Proc. Indian Acad. Sci. (Math. Sci.) Vol. 119, No. 5, November 2009, pp. 669–677.

Upper packing dimension of a measure and the limit distribution of products of i.i.d. stochastic matrices

ARUNAVA MUKHERJEA[1] and RICARDO RESTREPO[2]

[1]Department of Mathematics, The University of Texas – Pan American, 1201 West University Drive, Edinburg, TX 78539, USA
[2]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA
[2]Departamento de Matemáticas, Universidad de Antioquia, Calle 67 N°53-108, Medellín, Colombia, USA
E-mail: arunava.mukherjea@gmail.com; ricrest@gmail.com

MS received 17 March 2009; revised 9 April 2009

Abstract. This article gives sufficient conditions for the limit distribution of products of i.i.d. 2×2 stochastic matrices to be continuous singular, when the support of the distribution of the individual random matrices is countably infinite. It extends a previous result for which the support of the random matrices is finite. The result is based on adapting existing proofs in the context of attractors and iterated function systems to the case of infinite iterated function systems.

Keywords. Packing dimension; stochastic matrices; Erdös sum; products of random matrices; continuous singularity of the limit distribution.

1. Introduction

In Theorem 3.1, page 20 of [6] the following theorem was proven:

Theorem 1.1. *Let μ be a probability measure on $d \times d$ stochastic matrices such that $S(\mu)$ consists of only invertible matrices (that is, matrices with rank d) and the convolution sequence $\mu^{(n)}$ converges weakly to some probability measure λ. Then λ has exactly one of the following properties:*

(1) *λ is discrete (that is $\lambda(E) = 1$ for some countable subset E).*
(2) *λ is continuous singular with respect to m_0, where m_0 is the restriction of the Lebesgue measure to an appropriate subset of an Euclidean space of appropriate dimension.*
(3) *λ is absolutely continuous with respect to m_0.*

If, furthermore, the matrices in $S(\mu)$ are ergodic and λ is not degenerate, then λ is continuous.

Examples illustrating this result are also given in [6]. Now let μ be a probability measure on (the Borel subsets of) 2×2 stochastic matrices with support $S(\mu)$. Let S be the smallest closed (multiplicative) semigroup containing $S(\mu)$. Let (a, b) denote the matrix whose first column is (a, b). It is well known that the convolution sequence $\mu^{(n)}$ converges weakly if and only if μ is not the unit mass at the matrix $(0, 1)$. Let us assume that each point in $S(\mu)$ has a strictly positive column and is invertible. Let λ be the weak limit of $\mu^{(n)}$ as

$n \to \infty$. Then λ is a probability measure concentrated on matrices with rank one, and is continuous if and only if it is not degenerate (that is, with support at exactly one point), also this holds if and only if there are at least two points in the support of μ which are not collinear with the point $(1, 0)$. The proof of this fact, though implicit in the pages of [6], does not appear there explicitly. We present it here briefly.

Suppose λ is not degenerate. If it is not continuous, then $b \equiv \sup\{\lambda(\{P\})\colon P \in S(\lambda)\}$ is positive. After noticing that the set $\{P \in S(\lambda)\colon \lambda(\{P\}) > b/2\}$ is finite, it is clear that the supremum is attained, that is, for some $P \in S(\lambda)$, $\lambda(\{P\}) = b$, where this point P has rank one. For each $y \in S(\mu)$, y being invertible, the set $\{P\}y^{-1} \equiv \{Q \in S\colon Qy = P\}$ is a singleton. Now, $\lambda * \mu = \lambda$, and for each $y \in S(\mu)$, $\lambda(\{P\}) \geq \lambda(\{P\}y^{-1})$. Thus, $\int[\lambda(\{P\}) - \lambda(\{P\}y^{-1})]\mu(\mathrm{d}y) = 0$, and therefore, for μ-almost all y, $\lambda(\{P\}y^{-1}) = \lambda(\{P\}) = b > 0$. The map $y \to \lambda(\{P\}y^{-1})$ is upper semicontinuous, and therefore, for all $y \in S(\mu)$, $\lambda(\{P\}y^{-1}) = b > 0$. Similarly, since $\lambda * \mu^{(n)} = \lambda$, we can also show that for any $y \in S(\mu)$ and any positive integer n, $y^n \in S(\mu^{(n)})$, and therefore, $\lambda(\{P\}y^{-n}) = b > 0$.

