PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 17, Fasc. 1 (1997), pp. 115–138

INVARIANT MEASURES AND MARKOV CHAINS WITH RANDOM TRANSITION PROBABILITIES

BY

GUOQI LU AND ARUNAVA MUKHERJEA (TAMPA, FLORIDA)

Abstract. In this paper, sufficient conditions for the existence of (σ-finite) invariant measures for a class of Markov chains with random transition probabilities are given. A special class of Markov chains with random transition probabilities is also studied here to show the relevance of attractors for certain iterated function systems to the invariant measures for these chains, and some of these results are illustrated with computer-generated pictures.

1. Introduction. One of the aims of this paper is to generalize the model of Markov chains with random transition probabilities, extensively studied by Cogburn in a series of papers (here we mention only [5]) and later by Orey [16], to a locally compact Hausdorff state space. We provide a sufficient condition for the existence of a non-trivial σ-finite and locally finite (i.e. finite on compact subsets) invariant measure for the above-mentioned (skew) chains. In Section 3, where we present this result, the method and other details are described.

In Section 4, we consider a special class of the Markov chains we studied in Section 3. These chains are induced by a number of affine maps on $\boldsymbol{R}^d$. The Markov random walks (with values in these affine maps) from these Markov systems give rise to what has been called "attractors" in the literature. Such attractors in the context of iterated function systems, in various situations, have been studied extensively in recent years by M. Barnsley and his co-workers in numerous papers. Also, recently in [13], in the i.i.d. context, such attractors were studied in connection with invariant measures. The main purpose of this section is to establish a connection between these attractors and the invariant measures for the class of Markov chains that we study here. Our results in Section 4 are then illustrated by appropriate computer-generated pictures of attractors.

In Section 2, we present a result necessary and sufficient for the existence of invariant probability measures for general Markov chains that have the Feller property (i.e. takes bounded continuous functions into bounded continuous functions) and also the property of taking continuous functions vanishing at infinity into continuous functions vanishing at infinity.

2. Invariant measure for general Markov chains. In this section, we present a general theorem on the existence of an invariant probability measure for a Markov chain (X_n) with locally compact second countable state space X.

Let $P(x, A)$, $x \in X$ and $A \subset X$, be a Markov transition probability function so that, for each Borel subset $A \subset X$, $P(\cdot, A)$ is a Borel measurable function of x and, for each $x \in X$, $P(x, \cdot)$ is a probability measure on the Borel subsets of X.

In the rest of this section, we make the following two assumptions:

(i) If A is a compact subset of X, then the function $x \to P(x, A)$ vanishes at infinity.

(ii) P has the Feller property, that is, for each bounded continuous function f on X, the function $Pf(x) = \int f(y) P(x, dy)$ is continuous.

Let us remark that in the context of random matrices we are often in situations where the conditions (i) and (ii) hold. For instance, if μ is a (Borel) probability measure on X, the set of $d \times d$ non-negative matrices (with usual topology) which have no zero rows or zero columns, then the Markov kernel $P(x, A)$, defined by

$$P(x, A) = \mu\{y \in X: \; y \cdot x \in A\},$$

satisfies both (i) and (ii).

THEOREM 2.1. *Suppose that conditions* (i) *and* (ii) *hold. Then the following are equivalent*:

(a) *There is an invariant probability measure λ for P, that is,*

$$\lambda(A) = \int P(x, A) \lambda(dx)$$

for every Borel subset $A \subset X$.

(b) *There is a compact subset $K \subset X$ such that* $\sup_x n^{-1} \sum_{k=1}^{n} P^k(x, K)$ *does not converge to* 0 *as $n \to \infty$.*

(c) $$\sup_{A:\ \text{compact}} \limsup_{n \to \infty} \sup_{x} n^{-1} \sum_{k=1}^{n} P^k(x, A) = 1.$$

For the equivalence of (b) *and* (c), *we do not need condition* (ii).

Remark. Note that once we have established that (a) and (b) are equivalent, (c) follows trivially from (a). However, in the absence of condition (ii), the equivalence of (b) and (c) does not seem to be trivial. Let us also mention that our proof will show that the quantity on the left in (c) with "lim sup" there replaced by "lim inf" is also either 0 or 1, in the presence of condition (i).

