RANDOM ITERATION WITH PLACE DEPENDENT PROBABILITIES

BY

RAFAŁ KAPICA (KRAKÓW) AND MACIEJ ŚLĘCZKA (KATOWICE)

Abstract. Markov chains arising from random iteration of functions $S_\theta : X \to X$, $\theta \in \Theta$, where $X$ is a Polish space and $\Theta$ is an arbitrary set of indices are considered. At $x \in X$, $\theta$ is sampled from a distribution $\vartheta_x$ on $\Theta$ and $\vartheta_x$ are different for different $x$. Exponential convergence to a unique invariant measure is proved. This result is applied to the case of random affine transformations on $\mathbb{R}^d$ giving the existence of exponentially attractive perpetuities with place dependent probabilities.

2010 AMS Mathematics Subject Classification: Primary: 60J05; Secondary: 37A25.

Key words and phrases: random iteration of functions, exponential convergence, invariant measure, perpetuities

1. INTRODUCTION

We consider the Markov chain of the form $X_0 = x_0$, $X_1 = S_{\theta_0}(x_0)$, $X_2 = S_{\theta_1} \circ S_{\theta_0}(x_0)$ and inductively

$$X_{n+1} = S_{\theta_n}(X_n),$$

(1.1)

where $S_{\theta_0}, S_{\theta_1}, \ldots, S_{\theta_n}$ are randomly chosen from a family $\{S_\theta : \theta \in \Theta\}$ of functions that map a separable metric space $X$ into itself. If the chain is at $x \in X$ then $\theta \in \Theta$ is sampled from a distribution $\vartheta_x$ on $\Theta$, where $\vartheta_x$ are different for different $x$. The chain $(X_n)_{n \in \mathbb{N}_0}$ is defined in a rigorous way in Section 3. We are interested
in the rate of convergence to a stationary distribution $\mu_*$ on $X$, i.e.

$$\tag{1.2} P\{X_n \in A\} \to \mu_*(A) \text{ as } n \to \infty.$$ 

When the above convergence is exponential then $(X_n)_{n \in \mathbb{N}_0}$ is called exponentially ergodic. In the case of constant probabilities, i.e. $\vartheta_x = \vartheta_y$ for $x, y \in X$, the basic tool when studying asymptotics of (1.1) are backward iterations $Z_{n+1} = S_{\theta_0} \circ S_{\theta_1} \circ \cdots \circ S_{\theta_n}(x_0)$. Since $X_n$ and $Z_n$ are identically distributed and, under suitable conditions, $Z_n$ converge almost surely at exponential rate to some random element $Z$, one obtains exponential ergodicity of $(X_n)_{n \in \mathbb{N}_0}$ (see [6]). For place dependent $\vartheta_x$ we need different approach because distributions of $X_n$ and $Z_n$ are not equal.

The simplest case when $\Theta = \{1, \ldots, n\}$ is treated in [2] and [29], where the existence of a unique attractive invariant measure is established. Similar result holds true when $\Theta = [0, T]$ and $\vartheta_x$ are absolutely continuous (see [15]). Recently it was shown that the rate of convergence in the case of $\Theta = \{1, \ldots, n\}$ is exponential (see [30]).

In this paper we treat the general case of place dependent $\vartheta_x$ for arbitrary $\Theta$ and prove the existence of a unique exponentially attractive invariant measure for (1.1). Our approach is based on the coupling method which can be briefly described as follows. For arbitrary starting points $x_0, \bar{x}_0 \in X$ we consider chains $(X_n)_{n \in \mathbb{N}_0}$, $(\bar{X}_n)_{n \in \mathbb{N}_0}$ with $X_0 = x_0$, $\bar{X}_0 = \bar{x}_0$ and try to build correlations between $(X_n)_{n \in \mathbb{N}_0}$ and $(\bar{X}_n)_{n \in \mathbb{N}_0}$ in order to make their trajectories as close as possible. This can be done because the transition probability function $B_{x,y}(A) = P\{(X_{n+1}, \bar{X}_{n+1}) \in A | (X_n, \bar{X}_n) = (x, y)\}$ of the coupled chain $(X_n, \bar{X}_n)_{n \in \mathbb{N}_0}$ taking values in $X^2$ can be decomposed (see [11]) in the following way

$$B_{x,y} = Q_{x,y} + R_{x,y},$$

where subprobability measures $Q_{x,y}$ are contractive, i.e. $\int_{X^2} d(u, v) Q_{x,y}(du, dv) \leq \alpha d(x, y)$ for some constant $\alpha \in (0, 1)$. 

\textend{raw}
Since transition probabilities for (1.1) can be mutually singular for even very close points, one cannot expect that chains \((X_n)_{n \in \mathbb{N}_0}\) and \((\bar{X}_n)_{n \in \mathbb{N}_0}\) couple in finite time (i.e. \(X_n = \bar{X}_n\) for some \(n \in \mathbb{N}_0\)) as in classical coupling constructions ([21]) leading to the convergence in the total variation norm. On the contrary, they only couple at infinity (i.e. \(d(X_n, \bar{X}_n) \to 0\) as \(n \to \infty\)) so this method is sometimes called asymptotic coupling ([13]) and gives convergence in weak topology.

There are two reasons which make random iterations with place dependent probabilities in some sense the critical case. First, when passing from constant to place dependent probabilities one loses the backward iterations argument, which combined with mutual singularity of transition probabilities makes establishing exponential ergodicity a difficult task. Secondly, such iterations are an example of the simplest systems where exponential ergodicity must be understood in the sense of convergence in weak topology. The other example is much more sophisticated: stochastic partial differential equations (SPDEs). In both cases the main difficulty is the same: mutual singularity of transition probabilities. The emerging theory of weak exponential ergodicity consists of many results which can be found in vast SPDEs literature (see [19], [20], [12], [34]), some of them ([13]) indicating analogies to existing theory of strong ergodicity summarized in [23]. However, to our best knowledge, none of these results applies in our case.

One of the most important consequences of exponential ergodicity (understood here in the sense of weak topology) is that it implies the Law of Large Numbers and the Central Limit Theorem. Shirikyan showed in [27] that if the process is exponentially ergodic in norm equivalent to the Fortet-Mourier norm then the strong version of LLN holds. Since convergence we obtain in this paper is in Fortet-Mourier norm, Shirikyan’s result can be applied in our case. Establishing the CLT is much more delicate problem. As Komorowski and Walczuk proved in [18], exponential ergodicity implies the CLT, however they use stronger Wassertein norm.
Recent results [14], [16] establish the CLT and exponential ergodicity simultaneously, using methods similar to ours. It is thus reasonable to expect that general theorem in the spirit of Komorowski and Walczuk and concerning the Fortet-Mourier norm can be formulated.

The paper is organized as follows. In Section 2 we formulate and prove a theorem which assures exponential convergence to an invariant measure for a class of Markov chains. This theorem is applied in Section 3 to chains generated by random iteration of functions. In Section 4 we discuss special class of such functions, random affine transformations on $\mathbb{R}^d$, thus generalizing the notion of perpetuity to the place dependent case.

