PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 4, Number 3/2003, pp.000 - 000

A SIMPLE PROOF OF A BASIC THEOREM ON ITERATED RANDOM FUNCTIONS

Marius IOSIFESCU*

"Gheorghe Mihoc-Caius Iacob" Institute of Mathematical Statistics and Applied Mathematics of the Romanian Academy
Casa Academiei Române, Calea 13 Septembrie nr.13, RO-050711 Bucharest 5, Romania
E-mail: miosifes@acad.ro

We present a transparent proof of existence of a stationary probability for a Markov chain constructed by random iterations of functions on a complete separable metric space. Our proof is to be compared with that given under equivalent assumptions by Diaconis and Freedman [1, pp. 58-63]. We just use contraction properties of the two linear operators naturally associated with the Markov chain considered.

1. PRELIMINARIES

In what follows we shall be using the notation introduced in the Appendix at the end of this paper. Let also $\mathbf{N}_+ = \{1,2,...\}$ and $\mathbf{N} = \{0,1,...\}$

Let W be a metric space with metric δ and Borel σ-algebra $\mathcal{B}_W$, $(X, \mathcal{X})$ an arbitrary measurable space, $u: W \times X \to W$ a $(\mathcal{B}_W \otimes \mathcal{X}, \mathcal{B}_W)$-measurable mapping, and p a probability measure on $\mathcal{X}$. Write $u_x(w) := u(w,x)$, $w \in W$, $x \in X$, and note that for any $x \in X$ we have a $\mathcal{B}_W$-measurable mapping $u_x: W \to W$. The pair

$$(p, (u_x)_{x \in X}) \tag{1}$$

is called an iterated function system (IFS), at least in the case where X is a finite set. Actually, (1) is a special random system with complete connections (cf. [7, pp.5 & 15]; see also [3]). With IFS (1) we associate the linear operator U defined by

$$Uf(w) = \int_X f(u_x(w))\, p(\mathrm{d}x), w \in W. \tag{2}$$

Clearly, U maps into itself the linear space of $\mathcal{B}_W$-measurable extended real-valued functions f defined on W such that $Uf^+(w)$, and $Uf^-(w)$, $w \in W$, are not both equal to $+\infty$. An important special case where U is well defined for possibly unbounded functions f is described below.

Define

$$\ell(x) = \ell(x;\ \delta) = \sup_{\substack{w' \neq w'' \\ w', w'' \in W}} \frac{\delta(u_x(w'), u_x(w''))}{\delta(w', w'')}, x \in X$$

If the metric space W is assumed to be separable, then it is easy to see that the mapping $x \to \ell(x)$ of X into $\overline{\mathbf{R}}$ is $(\mathcal{X}, \mathcal{B}_{\overline{\mathbf{R}}})$-measurable. Assume that

$$\ell := \int_X \ell(x) p(\mathrm{d}x) < 1. \tag{3_δ}$$

It is known (see, e.g., [4, p.201]) that (3_δ) implies

---
* Member of the Romanian Academy

$$\int_X \log\left(\ell(x)\right) p(\mathrm{d}x) < 0 . \qquad (3'_\delta)$$

Conversely, if $\ell_\beta := \int_X \ell^\beta(x)\, p(\mathrm{d}x) < \infty$ for some $\beta > 0$ and $(3'_\delta)$ holds, then there exists $\alpha > 0$ such that $\ell_\alpha < 1$. Assume also that for some $w_0 \in W$ we have

$$\int_X \delta\left(w_0, u_x(w_0)\right) p(\mathrm{d}x) < \infty. \qquad (4_\delta)$$

Under assumptions (3_δ) and (4_δ), the operator U takes $\mathrm{Lip}_1(W)$ into itself. For, (4_δ) holds for *any* $w \in W$ in place of w_0 as

$$\delta(w, u_x(w)) \le \delta(w, w_0) + \delta(w_0, u_x(w_0)) + \delta(u_x(w_0), u_x(w)) \le (\ell(x)+1)\,\delta(w, w_0) + \delta(w_0, u_x(w_0)),$$

which yields

$$\int_X \delta(w, u_x(w))\, p(\mathrm{d}x) \le 2\delta(w_0, w) + \int_X \delta(w_0, u_x(w_0))\, p(\mathrm{d}x) < \infty, w \in W .$$

Next, for any $f \in \mathrm{Lip}_1(W)$ we have

$$\left| f(u_x(w)) \right| \le \left| f(w) \right| + \delta(w, u_x(w)),\ x \in X,\ w \in W,$$

hence

$$\left| Uf(w) \right| \le \int_X \left| f(u_x(w)) \right| p(\mathrm{d}x) < \infty, w \in W,$$

while $s(Uf) \le 1$ is an immediate consequence of (3_δ).