Proof of Theorem 2.1. (a) $\Rightarrow$ (b). This is easy since if λ is an invariant probability measure for P, then

$$\lambda(A) = \int P(x, A)\lambda(dx) = \int P^n(x, A)\lambda(dx) = \int \frac{1}{n}\sum_{k=1}^{n} P^k(x, A)\lambda(dx).$$

(b) $\Rightarrow$ (a). Note that (b) implies that there is a compact set K such that, for some subsequence (n_i), there exist elements $x_{n_i} \in X$ such that

$$\frac{1}{n_i}\sum_{k=1}^{n_i} P^k(x_{n_i}, K) > \delta > 0 \quad \text{for all } i \geqslant 1.$$

Define the probability measures (ν_{n_i}) by

$$\nu_{n_i}(\cdot) = \frac{1}{n_i}\sum_{k=1}^{n_i} P^k(x_{n_i}, \cdot).$$

Then there exists a subsequence (n_{i_j}) such that $(\nu_{n_{i_j}})$ converges (weak) to some measure ν, $\nu(X) \leqslant 1$. It follows that $\nu(K) \geqslant \delta$.

Now we observe that, for any Borel set A,

$$\left|\int P(y, A)\nu_{n_i}(dy) - \nu_{n_i}(A)\right|$$
$$= \left|\frac{1}{n_i}\sum_{k=1}^{n_i} P^{k+1}(x_{n_i}, A) - \frac{1}{n_i}\sum_{k=1}^{n_i} P^k(x_{n_i}, A)\right| \to 0 \quad \text{as } n_i \to \infty.$$

Now, writing $\lambda_{n_i}(A) = \int P(y, A)\nu_{n_i}(dy)$, we have: for any bounded continuous function f vanishing at infinity,

$$\int f d\lambda_{n_i} = \int \left[\int f(z) P(y, dz)\right] \nu_{n_i}(dy) \to \int \left[\int f(z) P(y, dz)\right] \nu(dy),$$

since the function $Pf(y)$ also vanishes at infinity. Thus, since

$$\left|\int f d\lambda_{n_i} - \int f d\nu_{n_i}\right| \to 0 \quad \text{as } i \to \infty,$$

we have

$$\int \left[\int f(z) P(y, dz)\right] \nu(dy) = \int f d\nu.$$

Now, given any compact set $B \subset X$, the set B being a G_δ-set, there exists f_n (continuous with compact support) such that $f_n \searrow I_B$ so that $\nu(B) = \int P(y, B)\nu(dy)$. Thus, (b) $\Rightarrow$ (a).

(b) $\Rightarrow$ (c). Suppose that $a > 0$, where

$$\sup_{A:\ \text{compact}} \limsup_{n\to\infty} \sup_{x} \frac{1}{n}\sum_{k=1}^{n} P^k(x, A) = a. \tag{2.1}$$

If possible, let $a < 1$. Choose $\varepsilon > 0$, $a < c < 1$, such that $\varepsilon < c/4$, $0 < \varepsilon < c - a$, $0 < a < c - \varepsilon < c < 1$, $c(1+c)/2 < a$. Let A be a given compact set and $x \in X$.

There exists a positive integer m_0 such that, for $m \geqslant m_0$,

$$\sup_y \frac{1}{m} \sum_{k=1}^{m} P^k(y, A) < c - \varepsilon. \tag{2.2}$$

Using the condition (i), we can now find a compact set E such that

$$P^s(y, A) < \varepsilon, \quad 1 \leqslant s \leqslant m_0, \tag{2.3}$$

whenever $y \notin E$. Now we have from (2.1): There exists r_0 such that, for $r \geqslant r_0$,

$$\frac{1}{m_0 r} \sum_{k=m_0+1}^{m_0 r + m_0} P^k(x, E) < c - \varepsilon \quad \text{and} \quad \frac{c(c+1)}{2} + \frac{3}{r_0} < a. \tag{2.4}$$

There exists a positive integer s_0, $1 \leqslant s_0 \leqslant m_0$, depending on r such that, for $r \geqslant r_0$,

$$\frac{1}{r} \sum_{k=1}^{r} P^{km_0 + s_0}(x, E) < c - \varepsilon. \tag{2.5}$$

To see this, just write

$$\frac{1}{m_0 r} \sum_{k=m_0+1}^{m_0 r + m_0} P^k(x, E) = \frac{1}{m_0 r} \{P^{m_0+1}(x, E) + P^{2m_0+1}(x, E) + \ldots + P^{rm_0+1}(x, E)$$
$$+ P^{m_0+2}(x, E) + P^{2m_0+2}(x, E) + \ldots + P^{rm_0+2}(x, E) + \ldots$$
$$+ P^{m_0+m_0}(x, E) + P^{2m_0+m_0}(x, E) + \ldots + P^{rm_0+m_0}(x, E)\}.$$

Now, let $n \geqslant m_0(r_0+1)$ and write $n = m_0(r+1)+t$, $0 \leqslant t < m_0$, where $r \geqslant r_0$. Let s_0 be such that (2.5) holds for this s_0.

