UNIVERSITATIS IAGELLONICAE ACTA MATHEMATICA, FASCICULUS XLIV
2006

GENERIC PROPERTIES OF ITERATED FUNCTION SYSTEMS WITH PLACE DEPENDENT PROBABILITIES

by Tomasz Bielaczyc

Abstract. It is shown that a learning system defined on a compact convex subset of $\mathbb{R}^n$ such that the Hausdorff dimension of its invariant measure is equal to zero is typical in the family of all systems satisfying the average contractivity condition.

1. Introduction. Generic properties of Markov operators have been studied in [**2, 6, 7, 10**]. Szarek [**10**] has proved that a typical continuous Markov operator acting on the space of Borel measures defined on a Polish space is asymptotically stable and its stationary distribution has the Hausdorff dimension equal to zero.

In this paper we investigate iterated function systems with place dependent probabilities defined on a compact convex subset of $\mathbb{R}^n$, known as learning systems. Generic properties of iterated function systems have been studied by Lasota and Myjak [**2**], who have proved that a typical nonexpansive iterated function system is asymptotically stable and its stationary distribution is singular. Szarek [**7**] has generalized this result to learning systems satisfying the average contractivity condition: $\lambda_{(S,p)} = \max_{x\in X} \sum_{i=1}^N p_i(x) L_i \le 1$.

We prove a more general result. Namely, for most of learning systems satisfying the average contractivity condition the Hausdorff dimension of the stationary distribution is equal to zero. We use the method developed by Lasota and Myjak [**2**], and Szarek [**7**].

1991 *Mathematics Subject Classification.* Primary 60J05, 28A80; Secondary 47A35, 58F08.

Key words and phrases. Markov operators, iterated function systems, Hausdorff dimension.

The organization of the paper is as follows. In Section 2 we introduce definitions and notation. Section 3 contains auxiliary lemmas which are used in proving the main result of the paper. The main theorem is proved in Section 4.

2. Preliminaries. Let $X \subset \mathbb{R}^k$ be a compact convex set. By $B(x,r)$ we denote the open ball with center at $x \in X$ and radius $r > 0$. Given a set $A \subset X$ and a number $r > 0$, by $\operatorname{diam} A$, we denote the diameter of the set A and by $B(A,r)$ the r-neighbourhood of the set A, i.e.

$$B(A,r) = \{x \in X : \rho(x,A) < r\},$$

where $\rho(x,A) = \inf\{\rho(x,y) : y \in A\}$.

Let $\mathcal{B}(X)$ denote the σ-algebra of all Borel subsets of X and let $\mathcal{M}$ denote the family of all finite Borel measures on X. By $\mathcal{M}_1$, we denote the set of all $\mu \in \mathcal{M}$ such that $\mu(X) = 1$. The elements of $\mathcal{M}_1$ will be called *distributions*.

Let $\mathcal{M}_s = \{\mu_1 - \mu_2 : \mu_1, \mu_2 \in \mathcal{M}\}$ be the space of all finite signed Borel measures on X. For every $l \geq 1$ we introduce the *Fortet–Mourier norm* (see [**1**])

$$\|\mu\|_l = \sup\{|\langle f, \mu\rangle| : f \in F_l\},$$

where $\langle f, \mu\rangle = \int_X f(x)\mu(dx)$ and F_l is the space of all continuous functions $f : X \to \mathbb{R}$ such that $\sup_{x\in X} |f(x)| \leq 1$ and $|f(x) - f(y)| \leq l\|x - y\|$ (here $\|\cdot\|$ denotes a norm in $\mathbb{R}^k$).

An operator $P : \mathcal{M} \to \mathcal{M}$ is called a *Markov operator* if it satisfies the following two conditions:

(i) *positive linearity*:

$$P(\lambda_1\mu_1 + \lambda_2\mu_2) = \lambda_1 P\mu_1 + \lambda_2 P\mu_2$$

for $\lambda_1, \lambda_2 \geq 0$ and $\mu_1, \mu_2 \in \mathcal{M}$,

(ii) *preservation of the norm*:

$$P\mu(X) = \mu(X) \quad \text{for } \mu \in \mathcal{M}.$$

An operator $P : \mathcal{M} \to \mathcal{M}$ is called *nonexpansive in the norm* $\|\cdot\|_l$, $l \geq 1$, if

$$\|P\mu_1 - P\mu_2\|_l \leq \|\mu_1 - \mu_2\|_l \quad \text{for } \mu_1, \mu_2 \in \mathcal{M}_1.$$