Actually, when restricted to the linear space $B(W)$ of $\mathcal{B}_W$-measurable real-valued bounded functions defined on W, U is the transition operator of a W-valued Markov chain $(\zeta_n)_{n \in \mathbf{N}}$ on a probability space $(\Omega, \mathcal{K}, P_{w_0, p})$ defined by $\zeta_0 = w_0$ (arbitrarily given in W) and

$$\zeta_n = u_{\xi_n} \circ \cdots \circ u_{\xi_1}(w_0), n \in \mathbf{N}_+, \qquad (5)$$

where $(\xi_n)_{n \in \mathbf{N}_+}$ is an i.i.d. X-valued sequence with common distribution p. The transition function Q of $(\zeta_n)_{n \in \mathbf{N}}$ is defined by $Q(w, A) = U\chi_A(w) = p(A_w)$, $w \in W$, $A \in \mathcal{B}_W$, where $A_w := \left\{ x \in X \mid u_x(w) \in A \right\}$ and χ_A is the indicator function of A. Then

$$Uf(w) = \int_W f(w')\, Q(w, \mathrm{d}w'), w \in W,$$

for any $f \in B(W)$ and, more generally,

$$U^n f(w) = \int_W f(w')\, Q^n(w, \mathrm{d}w'), w \in W,$$

for any $n \in \mathbf{N}_+$ and $f \in B(W)$, where Q^n is the n-step transition function associated with Q.

We shall also consider the more general case where $w_0 \in W$ is chosen at random according to a given probability distribution . More precisely, on a probability space $(\Omega, \mathcal{K}, P_{\lambda, p})$ let w_0 be a W-valued random variable with probability distribution $\lambda \in \mathrm{pr}(\mathcal{B}_W)$, independent of the ξ_i, $i \in \mathbf{N}_+$, which always are i.i.d. with common distribution p. In this case $(\zeta_n)_{n \in \mathbf{N}}$ defined by (5) is still a W-valued Markov chain with initial distribution λ and transition function Q.

Let us finally note that U is a bounded linear operator of norm 1 on $B(W)$, which is a Banach space when endowed with the supremum norm

$$\| f \| = \sup_{w \in W} |f(w)|,\ f \in B(W).$$

Another operator, closely related to U, is defined on $\mathrm{pr}(\mathcal{B}_W)$ by

$$V\mu(A) = \int_W \mu(\mathrm{d}w)\, Q(w, A),\ A \in \mathcal{B}_W,$$

for any $\mu \in \mathrm{pr}(\mathcal{B}_W)$. Actually, this is a kind of adjoint of U on $B(W)$, to mean that

$$(\mu, Uf) = (V\mu, f),\ \mu \in \mathrm{pr}(\mathcal{B}_W),\ f \in B(W), \tag{6}$$

where (μ, f) is defined as the integral $\int_W f\,\mathrm{d}\mu$. It is easy to check that V can be also expressed by means of an integral over X. We namely have

$$V\mu(A) = \int_X p(\mathrm{d}x)\, \mu u_x^{-1}(A),\ A \in \mathcal{B}_W,$$

for any $\mu \in \mathrm{pr}(\mathcal{B}_W)$, where $\mu u_x^{-1}(A) := \mu\left(u_x^{-1}(A)\right), x \in X, A \in \mathcal{B}_W$. Note that $V^n(A) = \int_W \mu(\mathrm{d}w)\, Q^n(w, A)$, $A \in \mathcal{B}_W$, or, alternatively,

$$V^n\mu(A) = \int_X \cdots \int_X p(\mathrm{d}x_1)\cdots p(\mathrm{d}x_n)\, \mu\left(u_{x_1} \circ \cdots \circ u_{x_n}\right)^{-1}(A),\ A \in \mathcal{B}_W, \tag{7}$$

for any $n \in \mathbf{N}_+$ and $\mu \in \mathrm{pr}(\mathcal{B}_W)$.