We will now use (2.2), (2.3) and (2.5). Now we have: for A in (2.2),

$$\frac{1}{n} \sum_{k=1}^{n} P^k(x, A)$$
$$= \frac{1}{n} [P(x, A) + \ldots + P^{m_0+s_0}(x, A)$$
$$+ \{\int_E \sum_{k=1}^{m_0} P^k(y, A) P^{m_0+s_0}(x, dy) + \int_{E^c} \sum_{k=1}^{m_0} P^k(y, A) P^{m_0+s_0}(x, dy)\}$$
$$+ \{\int_E \sum_{k=1}^{m_0} P^k(y, A) P^{2m_0+s_0}(x, dy) + \int_{E^c} \sum_{k=1}^{m_0} P^k(y, A) P^{2m_0+s_0}(x, dy)\} + \ldots$$
$$+ \{\int_E \sum_{k=1}^{m_0} P^k(y, A) P^{(r-1)m_0+s_0}(x, dy) + \int_{E^c} \sum_{k=1}^{m_0} P^k(y, A) P^{(r-1)m_0+s_0}(x, dy)\}$$
$$+ \sum_{k=1+s_0}^{t+m_0} P^{m_0 r + k}(x, A)]$$

$$\leqslant \frac{1}{n}\left[3m_0 + m_0 \left\{ (c-\varepsilon) P^{m_0+s_0}(x, E) + \left(\frac{c}{2}-\varepsilon\right)(1-P^{m_0+s_0}(x, E)) \right\} \right.$$

$$\left. + m_0 \left\{ (c-\varepsilon) P^{2m_0+s_0}(x, E) + \left(\frac{c}{2}-\varepsilon\right)(1-P^{2m_0+s_0}(x, E)) \right\} + \dots \right]$$

$$\leqslant \frac{1}{n}\left[3m_0 + m_0 \left\{ \frac{c}{2} + \frac{c}{2} P^{m_0+s_0}(x, E) \right\} + m_0 \left\{ \frac{c}{2} + \frac{c}{2} P^{2m_0+s_0}(x, E) \right\} + \dots \right]$$

$$\leqslant \frac{1}{n}\left[3m_0 + m_0 r \frac{c}{2} + m_0 \frac{c}{2} \sum_{k=1}^{r} P^{km_0+s_0}(x, E) \right]$$

$$\leqslant \frac{3m_0}{m_0(r+1)+t} + \frac{m_0 r}{m_0(r+1)+t}\left[\frac{c}{2} + \frac{c}{2}(c-\varepsilon) \right]$$

$$< \frac{3}{r_0} + \frac{c(1+c)}{2} < a,$$

which is a contradiction. This proves that (b) $\Rightarrow$ (c).

Notice that (c) $\Rightarrow$ (b) trivially. The proof of the theorem is now complete. ∎

3. Invariant measures for Markov chains with random transition probabilities. In this section, we generalize the model of Markov chains with random transition probabilities initially and extensively investigated by Cogburn [5] and later by Orey [16] to a Hausdorff topological state space. Cogburn points out in [5] that "the existence of a σ-finite invariant measure $\nu \ll \mu$ ($\mu = \kappa \times \pi$ a given distribution) is a more difficult question" and leaves this problem open. Orey also mentions this problem (see Problem 1.3.1 on p. 916 in [16]) as one of his open problems. We present here a sufficient condition for the existence of a non-trivial, σ-finite, and locally finite invariant measure for Cogburn's Markov chains. Under appropriate conditions (mentioned at the end of this section), this invariant measure will be absolutely continuous with respect to a given β.

We exploit here the method given by Skorokhod in [18]. Also, in the final step of the proof of the theorem in this section, the formula for the invariant measure ν in terms of the invariant measure λ for P_g that we use here was given in [18], and thus perhaps could have been omitted at the expense of creating, however, a hard-to-follow proof. Skorokhod's concern in his paper was to establish the uniqueness of the invariant measure, whereas our concern here is simply to establish the existence of a non-trivial, σ-finite, and locally finite invariant measure. Our idea is simply to find an appropriate function g so that $\{\int P_g^n((x, \vec{\theta}), \cdot)\beta(d(x, \vec{\theta}))\}$ is a tight sequence for a given distribution β, thus yielding an invariant probability measure for P_g, and then finding an invariant measure for P through Skorokhod's formula. We feel that the condition for the existence of a σ-finite invariant measure that we present here is new. Finally, even though we say our approach is as given in [18], Skorokhod's approach is essentially the same as Foguel's in [7] and Skorokhod's formula (that we use in Step V) is essentially the same as Harris' (see [7], equation (3.10)).