Now notice that if $y \notin \{(1, 0), (0, 1)\}$, then $\{P\} = \{P\}y^{-1}$ if and only if y^n converges to P if and only if the points P, y and $(1, 0)$ are collinear. If for each point $y \in S(\mu)$, $\{P\} = \{P\}y^{-1}$, then S will be contained in the line joining the points P and $(1, 0)$, and consequently, $S(\lambda) = \{P\}$, which implies that λ is degenerate. Thus, there is at least one point $y \in S(\mu)$ such that $\{P\}$ is different from $\{P\}y^{-1}$, this means that the points P, y and $(1, 0)$ are not collinear. We also know that for each positive integer $m > 1$, the points y, y^m and $(1, 0)$ are collinear; this means that for any $m > 1$, the points y^m, P and $(1, 0)$ are also not collinear. Therefore, for any two positive integers m and n, $\{P\}y^{-m} \neq \{P\}y^{-n}$, but this will contradict the fact that λ is a probability measure.

The following theorem is Theorem 3.2 on page 25 in [6]. By an oversight, the necessary assumption that λ is not degenerate in order to be continuous was not included in Theorem 3.2, in [6]. We add it below. The proof there is, however, correct and complete with this added assumption.

Theorem 1.2 [6, 7]. *Suppose that the support $S(\mu)$ of μ is given by*

$$S(\mu) = \{(x_i, y_i)\colon x_i \neq y_i, 1 \leq i \leq n, 0 < x_i < 1, 0 < y_i < 1\},$$

where (x, y) represents the stochastic matrix $\left[\begin{smallmatrix} x & 1-x \\ y & 1-y \end{smallmatrix}\right]$. Let $\mu\{(x_i, y_i)\} = p_i > 0$, $\sum_{i=1}^n p_i = 1$ and $|x_i - y_i| = \frac{1}{a_i}$. If λ is not degenerate, then λ is continuous singular with respect to the Lebesgue measure m_0 on $[0, 1]$ if one of the following conditions holds:

(1) $\sum_{i=1}^n |x_i - y_i| < 1$.
(2) $\sum_{i=1}^n |x_i - y_i| = 1$ *and for some* i, $p_i \neq |x_i - y_i|$.
(3) $\sum_{i=1}^n |x_i - y_i| > 1$ *and* $\prod_{i=1}^n (p_i a_i)^{p_i} > 1$.

The main purpose of this note is to describe a method which can be used to generalize Theorem 1.2 above when the support of μ is infinite and $d \geq 2$ (as the method used in [6] does not seem to be generalizable to this case). Here, we will use the concept of packing dimension of a measure. Our proofs are based on adapting existing and known proofs in the context of attractors and iterated function systems (see [2, 3]). For completeness and easy reading, we include details and clarifications. We do not claim any originality for the ideas used here but the use of these ideas in the context of Theorem 1.2 is likely new. Let us also mention that the problem of determining when, the limit measure λ in the theorem above, is continuous singular or absolutely continuous was mentioned and discussed briefly in page 160 of [10].

2. Necessary results

2.1 *Iterated systems*

Let (S, d) be a compact metric space and let $\mathfrak{K}(S)$ be the class of all closed subsets of S. For A, B in $\mathfrak{K}(S)$ define $D(A, B) \equiv \inf\{s > 0: A \subseteq N_s(B) \text{ and } B \subseteq N_s(A)\}$ where $N_s(B) \equiv \{y \in S: \text{for some } x \in B, d(x, y) < s\}$. Then $(\mathfrak{K}(S), D)$ is a complete metric space.

Let $\{f_i: S \to S\}_{i\in\Lambda}$, with Λ a countably infinite set, be a family of functions such that $d(f_i(x), f_i(y)) \le \pi_i d(x, y)$ for all $x, y \in S$ and some $\pi_i > 0$ and let $\pi = \sup\{\pi_i: i \in \Lambda\}$. Then, as is well known for finite Λ, here also if $\pi < 1$ then there is a unique compact subset $A \subseteq S$ such that

$$A = \overline{\bigcup_{i\in\Lambda} f_i(A)}. \tag{1}$$

The idea of the proof is the same as in the case when Λ is finite (see [2, 3]). Define for each $A \in \mathfrak{K}(S)$ the function $F: \mathfrak{K}(S) \to \mathfrak{K}(S)$ such that $F(A) = \overline{\bigcup_{i\in\Lambda} f_i(A)}$. It is then easy to verify that for $A, B \in \mathfrak{K}(S)$, $D(F(A), F(B)) \le \pi D(A, B)$ and therefore (1) follows from the classical contraction mapping theorem and the completeness of the space $(\mathfrak{K}(S), D)$.