2. AN EXPONENTIAL CONVERGENCE RESULT

2.1. Notation and basic definitions.

Let $(X, d)$ be a Polish space, i.e. a complete and separable metric space and denote by $B_X$ the $\sigma$-algebra of Borel subsets of $X$. By $B_b(X)$ we denote the space of bounded Borel-measurable functions equipped with the supremum norm, $C_b(X)$ stands for the subspace of bounded continuous functions. Let $\mathcal{M}_{\text{fin}}(X)$ and $\mathcal{M}_1(X)$ be the sets of Borel measures on $X$ such that $\mu(X) < \infty$ for $\mu \in \mathcal{M}_{\text{fin}}(X)$ and $\mu(X) = 1$ for $\mu \in \mathcal{M}_1(X)$. The elements of $\mathcal{M}_1(X)$ are called probability measures. The elements of $\mathcal{M}_{\text{fin}}(X)$ for which $\mu(X) \leq 1$ are called subprobability measures. By $\text{supp} \mu$ we denote the support of the measure $\mu$. We also define

$$\mathcal{M}_L(X) = \{\mu \in \mathcal{M}_1(X) : \int_X L(x)\mu(dx) < \infty\}$$

where $L : X \to [0, \infty)$ is an arbitrary Borel measurable function and

$$\mathcal{M}_1^x(X) = \{\mu \in \mathcal{M}_1(X) : \int_X d(\bar{x}, x)\mu(dx) < \infty\},$$

where $\bar{x} \in X$ is fixed. By the triangle inequality the definition of $\mathcal{M}_1^x(X)$ is independent of the choice of $\bar{x}$. 
The space $\mathcal{M}_1(X)$ is equipped with the Fortet-Mourier metric:

$$\|\mu_1 - \mu_2\|_{FM} = \sup_{X} \{|\int f(x)(\mu_1 - \mu_2)(dx)| : f \in \mathcal{F}\},$$

where $\mathcal{F} = \{f \in C_b(X) : |f(x) - f(y)| \leq 1 \text{ and } |f(x)| \leq 1 \text{ for } x, y \in X\}$. The space $\mathcal{M}^1_1(X)$ is equipped with the Wasserstein metric:

$$\|\mu_1 - \mu_2\|_{W} = \sup_{X} \{|\int f(x)(\mu_1 - \mu_2)(dx)| : f \in \mathcal{W}\},$$

where $\mathcal{W} = \{f \in C_b(X) : |f(x) - f(y)| \leq 1 \text{ for } x, y \in X\}$. By $\|\cdot\|$ we denote the total variation norm. If a measure $\mu$ is nonnegative then $\|\mu\|$ is simply the total mass of $\mu$. In Section 4 we will use Euclidean norm $|\cdot|$ in $\mathbb{R}^d$ and operator norm $\|\cdot\|_{op}$ given by $\|m\|_{op} = \sup\{|mx| : x \in \mathbb{R}^d, |x| = 1\}$.

Let $P : B_b(X) \to B_b(X)$ be a Markov operator, i.e. a linear operator satisfying $P1_X = 1_X$ and $Pf(x) \geq 0$ if $f \geq 0$. Denote by $P^*$ the the dual operator, i.e. operator $P^* : \mathcal{M}_{fin}(X) \to \mathcal{M}_{fin}(X)$ defined as follows

$$P^*\mu(A) := \int_X P1_A(x)\mu(dx) \text{ for } A \in B_X.$$ 

We say that a measure $\mu_* \in \mathcal{M}_1(X)$ is invariant for $P$ if $\int_X Pf(x)\mu_*(dx) = \int_X f(x)\mu_*(dx)$ for every $f \in B_b(X)$ or, alternatively, we have $P^*\mu_* = \mu_*$. By $\{P_x : x \in X\}$ we denote a transition probability function for $P$, i.e. a family of measures $P_x \in \mathcal{M}_1(X)$ for $x \in X$ such that the map $x \mapsto P_x(A)$ is measurable for every $A \in B_X$ and $Pf(x) = \int_X f(y)P_x(dy)$ for $x \in X$ and $f \in B_b(X)$ or equivalently $P^*\mu(A) = \int_X P_x(A)\mu(dx)$ for $A \in B_X$ and $\mu \in \mathcal{M}_{fin}(X)$.

### 2.2. Formulation of the theorem

**Definition 2.1.** A coupling for $\{P_x : x \in X\}$ is a family $\{B_{x,y} : x, y \in X\}$ of probability measures on $X \times X$ such that for every $B \in B_{X^2}$ the map $X^2 \ni (x,y) \mapsto B_{x,y}(B)$ is measurable and

$$B_{x,y}(A \times X) = P_x(A), \quad B_{x,y}(X \times A) = P_y(A)$$
for every \( x, y \in X \) and \( A \in \mathcal{B}_X \).

In the following we assume (see [11]) that there exists a family \( \{ Q_{x,y} : x, y \in X \} \) of subprobability measures on \( X^2 \) such that the map \( (x, y) \mapsto Q_{x,y}(B) \) is measurable for every \( B \in \mathcal{B}_X^2 \) and \( Q_{x,y}(A \times X) \leq P_x(A) \) and \( Q_{x,y}(X \times A) \leq P_y(A) \) for every \( x, y \in X \) and \( A \in \mathcal{B}_X \).

Measures \( \{ Q_{x,y} : x, y \in X \} \) allow us to construct a coupling for \( \{ P_x : x \in X \} \). Define on \( X^2 \) the family of measures \( \{ R_{x,y} : x, y \in X \} \) which on rectangles \( A \times B \) are given by
\[
R_{x,y}(A \times B) = \frac{1}{1 - Q_{x,y}(X^2)} (P_x(A) - Q_{x,y}(A \times X))(P_y(B) - Q_{x,y}(X \times B)),
\]
when \( Q_{x,y}(X^2) < 1 \) and \( R_{x,y}(A \times B) = 0 \) otherwise. A simple computation shows that the family \( \{ B_{x,y} : x, y \in X \} \) of measures on \( X^2 \) defined by
\[
(2.1) \quad B_{x,y} = Q_{x,y} + R_{x,y} \quad \text{for} \quad x, y \in X
\]
is a coupling for \( \{ P_x : x \in X \} \).

Now we list our assumptions on Markov operator \( P \) and transition subprobabilities \( \{ Q_{x,y} : x, y \in X \} \).

**A0** \( P \) is a Feller operator, i.e. \( P(C_b(X)) \subset C_b(X) \).

**A1** There exists a Lyapunov function for \( P \), i.e. continuous function \( L : X \to [0, \infty) \) such that \( L \) is bounded on bounded sets, \( \lim_{x \to \infty} L(x) = +\infty \) (for bounded \( X \) this condition is omitted) and for some \( \lambda \in (0, 1), c > 0 \)
\[
PL(x) \leq \lambda L(x) + c \quad \text{for} \quad x \in X.
\]

**A2** There exist \( F \subset X^2 \) and \( \alpha \in (0, 1) \) such that \( \text{supp} \, Q_{x,y} \subset F \) and
\[
(2.2) \quad \int_{X^2} d(u, v) Q_{x,y}(du, dv) \leq \alpha d(x, y) \quad \text{for} \quad (x, y) \in F.
\]

**A3** There exist \( \delta > 0, l > 0 \) and \( \nu \in (0, 1] \) such that
\[
(2.3) \quad 1 - \|Q_{x,y}\| \leq ld(x, y)^\nu \quad \text{and} \quad Q_{x,y}(\{(u, v) \in X^2 : d(u, v) < \alpha d(x, y)\}) \geq \delta
\]
for \((x,y) \in F\).

**A4** There exist \(\beta \in (0, 1), \hat{C} > 0\) and \(R > 0\) such that for

\[
\kappa((x_n, y_n)_{n \in \mathbb{N}_0}) = \inf\{n \in \mathbb{N}_0 : (x_n, y_n) \in F \text{ and } L(x_n) + L(y_n) < R\},
\]

where \((x_n, y_n)_{n \in \mathbb{N}_0}\) is a sequence of elements of \(X \times X\), we have

\[
\mathbb{E}_{x,y} \beta^{-\kappa} \leq \hat{C} \text{ whenever } L(x) + L(y) < \frac{4c}{1-\lambda},
\]

where \(\mathbb{E}_{x,y}\) denotes here the expectation with respect to the chain starting from \((x,y)\) and with transition function \(\{B_{x,y} : x, y \in X\}\).