A measure $\mu \in \mathcal{M}$ is called *stationary* or *invariant* if $P\mu = \mu$. A Markov operator P is called *asymptotically stable* if there exists a stationary distribution $\mu_\star$ such that

$$\lim_{n\to\infty} \langle f, P^n\mu\rangle = \langle f, \mu_\star\rangle \quad \text{for } \mu \in \mathcal{M}_1, f \in C(X)$$

(here $C(X)$ stands for the space of all continuous functions $f : X \to \mathbb{R}$). Clearly, the stationary distribution is unique provided P is asymptotically stable.

A Markov operator P is called a *Feller operator* if there is a linear operator $U\colon C(X) \to C(X)$ (dual to P) such that

$$\langle Uf, \mu\rangle = \langle f, P\mu\rangle \quad \text{for } f \in C(X), \mu \in \mathcal{M}.$$

For $A \subset X$ and $s, \delta > 0$, define

$$\mathcal{H}^s_\delta(A) = \inf\Big\{\sum_{i=1}^{\infty}(\operatorname{diam} U_i)^s\colon\ A \subset \bigcup_{i=1}^{\infty} U_i,\ \operatorname{diam} U_i \le \delta\Big\}$$

and

$$\mathcal{H}^s(A) = \lim_{\delta\to 0} \mathcal{H}^s_\delta(A).$$

The restriction of $\mathcal{H}^s$ to the σ-algebra of $\mathcal{H}^s$-measurable sets is called the *s-dimensional Hausdorff measure.* Note that all Borel sets are $\mathcal{H}^s$-measurable. The value

$$\dim_H A = \inf\{s > 0\colon \mathcal{H}^s(A) = 0\}$$

is called the *Hausdorff dimension* of the set A. As usual, we admit $\inf \emptyset = +\infty$.

The *Hausdorff dimension* of a measure $\mu \in \mathcal{M}_1$ is defined by the formula

$$\dim_H \mu = \inf\{\dim_H A\colon A \in \mathcal{B}(X), \mu(A) = 1\}.$$

Fix an integer $N \ge 1$. By an *iterated function system with place dependent probabilities* (or shorter, *learning system*)

$$(S, p) = (S_1, \dots, S_N, p_1, \dots .p_N)$$

we mean a finite sequence of continuous transformations $S_i\colon X \to X$ and continuous functions $p_i\colon X \to [0, 1], i = 1, \dots, N$, such that $\sum_{i=1}^N p_i(x) = 1$. A sequence $(p_i)_{i=1}^N$ as above is called a *probability vector.* We assume that S_i is lipschitzian with a Lipschitz constant L_i for $i = 1, \dots, N$.

For a learning system (S, p), we define the value

$$\lambda_{(S,p)} = \max_{x\in X} \sum_{i=1}^N p_i(x)L_i.$$

We denote by $\mathcal{F}$ the set of all learning systems (S, p) such that $\lambda_{(S,p)} \le 1$. In $\mathcal{F}$, we introduce a metric d defined by

$$d((S, p), (T, q)) = \sum_{i=1}^N \max_{x\in X} |p_i(x) - q_i(x)| + \sum_{i=1}^N \max_{x\in X} \|S_i(x) - T_i(x)\|$$

for $(S, p), (T, q) \in \mathcal{F}$. It is easy to prove that $\mathcal{F}$ endowed with the metric d is a complete metric space.

For a given learning system (S,p), we define the corresponding Markov operator $P_{(S,p)}\colon \mathcal{M} \to \mathcal{M}$ by

$$P_{(S,p)}\mu(A) = \sum_{i=1}^{N} \int_{S_i^{-1}(A)} p_i(x)\mu(dx) \quad \text{for } \mu \in \mathcal{M}, A \in \mathcal{B}(X)$$

and its dual $U_{(S,p)}\colon C(X) \to C(X)$ by

$$U_{(S,p)}f(x) = \sum_{i=1}^{N} p_i(x) f(S_i(x)) \quad \text{for } f \in C(X), x \in X.$$

We say that a learning system (S,p) has a *stationary distribution* (resp. is *asymptotically stable*) if the corresponding Markov operator $P_{(S,p)}$ has a stationary distribution (resp. is asymptotically stable).