We now present our theorem. A reader, familiar with [18], can skip easily Steps I, II, and V in the proof of Theorem 3.2.

Let $(X, \mathscr{A})$ be a locally compact Hausdorff second countable topological space and $(\Theta, \mathscr{B})$ be a complete metric separable space. We will write

$$\Omega = X \times \vec{\Theta}, \quad \mathscr{F} = \mathscr{A} \times \vec{\mathscr{B}},$$

where $\vec{\Theta} = \Theta^Z$ and $\vec{\mathscr{B}} = \prod_{i \in Z} \mathscr{B}_i$, $\mathscr{B}_i = \mathscr{B}$ with product topology in $\vec{\Theta}$ and Ω.

Let $\{P_\theta(\cdot, \cdot)\colon \theta \in \Theta\}$ be a given family of transition probabilities on $(X, \mathscr{A})$, and $Y_n = (X_n, \xi_n)$ be the Markov chain with initial distribution β and transition probability $P\big((x, \vec{\theta}), F\big)$ given by

$$P\big((x, \vec{\theta}), F\big) = P_{\theta_0}\big(x, (F)_{T\vec{\theta}}\big),$$

where $(F)_{T\vec{\theta}}$ is defined by

$$(F)_{T\vec{\theta}} = \{y \in X\colon (y, T\vec{\theta}) \in F\},$$

T being the left shift on $\vec{\Theta}$.

Notice that if X is discrete, then

$$P\big((x, \vec{\theta}), y \times B\big) = P_{\theta_0}(x, y) I_B(T\vec{\theta}),$$

and

$$P\big((x, \vec{\theta}), F\big) = \sum_{y \in X} P_{\theta_0}(x, y) I_{(F)_y}(T\vec{\theta}).$$

This is exactly the Markov chain investigated by Cogburn [5] and later by Orey [16].

PROPOSITION 3.1. (i) $P\{\vec{\xi}_n = T^n \vec{\theta} \mid Y_0 = (x, \vec{\theta})\} = 1$.
(ii) $P^n\big((x, \vec{\theta}), F\big) = P_{\theta_0} \dots P_{\theta_{n-1}}\big(x, (F)_{T^n\vec{\theta}}\big)$.

Proof. We have

$$P\{\vec{\xi}_1 = T\vec{\theta} \mid Y_0 = (x, \vec{\theta})\} = P\big((x, \vec{\theta}), X \times \{T\vec{\theta}\}\big) = P_{\theta_0}(x, X) = 1.$$

First, we use induction to prove (ii). For $n = 1$, by definition, we have

$$P\big((x, \vec{\theta}), F\big) = P_{\theta_0}\big(x, (F)_{T\vec{\theta}}\big).$$

Suppose that, for $k > 1$,

$$P^k\big((x, \vec{\theta}), F\big) = P_{\theta_0} \dots P_{\theta_{k-1}}\big(x, (F)_{T^k\vec{\theta}}\big);$$

then for $k+1$ we have

$$\begin{aligned} P^{k+1}\big((x, \vec{\theta})), F\big) &= \int P^k\big((y, \vec{\theta}'), F\big) P\big((x, \vec{\theta}), d(y, \vec{\theta}')\big) \\ &= \int P^k\big((y, T\vec{\theta}), F\big) P\big((x, \vec{\theta}), dy \times \{T\vec{\theta}\}\big) \\ &= \int P_{\theta_1} \dots P_{\theta_k}\big(y, (F)_{T^{k+1}\vec{\theta}}\big) P_{\theta_0}(x, dy) = P_{\theta_0} \dots P_{\theta_k}\big(x, (F)_{T^{k+1}\vec{\theta}}\big). \end{aligned}$$