Before proceeding to the formulation of the theorem we briefly comment above conditions. Assumption **A0** is standard. Existence of a Lyapunov function in **A1** allows to reduce the dynamics to some bounded region of \(X\). Assumption **A2** is well known contraction on average condition, written in the language of coupling. Measures \(Q_{x,y}\) are thus contractive part of coupled transition probabilities \(B_{x,y}\) with \(R_{x,y}\) being the uncontrollable part. Some version of such condition seems to be necessary for exponential ergodicity (see [25]). The first part of **A3** means that the total mass of the contractive part \(Q_{x,y}\) is close to 1 for \(x\) close to \(y\). Some higher order regularity here is necessary to ensure the uniqueness of invariant measure (see [28]). The second part of **A3** guarantees that some part of contractive on average measure \(Q_{x,y}\) is contractive in the strict sense. Condition **A4** means that the dynamics quickly enters the domain of contractivity \(F\). In this paper we discuss Markov chains generated by random iteration of functions for which always \(F = X^2\) and \(L(x) = d(x, \bar{x})\) with some fixed \(\bar{x} \in X\), so **A4** is trivially fulfilled when \(R = \frac{4c}{1-\lambda}\). There are, however, examples of random dynamical systems for which \(F\) is a proper subset of \(X^2\). Indeed, in *contractive Markov systems* introduced by I. Werner in [33] we have \(X = \sum_{i=1}^{n} X_i\) but \(F = \sum_{i=1}^{n} X_i \times X_i\). They are studied in [31].
Now we formulate the main result of this section. Its proof is given in Section 2.4.

**Theorem 2.1.** Assume \( A_0 - A_4 \). Then operator \( P \) possesses a unique invariant measure \( \mu^* \in \mathcal{M}_L(X) \), which is attractive, i.e.

\[
\lim_{n \to \infty} \int_X P^n f(x) \mu(dx) = \int_X f(x) \mu(dx) \quad \text{for} \quad f \in C_b(X), \mu \in \mathcal{M}_1(X).
\]

Moreover, there exist \( q \in (0, 1) \) and \( C > 0 \) such that

\[
\|P^* x \mu - \mu^*\|_{FM} \leq q^n C (1 + \int_X L(x) \mu(dx))
\]

for \( \mu \in \mathcal{M}_L(X) \) and \( n \in \mathbb{N} \).

**Remark.** In [13], Theorem 4.8, authors formulate sufficient conditions for the existence of a unique exponentially attractive invariant measure for continuous-time Markov semigroup \( \{P(t)\}_{t \geq 0} \), that do not refer to coupling. One of assumptions is that there exists distance-like (i.e. symmetric, lower semi-continuous and vanishing only on the diagonal) function \( d : X \times X \to [0, 1] \) which is contractive for some \( P(t_*) \), i.e. there exists \( \alpha < 1 \) such that for every \( x, y \in X \) with \( d(x, y) < 1 \) we have \( d(P(x, \cdot), P(y, \cdot)) \leq \alpha d(x, y) \), where \( P(\cdot, \cdot) : X \times B_X \to [0, 1] \) is transition kernel for \( P(t_*) \). This assumption is stronger than \( A_2 \), since measures \( R_{x,y} \) in [2.1] need not be contractive (i.e. \( \int_{X^2} d(u, v) R_{x,y}(du, dv) \leq \alpha d(u, v) \)) for any distance-like function \( d \).

2.3. Measures on the pathspace.

For fixed \((x_0, y_0) \in X^2\) the next step of a chain with transition probability function \( B_{x,y} = Q_{x,y} + R_{x,y} \) can be drawn according to \( Q_{x_0,y_0} \) or according to \( R_{x_0,y_0} \). To distinguish these two cases we introduce (see [11]) the augmented space \( \hat{X} = X^2 \times \{0, 1\} \) and the transition function \( \{(\hat{B}_{x,y,\theta} : (x, y, \theta) \in \hat{X})\} \) on \( \hat{X} \) given by \( \hat{B}_{x,y,\theta} = \hat{Q}_{x,y,\theta} + \hat{R}_{x,y,\theta} \), where \( \hat{Q}_{x,y,\theta} = Q_{x,y} \times \delta_1 \) and \( \hat{R}_{x,y,\theta} = R_{x,y} \times \delta_0 \). The parameter \( \theta \in \{0, 1\} \) is responsible for choosing measures \( Q_{x,y} \) and \( R_{x,y} \). If
Random iteration with place dependent probabilities

9

a Markov chain with transition function \( \{ \tilde{B}_{x,y,\theta} : (x, y, \theta) \in \tilde{X} \} \) stays in the set \( X^2 \times \{1\} \) at time \( n \) it means that the last step was drawn according to \( Q_{u,v} \), for some \((u, v) \in X^2 \).

For every \( x \in X \) finite-dimensional distributions \( P^0, \ldots, n \in M_1(X^{n+1}) \) are defined by

\[
P^0, \ldots, n_x(B) = \int_X \delta_x(dx_0) \int_X P_{x_1}(dx_2) \ldots \int_X P_{x_{n-1}}(dx_n) 1_B(x_0, \ldots, x_n)
\]

for \( n \in \mathbb{N}_0, B \in B_{X^{n+1}} \), where \( \delta_x \) is the Dirac measure at \( x \). By the Kolmogorov extension theorem we obtain the measure \( P_x^\infty \) on the pathspace \( X^\infty \). Similarly we define measures \( B_{x,y}^\infty, \tilde{B}_{x,y,\theta}^\infty \) on \( (X \times X)^\infty \) and \( \tilde{X}^\infty \). These measures have the following interpretation. Consider the Markov chain \( (X_n, Y_n)_{n \in \mathbb{N}_0} \) on \( X \times X \), starting from \((x_0, y_0)\), with the transition function \( \{ \mathcal{B}_{x,y} : x, y \in X \} \), obtained by canonical Kolmogorov construction, i.e. \( \Omega = (X \times X)^\infty \) is the sample space equipped with the probability measure \( \mathbb{P} = \mathcal{B}_{x_0,y_0}^\infty \), \( X_n(\omega) = x_n, Y_n(\omega) = y_n \), where \( \omega = (x_k, y_k)_{k \in \mathbb{N}_0} \in \Omega \), and \( n \in \mathbb{N}_0 \). Then \( (X_n)_{n \in \mathbb{N}_0}, (Y_n)_{n \in \mathbb{N}_0} \) are Markov chains in \( X \), starting from \( x_0 \) and \( y_0 \), with the transition function \( \{ \mathcal{P}_x : x \in X \} \), and \( \mathcal{P}_x^\infty, \mathcal{P}_y^\infty \) are their measures on the pathspace \( X^\infty \).

In this paper we often consider marginals of measures on the pathspace. If \( \mu \) is a measure on a measurable space \( X \) and \( f : X \rightarrow Y \) is a measurable map, then \( f^\# \mu \) is the measure on \( Y \) defined by \( f^\# \mu(A) = \mu(f^{-1}(A)) \). So, if we denote by \( pr \) the projection map from a product space to its component, then \( pr^\# \mu \) is simply the marginal of \( \mu \) on this component.

In the following we consider Markov chains on \( \tilde{X} \) with the transition function \( \{ \tilde{B}_{x,y,\theta} : x, y \in X, \theta \in \{0, 1\} \} \). We adopt the convention that \( \theta_0 = 1 \), so \( \Phi \) always starts from \( X^2 \times \{1\} \), and define \( \tilde{B}_{x,y}^\infty := \tilde{B}_{x,y,1}^\infty \).

For \( b \in \mathcal{M}_{fin}(X^2) \) we write: \( \tilde{B}_b^\infty(B) = \int_X \tilde{B}_{x,y}^\infty(B) b(dx, dy), B \in B_{X^\infty} \), \( Q_b(A) = \int_X Q_{x,y}(A) b(dx, dy), A \in B_{X^2} \) and \( Q_{x,y}^n(A) = Q_{x,y}^{n-1}(A), A \in B_{X^2} \).