Finally recall that a subset of a complete metric space $\mathcal{X}$ is called *residual* if its complement is a set of first Baire category. A property is said to be satisfied by most elements of the space $\mathcal{X}$ if it is satisfied on a residual subset. Such a property is also called *generic* or *typical*.

3. Auxiliary results.

LEMMA 3.1. *Let $\mu_1, \mu_2 \in \mathcal{M}_1$, $l \geq 1$ and $\epsilon > 0$. If $\|\mu_1 - \mu_2\|_l \leq \epsilon^2$, then*

$$\mu_1(B(A,\epsilon)) \geq \mu_2(A) - \epsilon$$

for every $A \in \mathcal{B}(X)$.

The lemma follows from [**11**, Lemma 3.1].

Consider now an asymptotically stable learning system (S,p). Let $P = P_{(S,p)}$ and $\mu = \mu_{(S,p)}$ denote the corresponding Markov operator and invariant distribution, respectively. We define

$$X_{(S,p)} = \{x \in X \colon \mu(\{x\}) > 0\}.$$

LEMMA 3.2. *For every $x \in X_{(S,p)}$ and $i \in \{1, \dots, N\}$,*

$$p_i(x) > 0 \Longrightarrow S_i(x) \in X_{(S,p)}. \tag{3.1}$$

PROOF. Fix $x \in X_{(S,p)}$ and $i \in \{1, \dots, N\}$. Assume that $p_i(x) > 0$. Using the definition of the corresponding Markov operator and the fact that μ is invariant, we obtain

$$\mu(S_i(x)) \geq p_i(x)\mu(S_i^{-1}(S_i(x))) \geq p_i(x)\mu(\{x\}) > 0.$$

The proof is complete. $\square$

LEMMA 3.3. *If $\mu(X_{(S,p)}) > 0$, then $\mu(X_{(S,p)}) = 1$.*

The lemma follows from Theorem 2.1 in [**3**]. For the reader's convienience we shall give a proof of the lemma (see also the proof of Proposition 2.3 in [**5**]).

Proof. Consider the measure $\widehat{\mu}\colon \mathcal{B}(X) \to I\!R$ given by

$$\widehat{\mu}(A) = \frac{\mu(A \cap X_{(S,p)})}{\mu(X_{(S,p)})} \quad \text{for } A \in \mathcal{B}(X).$$

Then

$$\begin{aligned} P\widehat{\mu}(A) &= \sum_{i=1}^{N} \int_{S_i^{-1}(A)} p_i(x)\widehat{\mu}(dx) \\ &= \frac{1}{\mu(X_{(S,p)})} \sum_{i=1}^{N} \int_{S_i^{-1}(A)\cap X_{(S,p)}} p_i(x)\mu(dx) \end{aligned}$$

for $A \in \mathcal{B}(X)$. From Lemma 3.2 there follows that

$$\{x \in X_{(S,p)}\colon p_i(x) > 0\} \subset S_i^{-1}(X_{(S,p)}) \quad \text{for } i \in \{1, \dots, N\}.$$

Using this inclusion, we obtain

$$\begin{aligned} P\widehat{\mu}(A) &\le \frac{1}{\mu(X_{(S,p)})} \sum_{i=1}^{N} \int_{S_i^{-1}(A\cap X_{(S,p)})} p_i(x)\mu(dx) \\ &= \frac{1}{\mu(X_{(S,p)})} P_{(S,p)}\mu(A \cap X_{(S,p)}) = \frac{\mu(A \cap X_{(S,p)})}{\mu(X_{(S,p)})} = \widehat{\mu}(A) \end{aligned}$$

for $A \in \mathcal{B}(X)$. Since $P\widehat{\mu}(X) = \widehat{\mu}(X) = 1$, this implies that $P\widehat{\mu} = \widehat{\mu}$. From the uniqueness of the invariant distribution it follows that $\widehat{\mu} = \mu$. Consequently, $\mu(X_{(S,p)}) = 1$. □

Corollary 3.1. *If $\mu(X_{(S,p)}) > 0$, then for every $\epsilon > 0$ there exists a finite set $Z_\epsilon \subset X_{(S,p)}$ such that $\mu(Z_\epsilon) > 1 - \epsilon$.*