Along the same lines, let $\{p_i\}_{i\in\Lambda}$ be a discrete probability measure on Λ and let I be a Λ-valued random variable distributed according to $\{p_i\}_{i\in\Lambda}$. Assume that $\mathbb{E}[\pi_I] < 1$. Then there is also a unique invariant probability measure λ on the Borel subsets of S such that

$$\lambda(A) = \sum_{i\in\Lambda} p_i \lambda(f_i^{-1}(A)) \tag{2}$$

for all Borel subsets $A \subseteq S$. Briefly, the idea of the proof is the following: Since S is compact, there exists $\beta > 0$ such that for any $x, y \in S$, $d(x, y) < \beta$. Define the $\mathfrak{F}_\beta$ by

$$\mathfrak{F}_\beta \equiv \{h: S \to \mathbb{R}: |h(x) - h(y)| \le d(x, y) \text{ and } |h(x)| \le \beta \text{ for all } x, y \in S\}$$

and define in $\mathbf{P}(S)$ (the space of all the probability measures defined on the Borel subsets of S) the metric $\rho(\mu_1, \mu_2) \equiv \sup\{|\int h\mathrm{d}\mu_1 - \int h\mathrm{d}\mu_2|: h \in \mathfrak{F}_\beta\}$. Then $(\mathbf{P}(S), \rho)$ is a complete metric space. To establish (2), we define $\Phi: \mathbf{P}(S) \to \mathbf{P}(S)$ such that for any Borel set $A \subseteq S$, $\Phi(\mu)(A) = \sum_{i\in\Lambda} p_i \mu(f_i^{-1}(A))$. Now, taking a fixed point $z \in S$, it follows that for $h \in \mathfrak{F}_\beta$, $\mu_1, \mu_2 \in \mathbf{P}(S)$,

$$\begin{aligned}\left|\int h\mathrm{d}\Phi(\mu_1) - \int h\mathrm{d}\Phi(\mu_2)\right| &= \left|\sum_{i\in\Lambda} p_i\bigg(\int [h\circ f_i(x) - h\circ f_i(z)]\mathrm{d}\mu_1(x)\right.\\ &\qquad \left. - \int [h\circ f_i(x) - h\circ f_i(z)]\mathrm{d}\mu_2(x)\bigg)\right|\\ &\le \left|\sum_{i\in\Lambda} p_i\pi_i\left[\int h_i\mathrm{d}\mu_1 - \int h_i\mathrm{d}\mu_2\right]\right|,\\ &\qquad\qquad \text{where } h_i(x) = \frac{1}{\pi_i}[h\circ f_i(x) - h\circ f_i(z)]\\ &\le \mathbb{E}[\pi_I]\rho(\mu_1, \mu_2), \text{ since } h_i \in \mathfrak{F}_\beta.\end{aligned}$$

Now (2) follows by applying the contraction mapping theorem on Φ. Notice that, in particular, $\mathbb{E}[\pi_I] < 1$ if π defined earlier (just preceding (1)) is less than 1.

2.2 *The model map*

To proceed, we need a 'model' map like the one presented in [2, 3]. Let S, Λ, $\{f_i, \pi_i: i \in \Lambda\}$ and π be defined as before. Consider the product space $\Lambda^{(\omega)} = \prod_{i\in\Lambda} \Lambda_i$, where $\Lambda_i \equiv \Lambda$ for each i. The elements σ in $\Lambda^{(\omega)}$ can be considered as a point $(e_1, e_2, e_3, \dots)$, $e_i \in \Lambda_i$ or as an infinite word $e_1e_2e_3\dots$ where the letters are taken from the alphabet Λ. If $\sigma = e_1e_2\dots e_k\sigma_k$ where $e_1, e_2, \dots, e_k$ are the first k letters of σ and $\sigma_k \in \Lambda^{(\omega)}$, then we denote the finite word $e_1e_2\dots e_k$ by simply $\sigma_{\restriction k}$, and by $[\sigma_{\restriction k}]$ we denote the set $\{\tau \in \Lambda^{(\omega)}: \tau_{\restriction k} = \sigma_{\restriction k}\}$, a natural topology over $\Lambda^{(\omega)}$ is the one generated by the sets of the form $[\sigma_{\restriction k}]$, we will denote this topology by $\mathcal{T}$. Also, in $\Lambda^{(\omega)}$ we are considering the distance function $\rho(\sigma, \tau) \equiv \prod_{i=1}^{k} \pi_{e_i}$, where k is the largest number such that $\tau_{\restriction k} = \sigma_{\restriction k}$.

Given a probability distribution $\{p_i\}_{i\in\Lambda}$ for Λ, it induces the product probability measure $\mathbb{P}^{(\infty)}$ on $\Lambda^{(\omega)}$. In particular, it follows that for any Borel subset $B \subseteq \Lambda^{(\omega)}$, we have that

$$\mathbb{P}^{(\infty)}(B) = \sum_{i\in\Lambda} p_i \mathbb{P}^{(\infty)}\{\sigma \in \Lambda^{(\omega)} | i\sigma \in B\}. \tag{3}$$

We will need (3) after we construct a model map $h: \Lambda^{(\omega)} \to S$, fashioned after Edgar (page 211 of [2]) in Proposition 4. In what follows, let I be a Λ-valued random variable distributed according to $\{p_i\}_{i\in\Lambda}$.