When studying the asymptotics of a chain \( (X_n)_{n \in \mathbb{N}_0} \) with a transition function
\{P_x : x \in X\} it is particularly interesting whether a coupled chain \((X_n, Y_n)_{n \in \mathbb{N}_0}\)
is moving only according to the contractive part \(Q_{x,y}\) of the transition function \(B_{x,y}\). For every subprobability measure \(b \in \mathcal{M}_{fin}(X^2)\) we define finite-dimensional subprobability distributions \(Q^0_{b,\ldots,n} \in \mathcal{M}_{fin}((X \times X)^{n+1})\)

\[
Q^0_{b,\ldots,n}(B) = \int_{X^2} b(dx_0, dy_0) \int_{X^2} Q_{x_0,y_0}(dx_1, dy_1) \ldots 
\]

\[
\ldots \int_{X^2} Q_{x_{n-1},y_{n-1}}(dx_n, dy_n) \mathbf{1}_B((x_0, y_0), \ldots, (x_n, y_n)),
\]

where \(B \in \mathcal{B}(X \times X)^{n+1}, n \in \mathbb{N}_0\). For \(n \geq 1\) define \(\iota_n : (X \times X)^n \rightarrow \hat{X}^n\),

\[
\iota_n((x_1, y_1), \ldots, (x_n, y_n)) = ((x_1, y_1, 1), \ldots, (x_n, y_n, 1))
\]

and take \(\hat{Q}^0_{b,\ldots,n} = \#_n Q^0_{b,\ldots,n}\). Since \(Q^0_{b,\ldots,n}((X \times X)^{n+1}) > Q^0_{b,\ldots,n+1}((X \times X)^{n+2})\) the family \(\{\hat{Q}^0_{b,\ldots,n} : n \in \mathbb{N}_0\}\) is not consistent and we cannot use the Kolmogorov extension theorem to obtain a measure on the whole pathspace \(\hat{X}^\infty\). However, defining for every \(b \in \mathcal{M}_{fin}(X^2)\) the measure \(Q^\infty_{b} \in \mathcal{M}_{fin}(\hat{X}^\infty)\) by \(Q^\infty_{b}(B) = \hat{B}^\infty_{b}(B \cap (X^2 \times \{1\})^\infty)\), where \(B \in \mathcal{B}(\hat{X}^\infty)\), one can easily check that for every cylindrical set \(B = A \times \hat{X}^\infty, A \in \mathcal{B}(\hat{X}^n+1)\) we have

\[
Q^\infty_{b}(B) = \lim_{n \to \infty} \hat{Q}^0_{b,\ldots,n}(A).
\]

### 2.4. Proof of Theorem 2.1

Before proceeding to the proof of Theorem 2.1 we formulate two lemmas. The first one is partially inspired by the reasoning which can be found in [26].

**Lemma 2.1.** Let \(Y\) be a metric space and let \((Y_n)_{n \in \mathbb{N}_0}\) be a family of Markov chains indexed by starting point \(y \in Y\), with common transition function \(\{\pi_y : y \in Y\}\). Let \(V : Y \rightarrow [0, \infty)\) be a Lyapunov function for \(\{\pi_y : y \in Y\}\). Assume that for some bounded and measurable \(A \subset Y\) there exist \(\lambda \in (0, 1)\) and \(C_\rho > 0\) such that for

\[
\rho((y_n)_{n \in \mathbb{N}_0}) = \inf \{n \geq 1 : y_n \in A\}
\]
we have
\[ E_y \lambda^{-\rho} \leq C \rho (V(y_0) + 1) \quad \text{for} \quad y \in Y, \]
where \( E_y \) is the expectation with respect to the measure \( \mathbb{P}_y \) on \( Y^\infty \) induced by \((Y_n^y)_{n \in \mathbb{N}_0}\). Moreover, assume that for some measurable \( B \subset Y \) and \( \epsilon((y_n)_{n \in \mathbb{N}_0}) = \inf\{n \geq 1 : y_n \notin B\} \)
there exist constants \( p \in (0, 1), \beta \in (0, 1) \) and \( C_\epsilon > 0 \) such that
\[ \mathbb{P}_y(\{n \in \mathbb{N}_0 : \forall n \geq 1 y_n \in B\}) > p \quad \text{and} \quad E_y 1_{\{\tau < \infty\}} \beta^{-e} \leq C_\epsilon, \]
for every \( y \in A \). Then there exist \( \gamma \in (0, 1) \) and \( C > 0 \) such that for
\[ \tau((y_n)_{n \in \mathbb{N}_0}) = \inf\{n \geq 1 : \forall k \geq n y_k \in B\} \]
we have
\[ E_y \gamma^{-\tau} \leq C (V(y) + 1) \quad \text{for} \quad y \in Y. \]

**Proof of Lemma 2.7** Define \( \kappa = \epsilon + \rho \circ T_\epsilon \), where \( T_\epsilon ((y_k)_{k \in \mathbb{N}_0}) = (y_{k+n})_{k \in \mathbb{N}_0} \).
Fix \( y \in Y \), \( \alpha \in (0, 1) \) and \( r > 1 \) such that \((\lambda \alpha)^{-\frac{1}{\gamma+1}} < \beta^{-1}\). The strong Markov property and the Hölder inequality for every \( y \in Y \) give
\[
E_y 1_{\{\epsilon < \infty\}} \lambda^{-\gamma} \leq \left[ E_y (1_{\{\epsilon < \infty\}} (\lambda \alpha)^{-\frac{1}{\gamma+1}})^{\frac{\gamma+1}{\gamma}} \right]^{\frac{\gamma}{\gamma+1}} \leq \left[ E_y (1_{\{\epsilon < \infty\}} \alpha^\epsilon \lambda^{-\rho \circ T_\epsilon}) \right]^{\frac{\gamma}{\gamma+1}}
\]
\[
= \left[ E_y (1_{\{\epsilon < \infty\}} \beta^{-e}) \right]^{\frac{\gamma}{\gamma+1}} \left[ E_y (1_{\{\epsilon < \infty\}} \alpha^\epsilon \lambda^{-\rho \circ T_\epsilon} (\mathcal{F}_\epsilon)) \right]^{\frac{\gamma}{\gamma+1}}
\]
\[
= \left[ E_y (1_{\{\epsilon < \infty\}} \beta^{-e}) \right]^{\frac{\gamma}{\gamma+1}} \left[ E_y (1_{\{\epsilon < \infty\}} \alpha^\epsilon \lambda^{-\rho \circ T_\epsilon} (\mathcal{F}_\epsilon)) \right]^{\frac{\gamma}{\gamma+1}}
\]
where \( \mathcal{F}_\epsilon \) is the \( \sigma \)-algebra generated by \( \epsilon \). Since \( \sup_{y \in A} V(y) < \infty \) and \( V \) satisfies
\[ E_y (1_{\{\epsilon < \infty\}} \alpha^\epsilon V(Y^y_\epsilon)) \leq C_1 (V(y) + 1) \quad \text{for} \quad y \in Y, \]
for some \( C_1 > 0 \), taking \( c = \lambda^\frac{1}{\gamma} \) we obtain \( E_y (1_{\{\epsilon < \infty\}} \alpha^\epsilon) \leq C_2 \) whenever \( y \in A \), for some constant \( C_2 > 0 \). Define \( \epsilon_0 = 0, \kappa_0 = \rho \) and
\[
\epsilon_n = \kappa_{n-1} + \epsilon \circ T_{\kappa_{n-1}},
\]
\[
\kappa_n = \kappa_{n-1} + \kappa \circ T_{\kappa_{n-1}} \quad \text{for} \quad n \geq 1.
\]
Choosing sufficiently large $s$ and setting $\gamma = \lambda^{\frac{1}{2}}$ we have $\mathbb{E}_y \gamma^{\frac{1}{2}} \leq C(1 + V(y))$ for $y \in Y$. Since $\tau < \hat{\tau}$, the proof is complete.
Lemma 2.2. Let \((Y_n^y)_{n \in \mathbb{N}_0}\) with \(y \in Y\) be a family of Markov chains on a metric space \(Y\). Suppose that \(V : Y \rightarrow [0, \infty)\) is a Lyapunov function for their transition function \(\{\pi_y : y \in Y\}\), i.e. there exist \(a \in (0, 1)\) and \(b > 0\) such that

\[
\int_Y V(x)\pi_y(dx) \leq aV(y) + b \quad \text{for} \quad y \in Y.
\]