Therefore (ii) is true. Now, we have

$$P\{\vec{\xi}_n = T^n\vec{\theta} \mid Y_0 = (x, \vec{\theta})\} = P\{Y_n \in X \times \{T^n\vec{\theta}\} \mid Y_0 = (x, \vec{\theta})\}$$
$$= P^n((x, \vec{\theta}), X \times \{T^n\vec{\theta}\}) = P_{\theta_0}\ldots P_{\theta_{n-1}}(x, X) = 1,$$

which is (i). ∎

Let $\beta = m \times \pi$, where m is a probability measure on X and π is a probability measure on $\vec{\Theta}$. Let $C_0(\Omega)$ be the class of all bounded real continuous functions on Ω that vanish at infinity and $g \in C_0(\Omega)$, $0 \leqslant g \leqslant 1$. In what follows, we define the function $P_g(\cdot, \cdot)$, as in [18], on $(\Omega, \mathscr{F})$ as follows:

$$(3.1)\quad P_g((x, \vec{\theta}), F) = E_{(x,\vec{\theta})}\Big[g(Y_1)I_F(Y_1) + \sum_{n=2}^{\infty}(1-g(Y_1))\ldots(1-g(Y_{n-1}))g(Y_n)I_F(Y_n)\Big].$$

Note that for any sequence a_i, $0 \leqslant a_i < 1$,

$$a_1 + (1-a_1)a_2 + (1-a_1)(1-a_2)a_3 + \ldots = 1 - \prod_{i=1}^{\infty}(1-a_i) = 1 \quad \text{iff} \quad \sum_{i=1}^{\infty} a_i = \infty.$$

Thus, it is clear that P_g is a transition probability on $(\Omega, \mathscr{F})$ if

$$(3.2)\qquad P_{(x,\vec{\theta})}\Big[\sum_{n=1}^{\infty} g(Y_n) = \infty\Big] = 1 \quad \text{for all } (x, \vec{\theta}) \in \Omega.$$

We can now state our results.

THEOREM 3.2. *Suppose that there exists a strictly positive function* $g \in C_0(\Omega)$, $0 \leqslant g \leqslant 1$, *such that*

(i) $P_{(x,\vec{\theta})}\big(\sum_{n=1}^{\infty} g(Y_n) = \infty\big) = 1$ *for all* $(x, \vec{\theta})$;

(ii) *for β almost $(x, \vec{\theta})$, there exists a Borel subset $B((x, \vec{\theta}))$ of Ω, containing $(x, \vec{\theta})$, such that*

$$P_{(x,\vec{\theta})}\big(Y_n \in B^c((x, \vec{\theta}))\ \text{i.o.}\big) = 0;$$

(iii) $\lim_{n\to\infty}\int\ldots\int\prod_{i=1}^{n}(1-g(y_i, T^i\vec{\theta}'))P_{\theta_0'}(x', dy_1)\ldots P_{\theta_{n-1}'}(y_{n-1}, dy_n) = 0$ *uniformly for $(x', \vec{\theta}')$ in each $B((x, \vec{\theta}))$*;

(iv) *for any bounded continuous function f on Ω,*

$$Pf(x, \vec{\theta}) = \int f(y, \vec{\theta}')P((x, \vec{\theta}), d(y, \vec{\theta}'))$$

is continuous.

Then there exists a non-trivial σ-finite and locally finite (that is, finite on compact subsets) invariant measure μ on $(\Omega, \mathscr{F})$ for the transition kernel P.

Remark. Let us remark that in [7] Foguel proved the above theorem for a Markov process (with non-random transition probabilities) under the Feller condition (our assumption (ii) above) and an assumption (assumption (2.1) in [7]) which implies our conditions (i), (ii) and (iii) above for some bounded continuous function with compact support with $B((x, \vec{\theta})) = \Omega$ for each $(x, \vec{\theta})$. [His function β in his Section 3 (see his limit statement (3.3)) happens to be this continuous function with compact support.] Note that our Theorem 3.2 under assumptions (i)–(iv) for some continuous g with compact support (when Θ is a singleton set) follows easily if we observe that the transition probability P_g, in this case, acts on a compact space and, as such, has an invariant probability measure, which, in turn, gives a σ-finite invariant measure for P.

Proof of Theorem 3.2. Notice that condition (i) implies that P_g is a transition probability on Ω. Now we separate the proof into several steps.