The probabilistic meaning of V^n is that $V^n\lambda(A) = P_{\lambda, p}(\zeta_n \in A)$ for any $\lambda \in \mathrm{pr}(\mathcal{B}_W)$, $A \in \mathcal{B}_W$, and $n \in \mathbf{N}_+$. From (7) we also have that

$$V^n\lambda(A) = P_{\lambda, p}\left(u_{\xi_1} \circ \cdots \circ u_{\xi_n}(w_0) \in A\right)$$

for any $n \in \mathbf{N}_+$, $A \in \mathcal{B}_W$, and $\lambda \in \mathrm{pr}(\mathcal{B}_W)$, with $P_{\lambda, p}(w_0 \in A) = \lambda(A)$.

The result below is well-known in the case where $f \in B(W)$, cf. (6). Its proof does not differ from that working when $f \in B(W)$.

Proposition 1. *If* $\int_W Uf\,\mathrm{d}\mu$ *exists for some real-valued* $\mathcal{B}_W$ *-measurable function* f *and probability* $\mu \in \mathrm{pr}(\mathcal{B}_W)$, *then* $\int_W f\,\mathrm{d}(V\mu)$ *also exists and the two integrals are equal.*

We shall deal here with the asymptotic behaviour as $n \to \infty$ of the distribution of ζ_n under $P_{\lambda, p}$. We actually reprove Theorem 5.1 in Diaconis and Freedman [1], which we give a simple, fully transparent proof by only using contraction properties of the operators U and V. In Section 2 we present the impact of the assumptions made on the contraction properties just alluded to, while Section 3 contains the proof of the main result. The Appendix gathers well known definitions and properties of different metrics in $\mathrm{pr}(\mathcal{B}_W)$.

2. AUXILIARY RESULTS

The key result on which our approach is based is

Proposition 2. *Assume that* (3_δ) *and* (4_δ) *hold. Let* $\mu, \nu \in \mathrm{pr}(\mathcal{B}_W)$ *such that* $\rho_H(\mu, \nu) < \infty$. *Then*

$$\tilde{n}_H(V\grave{\imath}, V\acute{\imath}) \le \ell\, \tilde{n}_H(\grave{\imath}, \acute{\imath}).$$

Proof. Under our assumptions, the operator U takes $\mathrm{Lip}_1(W)$ into itself. By Proposition 1 we then have

$$\tilde{n}_H(V\grave{\imath},\ V\acute{\imath}) = \sup\left\{ \int_W f\,\mathrm{d}(V\grave{\imath}) - \int_W f\,\mathrm{d}(V\acute{\imath}) \mid f \in \mathrm{Lip}_1(W)\right\} = \sup\left\{ \int_W Uf\,\mathrm{d}\grave{\imath} - \int_W Uf\,\mathrm{d}\acute{\imath} \mid f \in \mathrm{Lip}_1(W)\right\}. \tag{8}$$

Consider the function $g = Uf/\ell$. Note that $g \in \mathrm{Lip}_1(W)$ since for any $w', w'' \in W$, $w' \ne w''$, we have

$$\frac{|g(w') - g(w'')|}{\ddot{a}(w', w'')} = \frac{1}{\ell}\left| \int_X \frac{f(u_X(w')) - f(u_X(w''))}{\ddot{a}(w', w'')} p(\mathrm{d}x)\right| \le \frac{1}{\ell}\int_X \frac{\ddot{a}(u_X(w'), u_X(w''))}{\ddot{a}(w', w'')} p(\mathrm{d}x) \le \frac{1}{\ell}\int_X \ell(x)\,p(\mathrm{d}x) = 1.$$

Then, by (8),

$$\tilde{n}_H(V\grave{\imath}, V\acute{\imath}) = \ell \sup\left\{ \int_W g\,\mathrm{d}\grave{\imath} - \int_W g\,\mathrm{d}\acute{\imath} \;\middle|\; g = \frac{Uf}{\ell},\ f \in \mathrm{Lip}_1(W)\right\}$$
$$\le \ell \sup\left\{ \int_W f\,\mathrm{d} - \int_W^{\grave{\imath}} f \mid f \in {}_1(W)\right\} = \ell\ {}_H\ \tilde{n}\ (\grave{\imath}$$

and the proof is complete.