Then there exist \(\lambda \in (0, 1)\) and \(\tilde{C} > 0\) such that for

\[
\rho((y_k)_{k \in \mathbb{N}_0}) = \inf\{k \geq 1 : V(y_k) < \frac{2b}{1-a}\}
\]

we have

\[
\mathbb{E}_y \lambda^{-\rho} \leq \tilde{C}(V(y_0) + 1) \quad \text{for} \quad y \in Y.
\]

Proof of Lemma 2.2. Let us suppose that the chains \((Y_n^y)_{n \in \mathbb{N}_0}, y \in Y\) are defined on a common probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Fix \(\alpha \in (\frac{1+a}{2}, 1)\) and set \(V_0 = \frac{b}{\alpha-a}\). Define \(\tilde{\rho}((y_k)_{k \in \mathbb{N}_0}) = \inf\{k \geq 1 : V(y_k) \leq V_0\}\). Fix \(y \in Y\). Let \(\mathcal{F}_n \subset \mathcal{F}, n \in \mathbb{N}_0\) be the filtration induced by \((Y_n^y)_{n \in \mathbb{N}_0}\). Define

\[A_n = \{\omega \in \Omega : V(Y_i^y(\omega)) > V_0 \quad \text{for} \quad i = 0, 1, \ldots, n\}, \quad n \in \mathbb{N}_0.\]

Observe that \(A_{n+1} \subset A_n\) and \(A_n \in \mathcal{F}_n\). By the definition of \(V_0\) we have \(1_{A_n} \mathbb{E}(V(Y_{n+1}^y)|\mathcal{F}_n) \leq 1_{A_n}(aV(Y_{n}^y) + b) < \alpha 1_{A_n}V(Y_{n}^y) \mathbb{P}\text{-a.e. in } \Omega\). This gives

\[
\int_{A_n} V(Y_{n}^y)d\mathbb{P} \leq \int_{A_{n-1}} V(Y_{n}^y)d\mathbb{P} = \int_{A_{n-1}} \mathbb{E}(V(Y_{n}^y)|\mathcal{F}_{n-1})d\mathbb{P} \leq \alpha \int_{A_{n-1}} (aV(Y_{n-1}^y) + b)d\mathbb{P} \leq \alpha \int_{A_{n-1}} V(Y_{n-1}^y)d\mathbb{P}.
\]

By the Chebyshev inequality

\[
\mathbb{P}(V(Y_{0}^y) > V_0, \ldots, V(Y_{n}^y) > V_0) = \int_{A_{n-1}} \mathbb{P}(V(Y_{n}^y) > V_0|\mathcal{F}_{n-1})d\mathbb{P} \leq V_0^{-1} \int_{A_{n-1}} \mathbb{E}(V(Y_{n}^y)|\mathcal{F}_{n-1})d\mathbb{P} \leq \alpha^{n-1}V_0^{-1}(aV(y) + b),
\]

thus for some \(C > 0\) we have

\[
\mathbb{P}_y(\tilde{\rho} > n) \leq \alpha^n C(V(y) + 1), \quad n \in \mathbb{N}_0.
\]
Fix $\gamma \in (0, 1)$ and observe that for $\lambda = \alpha^\gamma$ we have

$$
E_y \lambda^{-\rho} \leq 2 + \sum_{n=1}^{\infty} \mathbb{P}_y(\lambda^{-\rho} > n) = 2 + \sum_{n=1}^{\infty} \mathbb{P}_y(\tilde{\rho} > \log_\alpha (n^{-\frac{1}{\gamma}}))
$$

$$
\leq 2 + \sum_{n=1}^{\infty} \mathbb{P}_y(\tilde{\rho} > \lceil \log_\alpha (n^{-\frac{1}{\gamma}}) \rceil) \leq 2 + \frac{C(V(y) + 1)}{\alpha} \sum_{n=1}^{\infty} n^{-\frac{1}{\gamma}}
$$

$$
= \tilde{C}(V(y) + 1)
$$

for properly chosen $\tilde{C}$. Since $\rho \leq \tilde{\rho}$, the proof is finished. \hfill \square

**Proof of Theorem 2.1.**

For every $r > 0$ define $D_r = \{(x, y) \in X^2 : d(x, y) < r\}$.

**Step I:** To simplify calculations define a new metric $\bar{d}(x, y) = d(x, y)^{\nu}$ and observe that for $\bar{D}_r = \{(x, y) \in X^2 : \bar{d}(x, y) < r\}$ we have $D_R = \bar{D}_\bar{R}$ with $\bar{R} = R^{\nu}$. By the Jensen inequality (2.2) takes form

$$
(2.6) \quad \int_{X^2} \bar{d}(u, v) Q_{x,y}(du, dv) \leq \bar{\alpha} \bar{d}(x, y) \quad \text{for} \quad (x, y) \in F,
$$

with $\bar{\alpha} = \alpha^\nu$. Assumption A3 implies that

$$
(2.7) \quad 1 - \|Q_{x,y}\| \leq l\bar{d}(x, y) \quad \text{and} \quad Q_{x,y}(D_{\bar{\alpha}\bar{d}(x,y)}) \geq \delta
$$

for $(x, y) \in F$.

**Step II:** Observe, that if $b \in M_{fin}(X^2)$ satisfies $\text{supp} \ b \subset F$ then (2.7) implies

$$
\|Q_b\| \geq \|b\| - l \int_{X^2} \bar{d}(u, v)b(du, dv).
$$

Since $\text{supp} \ b \subset F$, we have $\text{supp} \ Q_b^{0,\ldots, n} \subset F^{n+1}$. Iterating the above inequality and using (2.6) we obtain $\|Q_b^{0,\ldots, n}\| \geq \|b\| - \frac{l}{1-\bar{\alpha}} \int_{X^2} \bar{d}(u, v)b(du, dv)$. If $\text{supp} \ b \subset \{(u, v) \in X^2 : \bar{d}(u, v) < \frac{1-\bar{\alpha}}{2l}\} \cap F$ then from (2.5) it follows that

$$
(2.8) \quad \|Q_b^\infty\| \geq \frac{1}{2} \|b\|.
$$

Set $R_0 = \sup\{\bar{d}(x, y) : L(x) + L(y) < R\} < \infty$ and $n_0 = \min\{n \in \mathbb{N}_0 : \bar{\alpha}^n R_0 < \frac{1-\bar{\alpha}}{2l}\}$. Now (2.7) implies that for $(x, y) \in F$ such that $L(x) + L(y) < R$ we have

$$
Q_{x,y}^{n_0}(\{(u, v) \in X^2 : \bar{d}(u, v) < \frac{1-\bar{\alpha}}{2l}\} \cap F) > \delta^{n_0}
$$
and finally (2.8) gives

\[ \| Q_{x,y}^\infty \| \geq \frac{1}{2} \delta^{no}. \]

