1935 Sur les chaines de variables statistiques Bull. Sci. Math. 59 174–92

[37] Sinai Y G 1999 Simple random walks on tori J. Stat. Phys. 94 695–708

[38] Śleszka M 2011 The rate of convergence for iterated function systems Stud. Math. 205 201–14

[39] Solomyak B 1995 On the random series \( \sum_{n=0}^{\infty} \lambda^n \) (an Erdős problem) Ann. Math. 142 611–25

[40] Stenflo Ö 2001 A note on a theorem of Karlin Stat. Probab. Lett. 54 183–7

[41] Stenflo Ö 2012 A survey of average contractive iterated function systems J. Differ. Equ. Appl. 18 1355–80

[42] Szarek T and Zdunik A Attractors and invariant measures for random interval homeomorphisms