Clearly, the Appendix and the result just proved imply

Corollary 3. *Under the assumptions in Proposition* 2 *we have*

$$\tilde{n}_L(V^n\grave{\imath}, V^n\acute{\imath}) \le \ell^n \tilde{n}_H(\grave{\imath}, \acute{\imath})$$

for any $n \in \mathbf{N}_+$.

3. THE PROOF

We can now prove the main result.

Theorem 4. *Let* $(W, \ddot{a})$ *be a complete separable metric space. Assume that* $(3_{\ddot{a}})$ *and* $(4_{\ddot{a}})$ *hold. Then the associated Markov chain* $(\ae_n)_{n\in\mathbf{N}}$ *has a unique stationary distribution* ð *and*

$$\tilde{n}_L\left(Q^n(w, \cdot), \eth\right) \le \frac{\ell^n}{1-\ell}\int_X \ddot{a}(w, u_x(w))\,p(\mathrm{d}x) \tag{9}$$

for any $n \in \mathbf{N}$ *and* $w \in W$. *On* $(\Omega, \mathcal{K}, P_{\pi,p})$ *the sequence* $(\ae_n)_{n\in\mathbf{N}}$ *is an ergodic strictly stationary process.*

Proof. **Step 1.** Let $\grave{\imath} \in \mathrm{pr}(\mathcal{B}_W)$ such that $\tilde{n}_H(\grave{\imath}, V\grave{\imath}) < \infty$. By Corollary 3, for any $m, n \in \mathbf{N}_+$ we can write

$$\tilde{n}_L\left(V^{n+m}\grave{\imath}, V^n\grave{\imath}\right) \le \sum_{k=0}^{m-1} \tilde{n}_L\left(V^{n+k}\grave{\imath}, V^{n+k+1}\grave{\imath}\right) \le \sum_{k=0}^{m-1} \ell^{n+k}\tilde{n}_H(\grave{\imath}, V\grave{\imath}) \le \frac{\ell^n}{1-\ell}\tilde{n}_H(\grave{\imath}, V\grave{\imath}). \tag{10}$$

Since $(W, \ddot{a})$ is complete, so is $\left(\mathrm{pr}(\mathcal{B}_W),\ \tilde{n}_L\right)$, see Appendix. Hence the sequence $\left(V^n\grave{\imath}\right)_{n\in\mathbf{N}}$ is convergent in $\left(\mathrm{pr}(\mathcal{B}_W),\ \tilde{n}_L\right)$ to some, say, $\eth \in \mathrm{pr}(\mathcal{B}_W)$.

Consider another $\acute{\imath} \in \mathrm{pr}(\mathcal{B}_W)$ such that $\tilde{n}_H(\grave{\imath}, \acute{\imath}) < \infty$. Then since

$$\tilde{n}_H(\acute{\imath}, V\acute{\imath}) \le \tilde{n}_H(\acute{\imath}, \grave{\imath}) + \tilde{n}_H(\grave{\imath}, V\grave{\imath}) + \tilde{n}_H(V\grave{\imath}, V\acute{\imath}) \le (\ell + 1)\,\tilde{n}_H(\grave{\imath}, \acute{\imath}) + \tilde{n}_H(\grave{\imath}, Vm),$$

we also have $ñ_H(í, Ví) < \infty$. This allows to conclude that $\left(V^n í\right)_{n\in \mathbf{N}}$ is convergent to the same $ð$ as for any $n \in \mathbf{N}_+$ we have

$$ñ_L\left(V^n í,\ ð\right) \le ñ_L\left(V^n ì,\ ð\right) + ñ_L\left(V^n ì,\ V^n í\right) \le ñ_L\left(V^n ì,\ ð\right) + \ell^n ñ_H(ì,\ í).$$

To sum up, we have proved that if $ì \in \mathrm{pr}\left(\mathcal{B}_W\right)$ satisfies the condition $ñ_H(ì, Vì) < \infty$, then there exists $ð = ð(ì)$ such that

$$ñ_L\left(V^n ì,\ ð\right) \le \frac{\ell^n}{1-\ell} ñ_H(ì, Vì),\ \ n \in \mathbf{N}_+. \tag{11}$$

[The last inequality follows at once from (10).] The same conclusion holds, with the same $ð$, for any other $í \in \mathrm{pr}\left(\mathcal{B}_W\right)$ for which $ñ_H(ì,\ í) < \infty$.