**Step III:** Define \( \tilde{\rho}( (x_n, y_n)_{n \in \mathbb{N}_0} ) = \inf \{ n \geq 1 : L(x_n) + L(y_n) < \frac{4c}{1-\lambda} \} \). Since \( L(x) + L(y) \) is a Lyapunov function for a Markov chain in \( X^2 \) with transition probabilities \( \{ B_{x,y} : x, y \in X \} \), Lemma 2.2 shows that there exist constants \( \lambda_0 \in (0, 1) \) and \( C_0 \) such that

\[ (2.10) \quad \mathbb{E}_{x,y} \lambda_0^{-\tilde{\rho}} \leq C_0 (L(x) + L(y) + 1) \quad \text{for} \quad (x, y) \in X^2. \]

Define \( A = \{(x, y, \theta) \in \hat{X} : (x, y) \in F \quad \text{and} \quad L(x) + L(y) < R \} \) and

\[ \rho((x_n, y_n, \theta_n)_{n \in \mathbb{N}_0}) = \inf \{ n \in \mathbb{N}_0 : (x_n, y_n, \theta_n) \in A \}. \]

Slightly abusing notation, we identify \( \kappa : (X \times X)^\infty \rightarrow \mathbb{N}_0 \cup \{ \infty \} \), defined in A4, with its extension to \( \hat{X}^\infty : \kappa((x_n, y_n, \theta_n)_{n \in \mathbb{N}_0}) = \kappa((x_n, y_n)_{n \in \mathbb{N}_0}) \). Since \( \rho \leq \tilde{\rho} + \kappa \circ T_{\tilde{\rho}} \), where \( T_{\tilde{\rho}}((x_n, y_n, \theta_n)_{n \in \mathbb{N}_0}) = (x_{n+\tilde{\rho}}, y_{n+\tilde{\rho}}, \theta_{n+\tilde{\rho}})_{n \in \mathbb{N}_0} \), an argument similar to that in the proof of Lemma 2.1 shows that there exist \( \lambda \in (0, 1) \) such that

\[ \mathbb{E}_{x,y,\theta} \lambda^{-\rho} \leq \tilde{C} C_0 (L(x) + L(y) + 1) \quad \text{for} \quad x, y \in X, \theta \in \{0, 1\}. \]

Define \( B = \{(x, y, \theta) \in \hat{X} : \theta = 1 \} \) and

\[ \epsilon((x_n, y_n, \theta_n)_{n \in \mathbb{N}_0}) = \inf \{ n \geq 1 : (x_n, y_n, \theta_n) \notin B \}. \]

From Step II we obtain \( \mathbb{P}_{x,y,\theta}(B) \geq \frac{1}{2} \delta^{\epsilon_0} \) for \( (x, y, \theta) \in A \). From (2.6) and (2.7) it follows that

\begin{align*}
\hat{B}_{x,y,\theta}(\epsilon = n) &= \int_{\hat{X}^n} \int_{\hat{X}^n} (\hat{X}) Q_{x,y}^{d_0,\ldots,d_n} (d\alpha_0, \ldots, d\alpha_{n-1}) \\
&= \| Q_{x,y}^{d_0,\ldots,d_n} - \hat{Q}_{x,y}^{d_0,\ldots,d_n} \| \leq l \int_{X^2} d(u,v) Q_{x,y}^{d_0,\ldots,d_n} (du, dv) \\
&\leq l \alpha^{n-1} d(x,y) < \alpha^{n-1} l R_0,
\end{align*}
whenever \((x, y, \theta) \in A\). Finally Lemma 2.1 guarantees the existence of constants 
\(\gamma \in (0, 1), C_1 > 0\) such that for

\[
\tau((x_n, y_n, \theta_n)_{n \in \mathbb{N}_0}) = \inf \{n \geq 1 : \forall k \geq n (x_k, y_k, \theta_k) \in B\}
\]

we have

\[
E_{x, y, \theta} \gamma^{-\tau} \leq C_1 (L(x) + L(y) + 1) \quad \text{for} \quad x, y \in X, \theta \in \{0, 1\}.
\]

**STEP IV:** Define sets 
\(G_n = \{t \in (X^2 \times \{0, 1\})^\ast : \tau(t) \leq \frac{n}{2}\}\) and 
\(H_n = \{t \in (X^2 \times \{0, 1\})^\ast : \tau(t) > \frac{n}{2}\}\). For every 
\(n \in \mathbb{N}\) we have

\[
\hat{B}_{x, y, \theta} = \hat{B}_{x, y, \theta} \mid_{G_2} + \hat{B}_{x, y, \theta} \mid_{H_2} \quad \text{for} \quad x, y \in X, \theta \in \{0, 1\}.
\]

Fix \(\theta = 1\) and \((x, y) \in X^2\). From the fact that \(\| \cdot \|_{FM} \leq \| \cdot \|_W\) it follows that

\[
\|P^n \delta_x - P^n \delta_y\|_{FM} = \|P^n_x - P^n_y\|_{FM}
\]

\[
= \sup_{f \in F} \int_{X^2} (f(z_1) - f(z_2))(pr_{x} \hat{B}_{x, y, \theta} \mid_{G_2})(dz_1, dz_2)\mid
\]

\[
= \sup_{f \in W} \int_{X^2} (f(z_1) - f(z_2))(pr_{x} \hat{B}_{x, y, \theta} \mid_{G_2})(dz_1, dz_2)\mid
\]

\[
\leq \sup_{f \in W} \int_{X^2} (f(z_1) - f(z_2))(pr_{x} \hat{B}_{x, y, \theta} \mid_{G_2})(dz_1, dz_2)\mid + 2\hat{B}_{x, y, \theta}(H_2).
\]

From A2 we obtain

\[
\sup_{f \in W} \int_{X^2} (f(z_1) - f(z_2))(pr_{x} \hat{B}_{x, y, \theta} \mid_{G_2})(dz_1, dz_2)\mid
\]

\[
\leq \int_{X^2} d(z_1, z_2)(pr_{x} \hat{B}_{x, y, \theta} \mid_{G_2})(dz_1, dz_2)\mid
\]

\[
\leq \alpha \int_{X^2} d(z_1, z_2)(pr_{x} \hat{B}_{x, y, \theta} \mid_{G_2})(dz_1, dz_2)\mid \leq \alpha \frac{n}{2} R.
\]

Now Step III and the Chebyshev inequality imply that

\[
\hat{B}_{x, y, \theta}(H_2) \leq \gamma \frac{n}{2} C_1 (L(x) + L(y) + 1) \quad \text{for} \quad n \in \mathbb{N}.
\]
Taking $C_2 = 2C_1 + R$ and $q = \max\{\gamma x, \alpha x\}$ we obtain

$$\|P^{*n}\delta_x - P^{*n}\delta_y\|_{FM} \leq \gamma^n C_1 (L(x) + L(y) + 1) \quad \text{for} \quad x, y \in X, n \in \mathbb{N},$$

and so

$$\|P^{*n}\mu - P^{*n}\nu\|_{FM} \leq \gamma^n C_1 \left( \int_X L(x)\mu(dx) + \int_X L(y)\nu(dy) + 1 \right)$$

for $\mu, \nu \in \mathcal{M}_1(X)$ and $n \in \mathbb{N}$.