Proof. Since $\mu(\{x\}) > 0$ for every $x \in X_{(S,p)}$, it is easy to see that the set $X_{(S,p)}$ is countable. By Lemma 3.3, there is

$$1 = \mu(X_{(S,p)}) = \mu\Big(\bigcup_{x \in X_{(S,p)}} \{x\}\Big) = \sum_{x \in X_{(S,p)}} \mu(\{x\})$$

and the statement of the corollary follows. □

Lemma 3.4. *Let (S, p) be a learning system from $\mathcal{F}$ with the following properties:*

(3.2) $\qquad \lambda_{(S,p)} < 1,$

(3.3) $\qquad$ *p_i is lipschitzian and $p_i(x) > 0$ for $i \in \{1, \dots, N\}$ and $x \in X$.*

Then there exists $l \geq 1$ such that the corresponding Markov operator $P_{(S,p)}$ is nonexpansive in the norm $\| \cdot \|_l$.

Similarly as in the proof of Remark 3.1 in [**7**], a simple calculation shows that the statement of the lemma holds for

$$l = \max \Big\{ \frac{L_p}{1 - \lambda_{(S,p)}}, 1 \Big\},$$

where $L_p = \max_{1 \leq i \leq N} \operatorname{Lip} p_i$.

LEMMA 3.5. *Let P be a nonexpansive Markov operator. Assume that for every $\epsilon > 0$ there are a Borel set A with $\operatorname{diam} A \leq \epsilon$, a real number $\alpha > 0$ and an integer n such that*

$$P^n \mu(A) \geq \alpha \quad \text{for } \mu \in \mathcal{M}_1. \tag{3.4}$$

Then P is asymptotically stable and

$$\|P^{k \cdot n}(\mu_1 - \mu_2)\|_l \leq \epsilon + 2(1 - \alpha)^k \quad \text{for } \mu_1, \mu_2 \in \mathcal{M}_1, k \in \mathbb{N}. \tag{3.5}$$

For details, see the proof of Theorem 3.1 in [**8**]. In fact, Theorem 3.1 was proved for $l = 1$, but the same argument works for every $l \geq 1$.

The proof of the next lemma is based on the proof of Theorem 4.2 in [**9**].

LEMMA 3.6. *Assume that a learning system (S, p) satisfies conditions (3.2) and (3.3). Then (S, p) is asymptotically stable and for every $\epsilon > 0$ there exists an integer n such that*

$$\|P^n_{(S,p)}\mu_1 - P^n_{(S,p)}\mu_2\|_l < \epsilon \quad \text{for } \mu_1, \mu_2 \in \mathcal{M}_1,$$

where l is defined as in Lemma 3.4.

PROOF. Fix $\epsilon > 0$. From Lemma 3.4 there follows that $P_{(S,p)}$ is nonexpansive in the norm $\| \cdot \|_l$. From (3.2) it follows that there exists $i_0 \in \{1, \ldots, N\}$ such that S_{i_0} is contractive. Thus there exists an integer n such that $\operatorname{diam} S^n_{i_0}(X) < \epsilon/2$. By an induction argument, it is easy to verify that

$$\begin{aligned} P^n_{(S,p)}\mu(S^n_{i_0}(X)) &= \langle 1\!\!1_{S^n_{i_0}(X)}, P^n_{(S,p)}\mu \rangle = \langle U^n_{(S,p)} 1\!\!1_{S^n_{i_0}(X)}, \mu \rangle \\ &= \sum_{i_1, \ldots, i_n = 1}^{N} \int_X p_{i_1}(x) \ldots (p_{i_n} \circ S_{i_{n-1}} \circ \ldots \circ S_{i_1})(x) \\ &\qquad \times 1\!\!1_{S^n_{i_0}(X)}(S_{i_n} \circ \ldots \circ S_{i_1})(x) \mu(dx) \\ &\geq (\inf_{x \in X} p_{i_0}(x))^n \mu(X) = (\inf_{x \in X} p_{i_0}(x))^n \end{aligned}$$

for every $\mu \in \mathcal{M}_1$. Consequently, by (3.3) we obtain

$$P^n_{(S,p)}\mu(S^n_{i_0}(X)) \geq (\inf_{x\in X} p_{i_0}(x))^n > 0 \quad \text{for } \mu \in \mathcal{M}_1.$$

Hence the assumptions of Lemma 3.5 are satisfied. Thus the system (S,p) is asymptotically stable and inequality (3.5) holds. Let $k \in \mathbb{N}$ be such that

$$\big(1 - (\inf_{x\in X} p_{i_0}(x))^n\big)^k < \frac{\epsilon}{4}.$$

Using (3.5) with $\epsilon/2$ instead of ϵ finishes the proof. □

Let $\mathcal{F}_*$ be the set of all $(S,p) \in \mathcal{F}$ satisfying (3.2) and (3.3) and such that

$$\mu_{(S,p)}(X_{(S,p)}) > 0, \tag{3.6}$$

where $\mu_{(S,p)}$ is the stationary distribution corresponding to (S,p).