It is easy to prove that $ð = Vð$, that is, $ð$ is a stationary distribution for $(æ_n)_{n\in\mathbf{N}}$. We have $ñ_L(Vì,\ Vð) \le ñ_L(ì,\ í)$, $ì, í \in \mathrm{pr}\left(\mathcal{B}_W\right)$, by the very definition of the distance $ñ_L$ on account of Proposition 1. Then $ñ_L(V^{n+1} ì, Vð) \le ñ_L(V^n ì,\ ð) \to 0$ as $n \to \infty$. Hence both $Vð$ and $ð$ are equal to the limit in $\left(\mathrm{pr}\left(\mathcal{B}_W\right),\ ñ_L\right)$ of the sequence $(V^n ì)_{n\in\mathbf{N}}$, that is, $ð = Vð$.

Step 2. Clearly, d_w (probability measure concentrated at $w \in W$) satisfies $ñ_H(d_w,\ Vd_w) < \infty$ for any w since

$$ñ_H\left(d_w, Vd_w\right) = \sup\left\{ f(w) - \int_W f\,\mathrm{d}\left(Vd_w\right) \mid f \in \mathrm{Lip}_1(W)\right\} = \sup\left\{ f(w) - Uf(w) \mid f \in \mathrm{Lip}_1(W)\right\}$$

(by Proposition 1)

$$= \sup\left\{ \int_X \left(f(w) - f\left(u_x(w)\right)\right) p(\mathrm{d}x) \mid f \in \mathrm{Lip}_1(W, ä)\right\} \le \int_X ä\left(w, u_x(w)\right) p(\mathrm{d}x) < \infty$$

(by $(4_ä)$).

Note that since

$$ñ_H\left(d_{w'},\ d_{w''}\right) \le \sup\left\{ f(w') - f(w'') \mid f \in \mathrm{Lip}_1(W)\right\} \le ä\left(w', w''\right) < \infty$$

for any $w', w'' \in W$, it follows by Step 1 that the limiting $ð(d_w) := ð$ is the same for all $w \in W$.

Next, any finite linear combination $ī = \sum q_j d_{w_j}$ with positive rational coefficients such that $\sum q_j = 1$ satisfies the condition $ñ_H(ī, Vī) < \infty$ since, as it is easy to see,

$$ñ_H\left(ī, Vī\right) \le \sum q_j ñ_H\left(d_{w_j},\ Vd_{w_j}\right).$$

Moreover, $\left(\mathrm{pr}\left(\mathcal{B}_W\right),\ ñ_L\right)$ is separable since $(W, ä)$ was assumed to be, see Appendix, and it appears that the class of probability measures $ī = \sum q_j d_{w_j}$ just considered is dense in $\left(\mathrm{pr}\left(\mathcal{B}_W\right),\ ñ_L\right)$ if we start with a countable dense subset $\left\{ w_j \mid j \in \mathbf{N}_+\right\}$ in W. Cf. [5, p.83]. Let then $ë \in \mathrm{pr}\left(\mathcal{B}_W\right)$ be *arbitrary* and for any $å > 0$ consider a probability measure $ī_å$ from that class such that

$$ñ_L(ë,\ ī_å) < å.$$

We have $\lim_{n\to\infty} ñ_L(V^n ī_å, ð) = 0$ and since

$$ñ_L(V^n ë,\ ð) \le ñ_L(V^n ī_å,\ ð) + ñ_L(V^n ë,\ V^n ī_å) \le ñ_L(V^n ī_å, ð) + ñ_L(ë, ī_å),\ \ n \in \mathbf{N}_+,$$

it follows that

$$\limsup_{n\to\infty} \rho_L(V^n\lambda, \pi) \le \varepsilon .$$

As $\varepsilon > 0$ is arbitrary, we conclude that the sequence $(V^n\lambda)_{n\in\mathbf{N}}$ also converges to π in $\left(\mathrm{pr}(\mathcal{B}_W),\ \rho_L\right)$

Clearly, (9) follows from (11) with $\mu = d_w$, $w \in W$. For an arbitrary $\lambda \in \mathrm{pr}(\mathcal{B}_W)$ a similar conclusion holds if we assume that

$$\int_W \lambda(\mathrm{d}w) \int_X \delta\left(w, u_x(w)\right) p(\mathrm{d}x) < \infty .$$

Step 3. The uniqueness of π as stationary measure, $\pi = V\pi$, follows now easily. If $\pi' \in \mathrm{pr}(\mathcal{B}_W)$ satisfies $\pi' = V\pi'$, then by Step 2 we have

$$\lim_{n\to\infty} \rho_L(V^n\pi', \pi) = 0$$

and, at the same time, $V^n\pi' = \pi'$, $n \in \mathbf{N}_+$. Hence $\pi' = \pi$.