**Step V:** Observe that Step IV and A1 give

$$\|P^{*n}\delta_x - P^{*n+k}\delta_x\|_{FM} \leq \int_X \|P^{*n}\delta_x - P^{*n}\delta_y\|_{FM} P^{*k}\delta_x(dy)$$

$$\leq q^n C_2 \left( \int_X (L(x) + L(y))P^{*k}\delta_x(dy) \leq q^n C_3 (1 + L(x)), \right)$$

so $(P^{*n}\delta_x)_{n \in \mathbb{N}}$ is a Cauchy sequence for every $x \in X$. Since $\mathcal{M}_1(X)$ equipped with the norm $\| \cdot \|_{FM}$ is complete (see [8]), assumption A0 implies the existence of an invariant measure $\mu_*$. Assumption A1 gives $\mu_* \in \mathcal{M}_L(X)$. Applying inequality (2.11) we obtain (2.4). Observation that $\mathcal{M}_L(X)$ is dense in $\mathcal{M}_1(X)$ in the total variation norm finishes the proof. \qed

**Remark.** In steps IV and V of the above proof we follow M. Hairer (see [11]).

### 3. RANDOM ITERATION OF FUNCTIONS

Let $(X, d)$ be a Polish space and $(\Theta, \Xi)$ a measurable space with a family $\vartheta_x \in \mathcal{M}_1(\Theta)$ of distributions on $\Theta$ indexed by $x \in X$. Space $\Theta$ serves as a set of indices for a family $\{S_\theta : \theta \in \Theta\}$ of continuous functions acting on $X$ into itself. We assume that $(\theta, x) \mapsto S_\theta(x)$ is product measurable. In this section we study some stochastically perturbed dynamical system $(X_n)_{n \in \mathbb{N}_0}$. Its intuitive description is the following: if $X_0$ starts at $x_0$, then by choosing $\theta_0$ at random from $\vartheta_{x_0}$ we define $X_1 = S_{\theta_0}(x_0)$. Having $X_1$ we select $\theta_1$ according to the distribution $\vartheta_{X_1}$ and we put $X_2 = S_{\theta_1}(X_1)$ and so on. More precisely, the process $(X_n)_{n \in \mathbb{N}_0}$ can
be written as

\[ X_{n+1} = S_{Y_n}(X_n), \quad n = 0, 1, \ldots, \]

where \((Y_n)_{n \in \mathbb{N}_0}\) is a sequence of random elements defined on a probability space \((\Omega, \Sigma, \text{prob})\) with values in \(\Theta\) such that

\[ \text{prob}(Y_n \in B | X_n = x) = \vartheta_x(B) \quad \text{for} \quad x \in X, B \in \Xi, n = 0, 1, \ldots, \]

and \(X_0 : \Omega \to X\) is a given random variable. Denoting by \(\mu_n\) the probability law of \(X_n\), we will give a recurrence relation between \(\mu_{n+1}\) and \(\mu_n\). To this end fix \(f \in B_b(X)\) and note that

\[ \mathbb{E} f(X_{n+1}) = \int_X f d\mu_{n+1}. \]

By (3.1) we have

\[ \int_A \vartheta_x(B) \mu_n(dx) = \text{prob}(\{Y_n \in B\} \cap \{X_n \in A\}) \quad \text{for} \quad B \in \Xi, A \in B_X, \]

hence

\[ \mathbb{E} f(X_{n+1}) = \int \int_X f(S \theta)(x) \vartheta_x(d\theta) \mu_n(dx). \]

Putting \(f = 1_A, A \in B_X\), we obtain \(\mu_{n+1}(A) = P^* \mu_n(A)\), where

\[ P^* \mu(A) = \int \int_X 1_A(S \theta)(x) \vartheta_x(d\theta) \mu(dx) \quad \text{for} \quad \mu \in M_{fin}(X), A \in B_X. \]

In other words this formula defines the transition operator for \(\mu_n\). Operator \(P^*\) is adjoint of the Markov operator \(P : B_b(X) \to B_b(X)\) of the form

\[ Pf(x) = \int \vartheta_x(d\theta). \]

We take this formula as the precise formal definition of considered process. We will show that operator (3.2) has a unique invariant measure, provided the following conditions hold:

**B1** There exists \(\alpha \in (0, 1)\) such that

\[ \int \vartheta_x(d\theta) \leq \alpha d(x, y) \quad \text{for} \quad x, y \in X. \]
There exists $\bar{x} \in X$ such that

$$c := \sup_{x \in X} \int_{\Theta} d(S_{\theta}(\bar{x}), \bar{x}) \vartheta_x(d\theta) < \infty.$$ 

The map $x \mapsto \vartheta_x$, $x \in X$, is Hölder continuous in the total variation norm, i.e. there exists $l > 0$ and $\nu \in (0, 1]$ such that

$$\|\vartheta_x - \vartheta_y\| \leq l \, d(x, y)^\nu \quad \text{for} \quad x, y \in X.$$ 

There exists $\delta > 0$ such that

$$\vartheta_x \land \vartheta_y(\{\theta \in \Theta : d(S_{\theta}(x), S_{\theta}(y)) \leq \alpha d(x, y)\}) > \delta \quad \text{for} \quad x, y \in X,$$

where $\land$ denotes the greatest lower bound in the lattice of finite measures.

Remark. In the simplest case $\Theta = \{1, ..., n\}$ above conditions are standard and nearly optimal (see [30] for discussion).

**Proposition 3.1.** Assume B1 – B4. Then operator (3.2) possesses a unique invariant measure $\mu_\ast \in \mathcal{M}_1^1(X)$, which is attractive in $\mathcal{M}_1^1(X)$. Moreover there exist $q \in (0, 1)$ and $C > 0$ such that

$$\|P^n \mu - \mu_\ast\|_{FM} \leq q^n C(1 + \int_X d(\bar{x}, x) \mu(dx))$$

for $\mu \in \mathcal{M}_1^1(X)$ and $n \in \mathbb{N}$.

**Proof.** Define the operator $Q$ on $B_b(X^2)$ by

$$Q(f)(x, y) = \int_{\Theta} f(S_{\theta}(x), S_{\theta}(y)) \vartheta_x \land \vartheta_y(d\theta).$$

Since

$$\|\vartheta_x \land \vartheta_y - \vartheta_x \land \vartheta_y\| \leq 2(||\vartheta_x - \vartheta_x|| + ||\vartheta_y - \vartheta_y||)$$
it follows that

\[ |Q(f)(x', y') - Q(f)(x, y)| \leq \int |f(S_\theta(x'), S_\theta(y'))| (d_\theta^x \wedge d_\theta^y) \leq \alpha d(x, y) \]

for \( f \in B_b(X^2), x, y \in X \). Consequently, we see that \( Q(C_b(X^2)) \subset C_b(X^2) \), by Lebesgue's dominated convergence theorem. Put

\[
F = \{ f \in B_b(X^2) : \sup_{z \in X^2} |f(z)| \leq M, Q(f) \in B_b(X^2) \},
\]

where \( M > 0 \) is fixed, and observe that the family \( F \) contains all continuous functions bounded by \( M \) and the limit of any convergent sequence of functions in the class, i.e. \( F \) consists, by the definition, of all Baire functions bounded by \( M \). Since the class of Baire functions is identical with the class of Borel functions (see [22, Theorem 4.5.2]) it follows that \( Q(B_b(X^2)) \subset B_b(X^2) \). In particular, for the family \( \{Q_{x,y} : x, y \in X\} \) of (subprobability) measures given by

\[
Q_{x,y}(C) = \int_C 1_C(S_\theta(x), S_\theta(y)) d_\theta^x \wedge d_\theta^y \]

we have that maps \( (x, y) \mapsto Q_{x,y}(C) \) are measurable for every \( C \in \mathcal{B}_{X^2} \).