Since the set $\mathcal{F}_*$ is contained in $\mathcal{F}_0$ defined in [**7**], the next lemma is more general then Lemma 3.1 from [**7**].

Lemma 3.7. *The set $\mathcal{F}_*$ is dense in the space $(\mathcal{F}, d)$.*

Proof. Fix $(T,q) \in \mathcal{F}$. Let $\epsilon > 0$ be such that

$$\epsilon \cdot \left(\frac{1}{N} + \max_{1\leq i\leq N} \mathrm{Lip} T_i\right) < 2.$$

Fix $z \in X$. Since X is convex, for $i \in \{1,\dots,N\}$, we can define a new transformation $\widehat{T}_i\colon X \to X$ by

$$\widehat{T}_i(x) = \alpha z + (1-\alpha)T_i(x) \quad \text{for } x \in X,$$

where $\alpha = \epsilon(4N \operatorname{diam} X)^{-1}$. It follows immediately that

$$d((\widehat{T},q),(T,q)) \leq \epsilon/4 \tag{3.7}$$

and $\lambda_{(\widehat{T},q)} \leq 1-\alpha$. Thus there exists $i_0 \in \{1,\dots,N\}$ such that $\widehat{L}_{i_0} :=\mathrm{Lip}\widehat{T}_{i_0} < 1$. Let $x_0 \in X$ be a fixed point of $\widehat{T}_{i_0}$. By [**2**, Lemma 3.5], we can find a Lipschitz transformation $\widehat{S}\colon X \to X$ with a Lipschitz constant $L_{\widehat{S}}$ and with the following properties:

$$L_{\widehat{S}} < \widehat{L}_{i_0} + \eta, \tag{3.8}$$

$$\max_{x\in X} \|\widehat{S}(x) - \widehat{T}_{i_0}(x)\| < \epsilon/4, \tag{3.9}$$

$$\widehat{S}(x) = x_0 \quad \text{for } \|x - x_0\| \leq r, \tag{3.10}$$

where $\eta, r > 0$ and

$$\eta \leq \frac{\min\{1-\lambda_{(\widehat{T},q)}, 1-\widehat{L}_{i_0}\}}{2}.$$

By [**7**, Lemma 2.3], there exists a probability vector $(p_1, \dots, p_N)$ satisfying (3.3) and such that

$$\max_{x\in X} |p_i(x) - q_i(x)| < \frac{\epsilon \cdot \eta}{2N} \quad \text{for } i \in \{1, \dots, N\}. \tag{3.11}$$

Consider now the learning system $(S, p) = (S_1, \dots, S_N; p_1, \dots, p_N)$, where $S_i = \widehat{T}_i$ for $i \in \{1, \dots, N\} \setminus \{i_0\}$ and $S_{i_0} = \widehat{S}$. From (3.7), (3.9) and (3.11) there follows immediately that

$$d((S, p), (T, q)) < \epsilon.$$

Since

$$\eta\Big(\frac{\epsilon}{2} \max_{1\le i\le N} \mathrm{Lip}T_i + \sup_{x\in X} q_{i_0}(x) + \frac{\epsilon \cdot \eta}{2N}\Big) < 1 - \lambda_{(\widehat{T},q)},$$

there is

$$\sup_{x\in X} \sum_{i=1}^{N} q_i(x)\mathrm{Lip}\widehat{T}_i + N \cdot \frac{\eta \cdot \epsilon}{2N} \max_{1\le i\le N} \mathrm{Lip}T_i + \sup_{x\in X} q_{i_0}(x) \cdot \eta + \frac{\eta \cdot \epsilon}{2N} \cdot \eta < 1.$$

Thus the system (S, p) has property (3.2). It remains to verify that the stationary distribution $\mu_{(S,p)}$ satisfies condition (3.6).