Next, the ergodicity of π, that is, $(\zeta_n)_{n\in\mathbf{N}_+}$ is an ergodic strictly stationary sequence on $(\Omega, \mathcal{K}, P_{\pi,p})$, follows from Theorem 3(iii) in [2].

Remark. Equation (7) shows that the backward process

$$\zeta_n(w_0) = u_{x_1} \circ \cdots \circ u_{x_n}(w_0),\ w_0 \in W,\ n \in \mathbf{N}_+,$$

converges in distribution under any $P_{\lambda,p}$ to π as $n \to \infty$, that is,

$$\lim_{n\to\infty} P_{\lambda,p}(\zeta_n(w_0) \in A) = \pi(A),\ \lambda \in \mathrm{pr}(\mathcal{B}_W),\ A \in \mathcal{B}_W .$$

One can show more, namely, that $(\zeta'_n)_{n\in\mathbf{N}}$ converges $P_{\lambda,p}$-a.s. at a geometric rate to a W-valued random variable ζ_∞ not depending of $w_0 \in W$ (hence nor of $\lambda \in \mathrm{pr}(\mathcal{B}_W)$) such that $P_{\lambda,p}(\zeta_\infty \in A) = \pi(A)$, $A \in \mathcal{B}_W$. See [2] and [1, pp.59-62].

Corollary 5. *Under the assumptions in Theorem* 4, *for any real-valued bounded non-constant Lipschitz function* f *on* W *we have*

$$\left| U^n f(w) - \int_W f \mathrm{d}\pi \right| \le \frac{\ell^n}{1-\ell} \int_X \delta\left(w, u_x(w)\right) p(\mathrm{d}x) \max\left(\operatorname{osc} f, s(f)\right),\ n \in \mathbf{N}_+,\ w \in W,$$

with $\operatorname{osc} f = \sup_{w\in W} f(w) - \inf_{w\in W} f(w)$.

For the *proof* it is enough to note that for

$$g := \frac{f - \inf_{w\in W} f(w)}{\max\left(\operatorname{osc} f, s(f)\right)} \in \mathrm{Lip}_1(W),$$

we have $0 \le g \le 1$, and to recall the definition of $\rho_L(V^n d_w, \pi)$.

A more general version of Theorem 4 is obtained using the fact that δ^α is still a metric in W for any $0 < \alpha \le 1$. [It is enough to note that if $a, b, c \ge 0$ and $c \le a + b$, then $c^\alpha \le (a+b)^\alpha \le a^\alpha + b^\alpha$.] Write then (see Appendix) $\rho_{\alpha,L}$ and $\mathrm{Lip}_1^\alpha(W)$ for the items associated with the metric space $(W, {}^\alpha\delta)$, which correspond for $\alpha = 1$ to ρ_L and $\mathrm{Lip}_1(W)$, respectively. (Remark that $\mathcal{B}_W$ is not altered when replacing δ by δ^α.) Clearly, $\ell(x;\ \delta^\alpha) = [\ell(x;\ \delta)]^\alpha \coloneqq \ell^\alpha(x)$, $x \in X$, and then the conditions corresponding to (3_δ) and (4_δ) are

$$\ell_{\alpha} := \int_X \ell^{\alpha}(x)p(\mathrm{d}x) < 1 \qquad (3_{\alpha})$$

and

$$\int_X \delta^{a}(w_0, u_x(w_0))\, p(\mathrm{d}x) < \infty \qquad (4_{\alpha})$$

for some $w_0 \in W$ –hence for all $w_0 \in W$ –, respectively.