Arguing similarly as above we show that (3.2) is well defined Feller operator. It has Lyapunov function \( L(x) = d(x, \bar{x}) \), since \( \int_\Theta d(S_\theta(x), S_\theta(y)) d_\theta = \alpha d(x, \bar{x}) + c \). Now, observe that

\[
\|Q_{x,y}\| = \vartheta_x \wedge \vartheta_y(\Theta) = \inf_{A \in \Xi} \{ \vartheta_x(A) + \vartheta_y(\Theta \setminus A) \}
\]

\[
= 1 - \sup_{A \in \Xi} \{ \vartheta_x(A) - \vartheta_y(A) \} = 1 - \| \vartheta_x - \vartheta_y \| \geq 1 - l d(x, y)\]

for \( x, y \in X \). Moreover, we have

\[
\int_{X^2} d(u, v)Q_{x,y}(du, dv) = \int_\Theta d(S_\theta(x), S_\theta(y)) d_\theta^x \wedge d_\theta^y \leq \alpha d(x, y),
\]
and $Q_{x,y}(D_{ad(x,y)}) = \vartheta_x \land \vartheta_y(\{\theta \in \Theta : d(S_\theta(x), S_\theta(y)) \leq \alpha d(x, y)\}) > \delta$ for $x, y \in X$. In consequence $A_0 - A_3$ are fulfilled. The use of Theorem 2.1 (see also comments in subsection 2.2 concerning assumption $A_4$) ends the proof. □

4. PERPETUITIES WITH PLACE DEPENDENT PROBABILITIES

Let $X = \mathbb{R}^d$ and $G = \mathbb{R}^{d \times d} \times \mathbb{R}^d$, and consider the function $S_\theta : X \to X$ defined by $S_\theta(x) = M(\theta)x + Q(\theta)$, where $(M, Q)$ is a random variable on $(\Theta, \Xi)$ with values in $G$. Then (3.2) may be written as

\begin{equation}
Pf(x) = \int_G f(mx + q)d\vartheta \circ (M, Q)^{-1}(m, q)
\end{equation}

This operator is connected with the random difference equation of the form

\begin{equation}
\Phi_n = M_n\Phi_{n-1} + Q_n, \quad n = 1, 2, \ldots,
\end{equation}

where $(M_n, Q_n)_{n \in \mathbb{N}}$ is a sequence of independent random variables distributed as $(M, Q)$. Namely, the process $(\Phi_n)_{n \in \mathbb{N}_0}$ is a homogeneous Markov chain with the transition kernel $P$ given by

\begin{equation}
Pf(x) = \int_G f(mx + q)d\mu(m, q),
\end{equation}

where $\mu$ stands for the distribution of $(M, Q)$. Equation (4.2) arises in various disciplines as economics, physics, nuclear technology, biology, sociology (see e.g. [32]). It is closely related to a sequence of backward iterations $(\Psi_n)_{n \in \mathbb{N}}$, given by $\sum_{k=1}^n M_1 \ldots M_{k-1}Q_k, n \in \mathbb{N}$ (see e.g. [9]). Under conditions ensuring the almost sure convergence of the sequence $(\Psi_n)_{n \in \mathbb{N}}$ the limiting random variable

\begin{equation}
\sum_{n=1}^\infty M_1 \ldots M_{n-1}Q_n
\end{equation}

is often called perpetuity. It turns out that the probability law of (4.4) is a unique invariant measure for (4.3). The name perpetuity comes from perpetual payment
streams and recently gained some popularity in the literature on stochastic recurrence equations (see [7]). In the insurance context a perpetuity represents the present value of a permanent commitment to make a payment at regular intervals, say annually, into the future forever. The $Q_n$ represent annual payments, the $M_n$ cumulative discount factors. Many interesting examples of perpetuities can be found in [1]. Due to significant papers [17], [10], [32] and [9] we have complete (in the dimension one) characterization of convergence of perpetuities. The rate of this convergence has recently been extensively studied by many authors (see for instance [3]-[5], [24]). The main result of this section concerns the rate of convergence of the process $(X_n)_{n\in\mathbb{N}_0}$ associated with the operator $P : B_b(\mathbb{R}^d) \to B_b(\mathbb{R}^d)$ given by

\begin{equation}
(4.5) \quad Pf(x) = \int_G f(mx + q)d\mu_x(m, q),
\end{equation}

where $\{\mu_x : x \in \mathbb{R}^d\}$ is a family of Borel probability measures on $G$. In contrast to $(\Phi_n)_{n\in\mathbb{N}_0}$, the process $(X_n)_{n\in\mathbb{N}_0}$ moves by choosing at random $\theta$ from a measure depending on $x$. Taking into considerations the concept of perpetuities we may say that $(X_n)_{n\in\mathbb{N}_0}$ forms a perpetuity with place dependent probabilities.

**Corollary 4.1.** Assume that $\{\mu_x : x \in \mathbb{R}^d\}$ is a family of Borel probability measures on $G$ such that

\begin{equation}
(4.6) \quad \alpha := \sup_{x\in\mathbb{R}^d} \int_G ||m||_\text{op} d\mu_x(m, q) < 1, \quad c := \sup_{x\in\mathbb{R}^d} \int_G |q|d\mu_x(m, q) < \infty.
\end{equation}

Assume moreover that the map $x \mapsto \mu_x, x \in X$, is Hölder continuous in the total variation norm and there exists $\delta > 0$ such that

$$\mu_x \land \mu_y((m, q) \in G : ||m||_\text{op} \leq \alpha) > \delta \quad \text{for} \quad x, y \in \mathbb{R}^d.$$ 

Then operator (4.5) possesses a unique invariant measure $\mu_* \in \mathcal{M}_1(\mathbb{R}^d)$, which is attractive in $\mathcal{M}_1(\mathbb{R}^d)$. Moreover there exist $q \in (0, 1)$ and $C > 0$ such that

$$\|P^{*n}\mu - \mu_*\|_{FM} \leq q^nC(1 + \int_{\mathbb{R}^d} |x|\mu(dx)).$$
for $\mu \in M^1_1(\mathbb{R}^d)$ and $n \in \mathbb{N}$.

The proof of corollary is straightforward application of Proposition 3.1. We leave the details to the reader. We finish the paper by giving an example to illustrate Corollary 4.1.

**Example.** Let $\nu_0$, $\nu_1$ be distributions on $\mathbb{R}^2$. Assume that $p, q : \mathbb{R} \to [0, 1]$ are Lipschitz functions (with Lipschitz constant $L$) summing up to 1, for $x \leq 0, p(x) = 0$, for $x \geq 1$. Define $\mu_x$ by $\mu_x = p(x)\nu_0 + q(x)\nu_1, x \in \mathbb{R}$. Then:

1. $\|\mu_x - \mu_y\| \leq 2L|x - y|$ for $x, y \in \mathbb{R}$.
2. If $\int_{\mathbb{R}^2} |m|d\nu_i(m, q) < 1$ and $\int_{\mathbb{R}^2} |q|d\nu_i(m, q) < \infty$ for $i = 0, 1$, then (4.6) holds.
3. For every $A \in B_{\mathbb{R}^2}$, $x, y \in \mathbb{R}$ we have: $\mu_x \wedge \mu_y(A) \geq \nu_0 \wedge \nu_1(A) = (\nu_0 - \lambda^+)(A) = (\nu_1 - \lambda^-)(A) \geq \max\{\nu_0(A), \nu_1(A)\} - \|\nu_0 - \nu_1\|(A)$, where $(\lambda^+, \lambda^-)$ is the Jordan decomposition of $\nu_1 - \nu_0$.

**Acknowledgment.** This work was partially supported by the Faculty of Applied Mathematics AGH UST statutory tasks within subsidy of Ministry of Science and Higher Education.

**REFERENCES**


[24] M. Mirek, Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps,
Random iteration with place dependent probabilities


AGH University of Science and Technology, Faculty of Applied Mathematics
al. A. Mickiewicza 30, 30-059 Krakow, Poland
E-mail: rafal.kapica@agh.edu.pl

Institute of Mathematics, University of Silesia in Katowice
Bankowa 14, 40-007 Katowice, Poland
E-mail: sleczka@math.us.edu.pl

Received on 6.11.2017;
revised version on xx.xx.xxxx