Step I. Let $(\bar{Y}_n)$ be a copy of the Markov Chain (Y_n) with the same transition probability P. Then we have

$$E_{(x,\vec{\theta})}\left[f(Y_1, \ldots, Y_n)\, E_{Y_n}\big(h(\bar{Y}_1, \ldots, \bar{Y}_m)\big)\right]$$
$$= E_{(x,\vec{\theta})}\left[f(Y_1, \ldots, Y_n)\, h(Y_{n+1}, \ldots, Y_{n+m})\right],$$

where f and h are bounded Borel measurable functions on Ω. To prove this notice that

$$E_{(x,\vec{\theta})}\left[f(Y_1, \ldots, Y_n)\, E_{Y_n}\big(h(\bar{Y}_1, \ldots, \bar{Y}_m)\big)\right]$$
$$= \int \ldots \int \left[f\big((y_1, T\vec{\theta}), \ldots, (y_n, T^n\vec{\theta})\big)\right]$$
$$\times E_{(y_n, T^n\vec{\theta})}\big(h(\bar{Y}_1, \ldots, \bar{Y}_m)\big)\, P_{\theta_0}(x, dy_1)\, P_{\theta_1}(y_1, dy_2) \ldots P_{\theta_{n-1}}(y_{n-1}, dy_n)$$
$$= \int \ldots \int f\big((y_1, T\vec{\theta}), \ldots, (y_n, T^n\vec{\theta})\big) \Big[\int \ldots \int h\big((y_{n+1}, T^{n+1}\vec{\theta}), \ldots, (y_{n+m}, T^{n+m}\vec{\theta})\big)$$
$$\times P_{\theta_n}(y_n, dy_{n+1}) \ldots P_{\theta_{n+m-1}}(y_{n+m-1}, dy_{n+m})\Big]\, P_{\theta_0}(x, dy_1)\, P_{\theta_1}(y_1, dy_2) \ldots$$
$$\ldots P_{\theta_{n-1}}(y_{n-1}, dy_n) = E_{(x,\vec{\theta})}\left[f(Y_1, \ldots, Y_n)\, h(Y_{n+1}, \ldots, Y_{n+m})\right].$$

This establishes Step I.

Step II. In this step, we show that for $m \geqslant 1$

$$P_g^m\big((x, \vec{\theta}), F\big) = E_{(x,\vec{\theta})} \sum_{n_1, \ldots, n_m = 1}^{\infty} \Big[\Big(\prod_{k=1}^{n_1 - 1} \big(1 - g(Y_k)\big)\Big) g(Y_{n-1}) \ldots$$
$$\ldots \Big(\prod_{k=1}^{n_m - 1} \big(1 - g(Y_{n_1 + \ldots + n_{m-1} + k})\big)\Big) g(Y_{n_1 + \ldots + n_m})\, I_F(Y_{n_1 + \ldots + n_m})\Big].$$

To prove this, let $(\bar{Y}_n)$ be a copy of the Markov chain (Y_n) so that $\bar{Y}_n$ has also P as its transition probability. Now we use induction to establish the expression for P_g^m. Using induction, then we have

$$P_g^{m+1}((x,\vec{\theta}),F) = P_g^m(P_g I_F)(x,\vec{\theta})$$

$$= E_{(x,\vec{\theta})} \sum_{n_1,\dots,n_m} \Big(\prod_{k=1}^{n_1-1}(1-g(Y_k))\Big)g(Y_{n_1})\dots$$

$$\dots\Big(\prod_{k=1}^{n_m-1}(1-g(Y_{n_1+\dots+n_{m-1}+k}))\Big)g(Y_{n_1+\dots+n_m})P_g I_F(Y_{n_1+\dots+n_m})$$

$$= E_{(x,\vec{\theta})} \sum_{n_1,\dots,n_m} \Big(\prod_{k=1}^{n_1-1}(1-g(Y_k))\Big)g(Y_{n_1})\dots\Big(\prod_{k=1}^{n_m-1}(1-g(Y_{n_1+\dots+n_m+k}))\Big)g(Y_{n_1+\dots+n_m})$$

$$\times\Big[E_{Y_{n_1+\dots+n_m}} \sum_{n_{m+1}=1}^{\infty}\Big(\prod_{k=1}^{n_{m+1}-1}(1-g(\bar{Y}_k))\Big)g(\bar{Y}_{n_{m+1}})I_F(\bar{Y}_{n_{m+1}})\Big]$$

which, by Step I, is equal to

$$E_{(x,\vec{\theta})} \sum_{n_1,\dots,n_{m+1}} \Big(\prod_{k=1}^{n_1-1}(1-g(Y_k))\Big)g(Y_{n_1})\dots$$

$$\dots\Big(\prod_{k=1}^{n_{m+1}-1}(1-g(Y_{n_1+\dots+n_m+k}))\Big)g(Y_{n_1+\dots+n_{m+1}})I_F(Y_{n_1+\dots+n_{m+1}}).$$

This establishes Step II.