Remark 2.1. Notice that if $\pi \equiv \sup_{i\in\Lambda} \pi_i < 1$, then $\rho(\sigma, \tau)$ is a metric over $\Lambda^{(\infty)}$ that generates the topology $\mathcal{T}$ and that agrees with $\mathbb{P}^{(\infty)}$, meaning that all Borel sets of $\Lambda^{(\omega)}$ under ρ are measurable. Under the less restrictive condition $\pi \le 1$, we should include also the assumption $-\infty \le \mathbb{E}[\log \pi_I] < 0$, that is trivially attained in the cases $\pi < 1$ or $\mathbb{E}[\pi_I] < 1$. Under this condition, we have that for $\mathbb{P}^{(\infty)}$-almost all $\sigma = e_1e_2\cdots \in \Lambda^{(\omega)}$, $\prod_{i=1}^{k} \pi_{e_i} \downarrow 0$ as $k \to \infty$ i.e. $\rho(\sigma, \sigma) = 0$. Redefining ρ over those σ that do not satisfy $\rho(\sigma, \sigma) = 0$, we obtain a metric on $\Lambda^{(\omega)}$ that agrees with $\mathbb{P}^{(\infty)}$ and such that for $\mathbb{P}^{(\infty)}$-almost all $\sigma \in \Lambda^{(\omega)}$, $\rho(\sigma, \tau) = \prod_{i=1}^{k} \pi_{e_i}$ for all $\tau \in \Lambda^{(\omega)}$ that coincide with σ in exactly the first k letters.

PROPOSITION 2.2

Let S be a compact metric space and for each $i \in \Lambda$, a countably infinite set, let f_i be a map from S to S such that $d(f_i(x), f_i(y)) \le \pi_i d(x, y)$ for all $x, y \in S$, where $\pi_i > 0$ and $-\infty \le \mathbb{E}[\log \pi_I] < 0$. Consider the distance function ρ on $\Lambda^{(\omega)}$ and the product probability measure $\mathbb{P}^{(\infty)}$ induced by $\{p_i\}_{i\in\Lambda}$. Then, there exists a unique measurable function $h: \Lambda^{(\omega)} \to S$ such that $h(i\sigma) = f_i(h(\sigma))$ for each $\sigma \in \Lambda^{(\omega)}$ and each $i \in \Lambda$. This h also satisfies the following conditions:

1. *The measure $\lambda \in \mathbf{P}(S)$ such that for each Borel set $A \subseteq S$, $\lambda(A) = \mathbb{P}^{(\infty)}(h^{-1}(A))$ is the unique invariant probability measure given in* (2).
2. $\overline{h(\Lambda^{(\omega)})} = \overline{\bigcup_{i\in\Lambda} f_i(\overline{h(\Lambda^{(\omega)})})}$.
3. *For $\mathbb{P}^{(\infty)}$-almost all $\sigma, \tau \in \Lambda^{(\omega)}$, $d(h(\sigma), h(\tau)) \le \beta\rho(\sigma, \tau)$ where* $\beta = \sup\{d(x, y): x \in S, y \in S\}$.

4. *If* $\pi = \sup\{\pi_i : i \in \Lambda\} < 1$, *then* h *is continuous and* $\overline{h(\Lambda^{(\omega)})}$ *coincides with the unique invariant set* A *given in* (1).

The proof follows the same lines as in Edgar [2]. For completeness, we include it.

Proof. Fix an arbitrary $a \in S$. We define recursively a sequence of continuous functions g_k, $1 \leq k < \infty$, $g_k \colon \Lambda^{(\omega)} \to S$ such that $g_0(\sigma) = a$ for each $\sigma \in \Lambda^{(\omega)}$ and $g_{k+1}(i\sigma) = f_i(g_k(\sigma))$ for each $\sigma \in \Lambda^{(\omega)}$, $i \in \Lambda$ and $k \geq 0$. Notice that g_0 is continuous on $\Lambda^{(\omega)}$. Now, let us assume that for $k \geq 0$, g_k is also continuous in $(\Lambda^{(\omega)}, \mathcal{T})$. Let $i\sigma \in \Lambda^{(\omega)}$. Given $\epsilon > 0$, let l be such that if $\tau \in [\sigma_{\restriction l}]$ then $d(g_k(\sigma), g_k(\tau)) < \frac{\epsilon}{\pi_i}$. Now, for any $i\tau \in [(i\sigma)_{\restriction k+1}]$,