Let $P_{(S,p)}$ be the Markov operator corresponding to (S, p) and let $\mu_{(S,p)}$ be its stationary distribution. From (3.8) and (3.10) it follows that there exists an integer n such that

$$S^n_{i_0}(X) = \{x_0\}.$$

There is

$$\begin{aligned}
\mu_{(S,p)}(\{x_0\}) &= P^n_{(S,p)}\mu_{(S,p)}(\{x_0\}) \\
&= \sum_{i_1,\dots,i_n=1}^{N} \int_X p_{i_1}(x) \dots (p_{i_n} \circ S_{i_{n-1}} \circ \dots \circ S_{i_1})(x) \\
&\qquad\qquad \times 1\!\!1_{\{x_0\}}(S_{i_n} \circ \dots \circ S_{i_1})(x)\mu_{(S,p)}(dx) \\
&\ge (\min_{x\in X} p_{i_0}(x))^n \mu_{(S,p)}(X) = (\min_{x\in X} p_{i_0}(x))^n > 0.
\end{aligned}$$

The proof is complete. □

Lemma 3.8. *Assume that (S, p) is an asymptotically stable learning system and $\mu = \mu_{(S,p)}$ is the corresponding invariant distribution. Let $\alpha \ge 0$. If*

$$\mu\Big(\Big\{x \in X\colon \liminf_{r\to 0} \frac{\log \mu(B(x, r))}{\log r} \le \alpha\Big\}\Big) = 1, \tag{3.12}$$

then $\dim_H \mu \le \alpha$.

Proof. (See also the proof of Proposition 2.4 in [**5**].) Let $\alpha \geq 0$ be such that (3.12) holds. Define

$$K_\alpha = \bigl\{x \in X\colon \liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} \leq \alpha\bigr\}.$$

To complete the proof it is enough to show that $\dim_H K_\alpha \leq \alpha$. Let $s > \alpha$ be arbitrary. Fix $\delta > 0$. Since X is compact, we can choose a finite sequence $x_1, \dots, x_m \in X$ such that

$$K_\alpha \subset \bigcup_{k=1}^{m} B(x_k, r_k) \quad \text{and} \quad \frac{\ln \mu(B(x_k, r_k))}{\ln r_k} \leq s, \tag{3.13}$$

where $r_k < \min\{1, \delta/6\}$ for $k = 1, \dots m$. Without loss of generality, we may assume that $r_1 \geq r_2 \geq \dots \geq r_m$. By an induction argument, we can find a subset $\{k_1, \dots, k_q\}$ of $\{1, \dots, m\}$ such that

$$B(x_{k_i}, r_{k_i}) \cap B(x_{k_j}, r_{k_j}) = \emptyset \qquad \text{for } i, j \in \{1, \dots, q\}, i \neq j$$

and

$$K_\alpha \subset \bigcup_{i=1}^{q} B(x_{k_i}, 3r_{k_i}).$$

Consequently, using (3.13) we obtain

$$\sum_{i=1}^{q} (r_{k_i})^s \leq \sum_{i=1}^{q} \mu(B(x_{k_i}, r_{k_i})) \leq \mu(X) \leq 1$$

and

$$\mathcal{H}^s_\delta(K_\alpha) \leq \sum_{k=1}^{q} 6^s (r_{k_i})^s \leq 6^s.$$

Since $\delta > 0$ was arbitrary, it follows that $\mathcal{H}^s(K_\alpha) \leq 6^s < \infty$. In turn, since $s > \alpha$ was arbitrary, it follows that $\dim_H(K_\alpha) \leq \alpha$. The proof is complete. □

4. Main theorem.

Theorem 4.1. *The set $\widehat{\mathcal{F}}_*$ of all $(T, q) \in \mathcal{F}$ such that its unique invariant distribution $\mu_{(T,q)}$ satisfies $\dim_H \mu_{(T,q)} = 0$ is residual in $\mathcal{F}$.*

Proof. Fix $n \in \mathbb{N}$ and $(S, p) \in \mathcal{F}_*$. Let $P_{(S,p)}$ be the Markov operator corresponding to (S, p) and let $\mu_{(S,p)}$ be its stationary distribution. From Corollary 3.1 there follows that there exists a finite set $Z_{(S,p),n} \subset X_{(S,p)}$ such that

$$\mu_{(S,p)}(Z_{(S,p),n}) > 1 - 1/n.$$

Choose $r_{(S,p),n} \in (0,1/n)$ such that

$$r_{(S,p),n}^{1/n} \le \mu_{(S,p)}\big(B(x, r_{(S,p),n})\big) \quad \text{for } x \in Z_{(S,p),n}. \tag{4.1}$$