Step III. In this step, we establish that the sequence of probability measures

$$\{\textstyle\int P_g^m((x,\vec{\theta}),\cdot)\beta(d(x,\vec{\theta})):\ m\geqslant 1\}$$

is tight, that is, for $\varepsilon>0$ there is a compact set $F\subset\Omega$ such that for all $m\geqslant 1$

$$\int P_g^m((x,\vec{\theta}),F)\beta(d(x,\vec{\theta}))>1-\varepsilon.$$

To prove this, notice that for a sequence (a_i), $0\leqslant a_i\leqslant 1$,

$$\prod_{i=1}^{n}(1-a_i)=\sum_{j=n}^{\infty}\sum_{i=1}^{j}(1-a_i)\cdot a_{j+1}.$$

Therefore, with probability one, by assumption (i),

$$g(Y_1)+\sum_{n=2}^{\infty}\Big(\prod_{k=1}^{n-1}(1-g(Y_k))g(Y_n)\Big)=1-\prod_{k=1}^{\infty}(1-g(Y_k))=1. \tag{3.3}$$

Let $\varepsilon>0$. For each positive integer N, define the set $\Omega_N(\varepsilon)$ by

$$\Omega_N(\varepsilon)=\Big\{(x,\vec{\theta}):\ P_{(x,\vec{\theta})}\big(Y_n\in B((x,\vec{\theta}))\text{ for all }n\geqslant N\big)>1-\varepsilon,$$

$$\text{and }\int\dots\int\prod_{i=1}^{N}(1-g(y_i,T^i\vec{\theta}'))P_{\theta_0'}(x',dy_1)\dots P_{\theta_{N-1}'}(y_{N-1},dy_N)<\frac{\varepsilon}{2}$$

$$\text{for every }(x',\vec{\theta}')\in B((x,\vec{\theta}))\Big\}.$$

It is clear that $\Omega_N(\varepsilon) \subset \Omega_{N+1}(\varepsilon)$ and that

$$\lim_{N\to\infty} \beta\big(\overline{\Omega_N(\varepsilon)}\big) = 1$$

by condition (ii), where $\overline{\Omega_N(\varepsilon)}$ is the closure of $\Omega_N(\varepsilon)$. This means that there exists N_0 such that

$$\beta\big(\overline{\Omega_{N_0}(\varepsilon)}\big) > 1-\varepsilon.$$

Now, by condition (iii), for $(x, \vec{\theta}) \in \Omega_N(\varepsilon)$, whenever $(x', \vec{\theta}') \in B\big((x, \vec{\theta}')\big)$, we have

$$E_{(x',\vec{\theta}')}\Big[\sum_{n=N+1}^{\infty} \Big(\prod_{k=1}^{n-1} (1-g(Y_k))\Big) g(Y_n)\Big] < \frac{\varepsilon}{2}.$$

Choose $\delta < \varepsilon/[2(N_0+1)]$; since $g \in C_0(\Omega)$, there is a compact subset $F \subset \Omega$ such that $g(x, \vec{\theta}) < \delta$ for $(x, \vec{\theta}) \notin F$. Note that for $(x, \vec{\theta}) \in \Omega_{N_0}(\varepsilon)$ and all $(x', \vec{\theta}') \in B\big((x, \vec{\theta})\big)$, we have

$$E_{(x',\vec{\theta}')}\Big[g(Y_1) I_{F^c}(Y_1) + \sum_{n=2}^{\infty} \Big(\prod_{k=1}^{n-1} (1-g(Y_k))\Big) g(Y_n) I_{F^c}(Y_n)\Big] < \varepsilon. \tag{3.4}$$

Let $m \geqslant N_0+1$. We have, for $(x, \vec{\theta}) \in \Omega_{N_0}(\varepsilon)$,