$$d(g_{k+1}(i\sigma), g_{k+1}(i\tau)) = d(f_i(g_k(\sigma)), f_i(g_k(\tau))) \leq \pi_i d(g_k(\sigma), g_k(\tau)) < \epsilon,$$

proving that g_{k+1} is continuous in $(\Lambda^{(\omega)}, \mathcal{T})$. Similarly, it is possible to show that if $\beta \equiv \sup\{d(x, y) \colon x \in S, y \in S\}$, then for any $\sigma = e_1 e_2 \cdots \in \Lambda^{(\omega)}$ and $k \geq 1$, $d(g_{k+m}(\sigma), g_k(\sigma)) \leq \beta \prod_{i=1}^{k} \pi_{e_i}$, but the condition $\mathbb{E}[\log \pi_I] < 0$ implies (by the strong law of large numbers) that for $\mathbb{P}^{(\infty)}$-almost all $\sigma \in \Lambda^{(\omega)}$,

$$\limsup_{k\to\infty} \prod_{i=1}^{k} \pi_{e_i} = \limsup_{k\to\infty} \exp\left\{\sum_{i=1}^{k} \log \pi_{e_i}\right\} = 0, \tag{4}$$

proving that the sequence $\{g_k(\sigma)\}_{k\geq 1}$ is Cauchy for $\mathbb{P}^{(\infty)}$-almost all $\sigma \in \Lambda^{(\omega)}$. Therefore, there exists a measurable function $h(\sigma)$ such that for $\mathbb{P}^{(\infty)}$-almost all $\sigma \in \Lambda^{(\omega)}$, $\lim_{k\to\infty} g_k(\sigma) = h(\sigma)$. Notice that if $\pi \equiv \sup\{\pi_i | i \in \Lambda\} < 1$, then $d(g_{k+m}(\sigma), g_k(\sigma)) \leq \beta\pi^k$ implies that the convergence is uniform and therefore that $h(\sigma)$ is continuous.

Now, since $g_{k+1}(i\sigma) = f_i(g_k(\sigma))$ for each $\sigma \in \Lambda^{(\omega)}$, $i \in \Lambda$ and $k \geq 1$, it follows that for $\mathbb{P}^{(\infty)}$-almost all $\sigma \in \Lambda^{(\omega)}$, $h(i\sigma) = f_i(h(\sigma))$. Also, from (3) it follows that for any Borel subset $A \subseteq S$, we have

$$\begin{aligned}
\mathbb{P}^{(\infty)}(h^{-1}(A)) &= \sum_{i\in\Lambda} p_i \mathbb{P}^{(\infty)}\{\sigma \in \Lambda^{(\omega)} | i\sigma \in h^{-1}(A)\} \\
&= \sum_{i\in\Lambda} p_i \mathbb{P}^{(\infty)}\{\sigma \in \Lambda^{(\omega)} | f_i(h(\sigma)) \in A\} \\
&= \sum_{i\in\Lambda} p_i (\mathbb{P}^{(\infty)} \circ h^{-1})(f^{-1}(A)),
\end{aligned}$$

establishing property (1) in the proposition. It is now clear that $\overline{h(\Lambda^{(\omega)})} = \overline{\bigcup_{i\in\Lambda} f_i(\overline{h(\Lambda^{(\omega)})})}$, and in the case that $\pi < 1$, $\overline{h(\Lambda^{(\omega)})}$ coincides with the unique invariant set given in (1).

Finally, to prove property (3), let k be the greatest integer such that $\tau_{\restriction k} = \sigma_{\restriction k}$. If $\tau = e_1 e_2 \ldots e_k \tau_k$ and $\sigma = e_1 e_2 \ldots e_k \sigma_k$, then

$$\begin{aligned}
d(h(\tau), h(\sigma)) &= d(f_{e_1} \circ f_{e_2} \circ \cdots \circ f_{e_k} \circ h(\tau_k), f_{e_1} \circ f_{e_2} \circ \cdots \circ f_{e_k} \circ h(\sigma_k)) \\
&\leq \prod_{i=1}^{k} \pi_{e_i} d(h(\tau_k), h(\sigma_k)) \\
&\leq \beta\rho(\tau, \sigma).
\end{aligned}$$