Further, by Lemma 3.4, there exists $l_{(S,p)} \ge 1$ such that $P_{(S,p)}$ is nonexpansive in the norm $\|\cdot\|_{l_{(S,p)}}$. By Lemma 3.6, there exists $k_{(S,p),n} \in I\!N$ such that

$$\|P_{(S,p)}^{k_{(S,p),n}}\mu_1 - P_{(S,p)}^{k_{(S,p),n}}\mu_2\|_{l_{(S,p)}} \le \frac{r_{(S,p),n}^2}{8} \quad \text{for } \mu_1, \mu_2 \in \mathcal{M}_1. \tag{4.2}$$

By Lemma 3.2 in [**7**], there exists $\delta_{(S,p),n} > 0$ such that for all $(T,q) \in \mathcal{F}$,

$$d\big((S,p),(T,q)\big) < \delta_{(S,p),n} \Longrightarrow$$

$$\sup_{f \in F_{l_{(S,p)}}, x \in X} |U_{(S,p)}^{k_{(S,p),n}} f(x) - U_{(T,q)}^{k_{(S,p),n}} f(x)| < \frac{r_{(S,p),n}^2}{8}. \tag{4.3}$$

Define

$$\widehat{\mathcal{F}} = \bigcap_{n=1}^{\infty} \bigcup_{(S,p) \in \mathcal{F}_*} B_{\mathcal{F}}\big((S,p), \delta_{(S,p),n}\big),$$

where $B_{\mathcal{F}}((S,p),\delta_{(S,p),n})$ is the open ball in $(\mathcal{F}, d)$ with center at (S,p) and radius $\delta_{(S,p),n}$. From Lemma 3.7 it follows that $\widehat{\mathcal{F}}$ is the intersection of a countably family of open dense sets. Consequently, $\widehat{\mathcal{F}}$ is residual. We are going to show that $\widehat{\mathcal{F}} \subset \widehat{\mathcal{F}}_*$.

Fix $(T,q) \in \widehat{\mathcal{F}}$. By Theorem 3.1 in [**4**], (T,q) has an invariant distribution. Let $P_{(T,q)}$ and $\mu_{(T,q)}$ denote the corresponding Markov operator and stationary distribution, respectively. Let $((S,p)_n)_{n \in I\!N}$ be a sequence of learning systems of $\mathcal{F}_*$ such that

$$(T,q) \in B_{\mathcal{F}}\big((S,p)_n, \delta_{(S,p)_n,n}\big) \quad \text{for } n \in I\!N.$$

Assume that $r_{(S,p)_n,n} \in (0,1/n)$, $l_{(S,p)_n} > 1$ and $k_{(S,p)_n,n} \in I\!N$ are such that (4.1), (4.2) hold for $P_{(S,p)_n}$ and $\mu_{(S,p)_n}$. We set more compact notation:

$$\begin{aligned} P_n &= P_{(S,p)_n}, & \mu_n &= \mu_{(S,p)_n}, & Z_n &= Z_{(S,p)_n,n}, \\ r_n &= r_{(S,p)_n,n}, & l_n &= l_{(S,p)_n}, & k_n &= k_{(S,p)_n,n}. \end{aligned}$$

By (4.2) and (4.3), for every $n \in I\!N$, $\mu \in \mathcal{M}_1$ and $m \ge k_n$ there holds

$$\begin{aligned} \|P_{(T,q)}^m\mu - \mu_{(T,q)}\|_1 &\le \|P_{(T,q)}^m\mu - \mu_{(T,q)}\|_{l_n} \\ &\le \|P_{(T,q)}^{k_n}P_{(T,q)}^{m-k_n}\mu - P_n^{k_n}P_{(T,q)}^{m-k_n}\mu\|_{l_n} \\ &+ \|P_n^{k_n}P_{(T,q)}^{m-k_n}\mu - P_n^{k_n}\mu_{(T,q)}\|_{l_n} \\ &+ \|P_n^{k_n}\mu_{(T,q)} - P_{(T,q)}^{k_n}\mu_{(T,q)}\|_{l_n} \le \frac{1}{n}. \end{aligned}$$