We can now state

Theorem 4′. *Let* (W, δ) *be a complete separable metric space. Assume that* (3_{α}) *and* (4_{α}) *hold. Then the associated Markov chain* $(\xi_n)_{n\in\mathbf{N}}$ *has a unique stationary distribution* π *and*

$$\rho_L(Q^n(w,\cdot), \pi) \le \frac{\ell_{\alpha}^n}{1-\ell_{\alpha}} \int_X \delta^{\alpha}(w, u_x(w))\, p(\mathrm{d}x) \qquad (9)$$

for any $n \in \mathbf{N}_+$ *and* $w \in W$. *On* $(\Omega, \mathrm{K}, P_{\pi,p})$ *the sequence* $(\xi_n)_{n\in\mathbf{N}}$ *is an ergodic strictly stationary process.*

Proof. It follows from Theorem 4 that (9) holds with $\rho_{\alpha,L}$ in place of ρ_L. The validity of (9) will follow from the inequality $\rho_{a,L} \ge \rho_L$ for any $0<\alpha\le 1$. We shall in fact prove that

$$\{f \mid f \in \mathrm{Lip}_1(W),\ 0 \le f \le 1\} \subset \{f \mid f \in \mathrm{Lip}_1^{\alpha}(W),\ 0 \le f \le 1\} \qquad (12)$$

for any $0<\alpha\le 1$, which clearly implies $\rho_{\alpha,L} \ge \rho_L$.

To proceed note that if $f \in \mathrm{Lip}_1(W)\ \left(= \mathrm{Lip}_1^1(W)\right)$ and $0 \le f \le 1$, then for any $0<\alpha\le 1$ we can write

$$\sup_{w'\ne w''} \frac{|f(w')-f(w'')|}{\delta^{\alpha}(w',w'')} = \max\left(\sup_{\substack{w'\ne w''\\ \delta(w',w'')\le 1}} \frac{|f(w')-f(w'')|}{\delta^{\alpha}(w',w'')},\ \sup_{\substack{w',w''\\ \delta(w',w'')>1}} \frac{|f(w')-f(w'')|}{\delta^{\alpha}(w',w'')}\right)$$

$$\le \max\left(\sup_{\substack{w'\ne w''\\ \delta(w',w'')\le 1}} \frac{|f(w')-f(w'')|}{\delta(w',w'')},\ \text{some quantity not exceeding } 1\right) \le \max\left(s(f),1\right) \le 1.$$

(We used the inequality $x^{\alpha} > x$ which holds for $0<\alpha,\ x<1$.) Hence $f \in \mathrm{Lip}_1^{\alpha}(W)$, showing that (12) holds.

Remarks. 1. It is obvious that the assumptions in Theorem 4′ are weaker than those in Theorem 4, so that the latter is a real generalization of the former.

2. P.Diaconis and D.Freedman's assumptions in their Theorem 5.1 (see [1, pp.58-59]) are $(3'_{\alpha})$ in conjunction with a so-called "algebraic-tail" condition on ℓ and δ which amounts to the existence of positive constants a and b such that

$$p\left(\{x \mid \ell(x) > y\}\right) < ay^{-b},\ p\left(\{x \mid \delta(w_0, u_x(w_0)) > y\}\right) < ay^{-b} \qquad (13)$$

for $y > 0$ large enough and some $w_0 \in W$, hence for all $w_0 \in W$. We are going to prove that these assumptions are equivalent to ours in Theorem 4′.

First, on account of the equation

$$E\eta = \int_0^{\infty} P(\eta > y)\,\mathrm{d}y \qquad (14)$$

which holds for any non-negative random variable η, it is clear that (3_{α}) and (4_{α}) imply both $(3'_{\alpha})$ and, via Markov's inequality, (13). Second, if (13) holds, then for any $\alpha > 0$ we have

$$p\left(\left\{x \mid \ell^{\alpha}(x) > y\right\}\right) < ay^{-b/\alpha},\ p\left(\left\{x \mid \delta^{\alpha}(w_0, u_x(w_0)) > y\right\}\right) < ay^{-b/\alpha}$$

for $y > 0$ large enough. Choosing $\alpha < \min(b, 1)$, it follows from (14) that both ℓ_α and $\int_X \delta^{\alpha}(w_0, u_x(w_0))\, dx$ are finite. But $\ell_\alpha < \infty$ in conjunction with $(3'_\delta)$ implies the existence of $0 < \alpha' < \alpha$ such that $\ell_{\alpha'} < 1$. (Cf. our Section 1.) The proof is complete.