□

2.3 *A bound for the upper packing dimension*

Let λ be a Borel probability measure on a metric space S. Then the essential supremum (with respect to λ) of $\bar{L}_\lambda(x)$, $x \in S$ defined by $\bar{L}_\lambda(x) \equiv \limsup_{\delta\to 0} \frac{\log \lambda(\overline{B_\delta(x)})}{\log \delta}$, where $B_\delta(x)$ is the open ball with center x and radius δ, is called the upper packing dimension of the measure λ. When $S = \mathbb{R}^d$ and λ-ess sup $\bar{L}_\lambda(x) < d$, then λ is singular with respect to the Lebesgue measure m on $\mathbb{R}^d$. To see this, suppose that $\lambda << m$ where m is the Lebesgue measure in $\mathbb{R}^d$ and let $f(x) = \frac{\mathrm{d}\lambda}{\mathrm{d}m}$. Define

$$\alpha_r(x) \equiv \frac{\int_{B_r(x)} f(y)\mathrm{d}y}{m(B_r(x))}.$$

Notice that for almost all $x \in \mathbb{R}^d$ (and therefore for λ-almost all $x \in \mathbb{R}^d$), $f(x) < +\infty$ and $\lim_{r\to 0} \alpha_r(x) = f(x)$. Now, let $S \equiv \{x \in \mathbb{R}^d : f(x) > 0\}$. For λ-almost all $x \in S$,

$$\begin{aligned}\bar{L}_\lambda(x) &= \limsup_{r\to 0} \frac{\log\left(\int_{\overline{B_r(x)}} f(y)\mathrm{d}y\right)}{\log r}\\ &= \limsup_{r\to 0} \frac{\log\left(\int_{B_r(x)} f(y)\mathrm{d}y\right)}{\log r}\\ &= \limsup_{r\to 0} \left(\frac{\log \alpha_r(x)}{\log r} + \frac{\log m(B_r(x))}{\log r}\right)\\ &= d.\end{aligned}$$

But $\lambda(S) = \int_S f(x)\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\mathrm{d}x = \lambda(\mathbb{R}^d)$, so that for λ-almost all $x \in \mathbb{R}^d$, $\bar{L}_\lambda(x) = d$ and therefore, λ-ess sup $\bar{L}_\lambda(x) = d$, which is a contradiction.

Now, let S, Λ, $\{f_i : i \in \Lambda\}$, $\{p_i : i \in \Lambda\}$, $\{\pi_i : i \in \Lambda\}$ be as in Proposition (2.2). The following proposition holds.

PROPOSITION 2.3

Let λ be the unique invariant probability measure satisfying (*see* (2)) *such that* $\lambda(A) = \sum_{i\in\Lambda} p_i \lambda(f_i^{-1}(A))$ *for all Borel subsets $A \subseteq S$, and let I be a Λ-valued random variable distributed according to $\{p_i : i \in \Lambda\}$. Suppose that*

1. $-\infty \le \mathbb{E}[\log \pi_I] < 0$.
2. $\mathbb{E}[|\log p_I|] < \infty$.

Then, λ-ess sup $\bar{L}_\lambda(x) \le \frac{\mathbb{E}[\log p_I]}{\mathbb{E}[\log \pi_I]} = \frac{\sum_{i\in\Lambda} p_i \log p_i}{\sum_{i\in\Lambda} p_i \log \pi_i}$.

Proof. Consider the product space $\Lambda^{(\omega)}$ together with the product measure $\mathbb{P}^{(\infty)}$ induced by $\{p_i : i \in \Lambda\}$. Using the strong law of large numbers we have that for $\mathbb{P}^{(\infty)}$-almost all $\sigma \in \Lambda^{(\omega)}$, $\sigma = e_1 e_2 e_3, \ldots,$

$$\lim_{k\to\infty} \frac{\sum_{i=1}^k \log p_{e_i}}{k} = \mathbb{E}[\log p_I] \tag{5}$$

and

$$\lim_{k\to\infty}\frac{\sum_{i=1}^{k}\log\pi_{e_i}}{k}=\mathbb{E}[\log\pi_I]. \tag{6}$$

Also, we have from (4) that the condition $\mathbb{E}[\log\pi_I]<0$ implies that for $\mathbb{P}^{(\infty)}$-almost all $\sigma\in\Lambda^{(\omega)}$, $\sigma=e_1e_2e_3,\ldots,$

$$\lim_{k\to\infty}\prod_{i=1}^{k}\pi_{e_i}=0. \tag{7}$$

Now, let $F\subseteq\Lambda^{(\omega)}$ with $\mathbb{P}^{(\infty)}(F)=1$ be such that (5),(6), (7) and the property (3) of Proposition 2.2 hold. Fix $x=h(\sigma)$ where $\sigma\in F$, $\sigma=e_1e_2e_3,\ldots.$ For each $\epsilon>0$, define $k_\epsilon\equiv\max\left\{k\colon \prod_{i=1}^{k}\pi_{e_i}<\frac{\epsilon}{\beta}\le\prod_{i=1}^{k-1}\pi_{e_i}\right\}$. Now, if $\sigma'\in[\sigma_{\restriction k_\epsilon}]\cap F$ we have by the property (3) of Proposition 2.2 that