Consequently, (T,q) is asymptotically stable. From (4.2) and (4.3), we also derive:

$$\begin{aligned}\|\mu_{(T,q)}-\mu_n\|_{l_n} &= \|P^{k_n}_{(T,q)}\mu_{(T,q)}-P_n^{k_n}\mu_n\|_{l_n}\\ &\le \|P^{k_n}_{(T,q)}\mu_{(T,q)}-P_n^{k_n}\mu_{(T,q)}\|_{l_n}+\|P_n^{k_n}\mu_{(T,q)}-P_n^{k_n}\mu_n\|_{l_n}\\ &\le \frac{r_n^2}{4}.\end{aligned}$$

Moreover, for every $y\in B(Z_n,r_n)$, there exists $x\in Z_n$ such that

$$B(x,r_n)\subset B(y,2r_n).$$

Consequently, by (4.1) and Lemma 3.1, we obtain

$$\begin{aligned}\mu_{(T,q)}\big(B(y,3r_n)\big) &\ge \mu_n\big(B(y,2r_n)\big)-\frac{r_n}{2}\ge \mu_n\big(B(x,r_n)\big)-\frac{r_n}{2}\\ &\ge r_n^{1/n}-\frac{r_n^{1/n}}{2}=\frac{r_n^{1/n}}{2}.\end{aligned}\tag{4.4}$$

Define

$$Y=\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}B(Z_n,r_n).$$

By the definition of Z_n and Lemma 3.1, we obtain

$$\mu_{(T,q)}\big(B(Z_n,r_n)\big)\ge \mu_n(Z_n)-r_n\ge 1-1/n-1/n=1-2/n.$$

Consequently, $\mu_{(T,q)}(Y)=1$. On the other hand, if $y\in Y$, then $y\in B(Z_n,r_n)$ for infinitely many $n\in I\!N$ and by (4.4), we can choose a sequence of integers $(s_n)_{n\in I\!N}$ such that

$$\mu_{(T,q)}\big(B(y,3r_{s_n})\big)\ge \frac{r_{s_n}^{1/s_n}}{2}\qquad \text{for } n\in I\!N.$$

Hence

$$\liminf_{n\to\infty}\frac{\log\mu_{(T,q)}\big(B(y,3r_{s_n})\big)}{\log 3r_{s_n}}\le \lim_{n\to\infty}\frac{\log(\frac{r_{s_n}^{1/s_n}}{2})}{\log 3r_{s_n}}=0.$$

Since $y\in Y$ was arbitrarily chosen and $\mu_{(T,q)}(Y)=1$, by Lemma 3.8, there is $\dim_H\mu_{(T,q)}=0$. The proof is complete. □

References

1. Dudley R.M., *Probabilities and Metrics*, Aarhus Universitet, 1976.
2. Lasota A., Myjak J., *Generic properties of fractal measures*, Bull. Pol. Acad. Sci.; Math., **42** (1994), 283–296.
3. Lasota A., Myjak A., Szarek T., *Markov operators with a unique invariant measure*, J. Math. Anal. Appl., **276** (2002), 343–356.

4. Lasota A., Yorke J.A., *Lower bound technique for Markov operators and iterated function systems*, Random. Comput. Dynam., **2** (1994), 41–77.
5. Myjak J., Szarek T., *On Hausdorff dimension of invariant measures arising from non-contractive iterated function systems*, Ann. Mat. Pura Appl. IV, Ser. **181** (2002), 223–237.
6. Szarek T., *Generic properties of continous iterated function systems*, Bull. Pol. Acad. Sci.; Math., **57** (1999), 77–89.
7. Szarek T., *Generic properties of learning systems*, Ann. Polon. Math., **73** (2000), 93–102.
8. Szarek T., *Iterated function systems depending on a previous transformation*, Univ. Iagel. Acta Math., **33** (1996), 161–172.
9. Szarek T., *Markov operators acting on Polish spaces*, Ann. Polon. Math., **67** (1997), 247–257.
10. Szarek T., *On typical Markov operators acting on Borel measures*, Abstr. Appl. Anal., **5** (2005), 489–497.
11. Szarek T., *The stability of Markov operators on Polish spaces*, Studia Math., **143** (2000), 145–152.

Received May 28, 2004

Silesian University
Institute of Mathematics
40-007 Katowice
Poland
e-mail: bielaczyc@ux2.math.us.edu.pl