APPENDIX

Given a metric space W with metric δ and Borel σ-algebra $\mathcal{B}_W$, let us denote by $\mathrm{pr}(\mathcal{B}_W)$ the collection of all probability measures on $\mathcal{B}_W$. In $\mathrm{pr}(\mathcal{B}_W)$ a distance ρ_H is defined by

$$\rho_H(\mu, \nu) = \sup\left\{\int_W f\,d\mu - \int_W f\,d\nu \;\middle|\; f \in \mathrm{Lip}_1(W)\right\}$$

for any $\mu, \nu \in \mathrm{pr}(\mathcal{B}_W)$, where $\mathrm{Lip}_1(W) = \{f : W \to \mathbf{R} \mid s(f) \le 1\}$ with

$$s(f) = s(f, \delta) = \sup_{\substack{w' \ne w'' \\ w', w'' \in W}} \frac{|f(w') - f(w'')|}{\delta(w', w'')}$$

We speak of a 'distance' (cf.[8, p.9]) and not of a metric since it is possible that $\rho_H(\mu, \nu) = \infty$ for some $\mu, \nu \in \mathrm{pr}(\mathcal{B}_W)$ However, we have $\rho_H(\mu, \nu) < \infty$ when, for instance, both μ and ν have bounded supports. Cf.[6, p.732].

A genuine well-known metric in $\mathrm{pr}(\mathcal{B}_W)$ is the Lipschitz metric ρ_L which is defined by

$$\rho_L(\mu, \nu) = \sup\left\{\left|\int_W f\,d\mu - \int_W f\,d\nu\right| \;\middle|\; f \in \mathrm{Lip}_1(W),\ 0 \le f \le 1\right\}$$

for any $\mu, \nu \in \mathrm{pr}(\mathcal{B}_W)$. If (W, δ) is a separable (complete) metric space, then $(\mathrm{pr}(\mathcal{B}_W), \rho_L)$ is a separable (complete) metric space, too. Another usual metric in $\mathrm{pr}(\mathcal{B}_W)$ is the Prokhorov metric ρ_P which is defined by

$$\rho_P(\mu, \nu) = \inf\left\{\varepsilon > 0 \mid \mu(A) \le \varepsilon + \nu(A^{\varepsilon}),\ A \in \mathcal{B}_W\right\}$$

for any $\mu, \nu \in \mathrm{pr}(\mathcal{B}_W)$, where $A^{\varepsilon} = \left\{ w \mid \delta(w, A) := \inf_{a \in A} \delta(w, a) < \varepsilon \right\}$. We have

$$\frac{1}{2}\rho_L(\mu, \nu) \le \rho_P(\mu, \nu) \le \rho_L^{1/2}(\mu, \nu)$$

for any $\mu, \nu \in \mathrm{pr}(\mathcal{B}_W)$. Cf. [5, pp.81-82].

Clearly, $\rho_L(\mu, \nu) \le \rho_H(\mu, \nu)$ and $\rho_P(\mu, \nu) \le \rho_H^{1/2}(\mu, \nu)$ for any $\mu, \nu \in \mathrm{pr}(\mathcal{B}_W)$.

ACKNOWLEDGEMENT

The author gratefully acknowledges support from the Deutsche Forschungsgemeinschaft under Grant 936 RUM 113/21/0-1.

REFERENCES

1. DIACONIS, P., FREEDMAN, D., *Iterated random functions,* SIAM Review, **41**, pp.45-76, 1999.
2. ELTON, J.H., *A multiplicative ergodic theorem for Lipschitz maps,* Stochastic Process. Appl., **34**, pp. 39-47, 1990.
3. HERKENRATH, U., IOSIFESCU, M., RUDOLPH, A., *Random systems with complete connections and iterated function systems,* Math.Reports, **5**(**55**), pp.127-140,2003.
4. HEWITT,E., STROMBERG, K., *Real and Abstract Analysis,* New York, Springer-Verlag, 1965.
5. HOFFMANN-JØRGENSEN, J., *Probability with a View Toward Statistics,* Vol. II, New York, Chapman & Hall, 1994.
6. HUTCHINSON, J., *Fractals and self-similarity,* Indiana Univ. Math. J., **30**, pp.713-747, 1981.
7. IOSIFESCU, M., GRIGORESCU, S., *Dependence with Complete Connections and its Applications,* Cambridge, Cambridge Univ. Press, 1990.
8. RACHEV, S.T., *Probability Metrics and the Stability of Stochastic Models,* Chichester, Wiley, 1991.

Received July 25, 2003