$$d(h(\sigma'),x)\le\beta\rho(\sigma',\sigma)\le\beta\prod_{i=1}^{k_\epsilon}\pi_{e_i}<\epsilon$$

and therefore $h([\sigma_{\restriction k_\epsilon}]\cap F)\subseteq B_\epsilon(x)$, so that

$$\begin{aligned}\log\lambda(\overline{B_\epsilon(x)})&\ge\log\lambda(h([\sigma_{\restriction k_\epsilon}]\cap F))\\&=\log\mathbb{P}^{(\infty)}(h^{-1}(h([\sigma_{\restriction k_\epsilon}]\cap F)))\\&\ge\log\mathbb{P}^{(\infty)}([\sigma_{\restriction k_\epsilon}]\cap F)\\&=\sum_{i=1}^{k_\epsilon}\log p_{e_i}.\end{aligned}$$

Now,

$$\begin{aligned}\limsup_{\epsilon\to0}\frac{\log\lambda(\overline{B_\epsilon(x)})}{\log\epsilon}&\le\lim_{\epsilon\to0}\frac{\frac{1}{k_\epsilon}\sum_{i=1}^{k_\epsilon}\log p_{e_i}}{\frac{1}{k_\epsilon-1}\sum_{i=1}^{k_\epsilon-1}\log\pi_{e_i}}\left(\frac{k_\epsilon}{k_\epsilon-1}\right)\\&=\lim_{k\to\infty}\frac{\frac{1}{k}\sum_{i=1}^{k}\log p_{e_i}}{\frac{1}{k-1}\sum_{i=1}^{k-1}\log\pi_{e_i}}\left(\frac{k}{k-1}\right)\\&=\frac{\mathbb{E}[\log p_I]}{\mathbb{E}[\log\pi_I]}.\end{aligned}$$

This means that for all $x\in h(F)$, $\bar{L}_\lambda(x)\le\frac{\sum_{i\in\Lambda}p_i\log p_i}{\sum_{i\in\Lambda}p_i\log\pi_i}$, but $\lambda(h(F))=\mathbb{P}^{(\infty)}(h^{-1}(h(F)))\ge\mathbb{P}^{(\infty)}(F)=1$ implies that

$$\lambda\text{-}\operatorname{ess\,sup}\bar{L}_\lambda(x)\le\frac{\sum_{i\in\Lambda}p_i\log p_i}{\sum_{i\in\Lambda}p_i\log\pi_i}.$$

□

3. Limit distribution of i.i.d. stochastic matrices

Now we will apply the above results (specially Proposition 2.3) to determine conditions for singularity of the limit distribution for products of i.i.d. 2×2 stochastic matrices. More specifically, we wish to find an analogue of Theorem 1.2 when $S(\mu)$ is countably infinite. Let us consider the same assumptions as in Theorem 1.2, with the only exception that this time

$$S(\mu) = \{(x_i, y_i): x_i \neq y_i, 1 \leq i < \infty, 0 < x_i < 1, 0 < y_i < 1\}.$$

We take, as before, $\mu\{(x_i, y_i)\} = p_i, 0 < p_i < 1$ and $\sum_{i=1}^{\infty} p_i = 1$. Let $|x_i - y_i| \equiv \frac{1}{a_i}$. Then $\lambda = \lim_{n\to\infty} \mu^{(n)}$ satisfies the convolution equation $\lambda = \lambda * \mu$ and is the self-similar probability measure satisfying $\lambda = \sum_{i=1}^{\infty} p_i \lambda \circ f_i^{-1}$, where $f_i(x) = y_i + (x_i - y_i)x$.

In the present context, $f_i(h(\sigma)) = h(\sigma) \cdot (x_i, y_i)$, where $h(\sigma)$ represents a stochastic matrix where the first element in each row is $h(\sigma)$ and (x_i, y_i) represents a stochastic matrix where the first elements in the two rows are respectively x_i and y_i. With this observation, it is easy to verify that the last equation in §2.2 involving $\mathbb{P}^{(\infty)} \circ h^{-1}$ can be written as $(\mathbb{P}^{(\infty)} \circ h^{-1}) = (\mathbb{P}^{(\infty)} \circ h^{-1}) * \mu$, implying that the invariant measure λ, being the unique solution of the convolution equation $\lambda = \lambda * \mu$, must be the measure $(\mathbb{P}^{(\infty)} \circ h^{-1})$.

Note that as is known (see [6]), λ can be considered as a probability measure on the closed unit interval [0, 1], and its support K will then satisfy (as was shown earlier) the equality $K = \overline{\bigcup_{i=1}^{\infty} f_i(K)}$. Suppose now that m is the Lebesgue measure on [0, 1]. Let us assume one of the following three conditions (recall that $\pi_i = |x_i - y_i|$ for each i):

1. $\sum_{i=1}^{\infty} \pi_i < 1$.
2. $\sum_{i=1}^{\infty} \pi_i = 1$, but $\pi_i \neq p_i$ for at least one i.
3. $\prod_{i=1}^{\infty} \left(\frac{\pi_i}{p_i}\right)^{p_i} < 1$.

We also assume that $0 < -\sum_{i=1}^{\infty} p_i \log p_i < \infty$. Now, if any of the conditions (1) or (2) holds, then Jensen's inequality implies that $\sum_{i=1}^{\infty} p_i \log \frac{\pi_i}{p_i} \leq \log\left(\sum_{i=1}^{\infty} \pi_i\right) \leq 0$, with equality if and only if $\sum_{i=1}^{\infty} \pi_i = 1$ and $\pi_i = p_i$ for all $i \geq 1$. Therefore, $-\sum_{i=1}^{\infty} p_i \log p_i < -\sum_{i=1}^{\infty} p_i \log \pi_i$. So that in any case (1), (2) or (3), it holds that $\frac{\sum_{i=1}^{\infty} p_i \log p_i}{\sum_{i=1}^{\infty} p_i \log \pi_i} < 1$ and therefore, it follows from Proposition 2.3 that on [0, 1], λ-ess sup $\overline{L}_\lambda(x) < 1$.

It follows that λ is then singular with respect to m. This generalizes Theorem 1.2 to the case when $S(\mu)$ is countably infinite.

4. Connections with the Erdös problem [4, 5]

Let t be a number in (0, 1). Let $X_1, X_2, \ldots$ be i.i.d. random variables taking values 0 and 1 each with probability 1/2 and let $Y(t) \equiv (1-t)\sum_{n=0}^{\infty} t^n X_n$. This random variable is related to the random (Erdös) sum $Z(t) \equiv \sum_{n=0}^{\infty} t^n Z_n$, where the Z_n's are i.i.d. random variables taking the values -1 and $+1$ each with probability 1/2. It is easy to notice that $Z(t)$ is equal in distribution to $\frac{2Y(t)-1}{1-t}$. Now, following [1], when in the equation (2), Λ consists of only two points, $f_1(x) = tx$ and $f_2(x) = tx + (1-t)$, each with probability 1/2, the invariant measure turns out to be the distribution of the random variable $Y(t)$. More generally, if μ is the probability measure on two different stochastic matrices (a, b)

and (c, d), giving equal mass to each, where $a - b = c - d = t > 0$, then the weak limit of the convolution sequence $\mu^{(n)}$ is, up to a linear scaling, equal to the distribution of $Y(t)$. Therefore, the limit of $\mu^{(n)}$ is continuous singular if and only if the distribution of $Y(t)$ or that of the random Erdös sum is continuous singular.

The previous observation connects our problem to the Erdös problem on symmetric Bernoulli convolutions (see [4, 5, 9]). Using this connection and Erdös observation in [4, 5] that the distribution of his random sum is continuous singular when $t = (-1+\sqrt{5})/2$, it is easily seen that in Theorem 1.2, condition 3 is not necessary when $|x_1 - y_1| + |x_2 - y_2| > 1$. Finally, it should be mentioned that in the present context, the references [1] and [8] are relevant.

References

[1] Dubins L E and Freedman D A, Invariant probabilities for certain Markov processes, *Ann. Math. Stat.* **37** (1966) 837–847

[2] Edgar G A, Integral, probability, and fractal measures (New York: Springer-Verlag) (1998)

[3] Edgar G A, Measure, topology, and fractal geometry, Undergraduate Texts in Mathematics (New York: Springer-Verlag) (1990)

[4] Erdös P, On a family of symmetric Bernoulli convolutions, *Am. J. Math.* **61** (1939) 974–976

[5] Erdös P, On the smoothness properties of Bernoulli convolutions, *Am. J. Math.* **62** (1940) 180–186

[6] Mukherjea A, Topics in products of random matrices, Tata Institute of Fundamental Research Lectures on Mathematics 87 (Mumbai) (2000)

[7] Mukherjea A, Nakassis A and Ratti J, On the distribution of the limit of products of i.i.d. 2 by 2 random stochastic matrices, *J. Theoret. Probab.* **12** (1999) 571–583

[8] Mukherjea A and Tserpes N A, Measures on topological semigroups, Lecture Notes in Math. 547 (Springer) (1976)

[9] Peres Y and Solomyak B, Absolute continuity of Bernoulli convolutions, a simple proof, *Math. Res. Lett.* **3** (1996)

[10] Rosenblatt, Murray, Markov Processes: Structure and Asymptotic Behavior (Springer Verlag) (1971)