$$\begin{aligned}
P_g^m\big((x, \vec{\theta}) F^c\big) &= E_{(x,\vec{\theta})} \sum_{n_1,\ldots,n_{m-1}} \Big(\prod_{k=1}^{n_1-1} (1-g(Y_k)) g(Y_{n_1})\Big) \ldots \\
&\quad \ldots \Big(\prod_{k=1}^{n_{m-1}-1} (1-g(Y_{n_1+\ldots+n_{m-2}+k})) g(Y_{n_1+\ldots+n_{m-1}})\Big) \\
&\quad \times E_{Y_{n_1+\ldots+n_{m-1}}} \sum_{n_m} \prod_{k=1}^{n_m-1} (1-g(\bar{Y}_k)) g(\bar{Y}_{n_m}) I_{F^c}(\bar{Y}_{n_m}) \\
&= E_{(x,\vec{\theta})} \sum_{n_1,\ldots,n_{m-1}} \Big(\prod_{k=1}^{n_1-1} (1-g(Y_k)) g(Y_{n_1})\Big) \ldots \\
&\quad \ldots \Big(\prod_{k=1}^{n_{m-1}-1} (1-g(Y_{n_1+\ldots+n_{m-2}+k})) g(Y_{n_1+\ldots+n_{m-1}})\Big) \\
&\quad \times \big[I_{B((x,\vec{\theta}))}(Y_{n_1+\ldots+n_{m-1}}) + I_{B^c((x,\vec{\theta}))}(Y_{n_1+\ldots+n_{m-1}})\big] \\
&\quad \times E_{Y_{n_1+\ldots+n_{m-1}}} \sum_{n_m} \prod_{k=1}^{n_m-1} (1-g(\bar{Y}_k)) g(\bar{Y}_{n_m}) I_{F^c}(\bar{Y}_{n_m}) \\
&\leqslant \varepsilon + P_g^{m-1} I_{B^c((x,\vec{\theta}))}\big((x, \vec{\theta})\big) \quad \text{(by (3.4))}.
\end{aligned}$$

Let

$$D = \big\{\omega:\ Y_n(\omega) \in B^c\big((x, \vec{\theta})\big) \text{ for some } n \geqslant N_0\big\}.$$

By (3.3), we obtain

$$P_g^{m-1} I_{B^c((x,\vec{\theta}))}(x, \vec{\theta}) \leqslant E_{(x,\vec{\theta})}(I_D)$$

since $(x, \vec{\theta}) \in \Omega_{N_0}(\varepsilon)$. Therefore,

$$P_g^m((x, \vec{\theta}), F^c) \leqslant \varepsilon + P_{(x,\vec{\theta})}(D) \leqslant 2\varepsilon.$$

Now we show that, for $m \geqslant 1$ and any $(x, \vec{\theta}) \in \overline{\Omega_{N_0}(\varepsilon)}$,

$$P_g^m((x, \vec{\theta}), F) \geqslant 1 - 3\varepsilon.$$

To this end, we need the Feller property for the transition kernel P_g that we will establish in Step IV. Notice that the compact subset F above has the property that F is a G_δ-set and that, for any $(x, \vec{\theta}) \in \Omega_{N_0}(\varepsilon)$,

$$P_g^m((x, \vec{\theta}), F) \geqslant 1 - 2\varepsilon.$$

Let us suppose that there is some $(x_0, \vec{\theta}_0) \in \overline{\Omega_{N_0}(\varepsilon)} - \Omega_{N_0}(\varepsilon)$ such that, for a positive integer m,

$$P_g^m((x_0, \vec{\theta}_0), F) < 1 - 3\varepsilon.$$

This means that there is an open set V, $F \subset V$, such that

$$P_g^m((x_0, \vec{\theta}_0), V) < 1 - 3\varepsilon.$$

Let f be a continuous function on Ω such that $0 \leqslant f \leqslant 1$, $f = 1$ on F and 0 on V^c. Then we have

$$\int f(y, \vec{\theta}') P_g^m((x_0, \vec{\theta}_0), d(y, \vec{\theta}')) < 1 - 3\varepsilon.$$

Since the transition kernel P_g^m has also the Feller property, this means that there must exist some $(x, \vec{\theta}) \in \Omega_{N_0}(\varepsilon)$ such that

$$\int f(y, \vec{\theta}') P_g^m((x, \vec{\theta}), d(y, \vec{\theta}')) < 1 - 3\varepsilon,$$

which implies that

$$P_g^m((x, \vec{\theta}), F) < 1 - 3\varepsilon.$$

This is a contradiction.

The rest of the proof of Step III now follows easily.

Step IV. In this step, we show that there exists an invariant probability measure for P_g.

First, notice that for any bounded continuous function f on Ω, $f \geqslant 0$,

$$P_g f((x, \vec{\theta})) = E_{(x,\vec{\theta})} \left[\sum_{n=1}^{\infty} \prod_{k=1}^{n-1} (1 - g(Y_k)) g(Y_n) f(Y_n) \right],$$

which is an infinite sum of non-negative bounded continuous functions (because of condition (iv)) and, as such, is lower semicontinuous. Also, notice that if

f above is bounded by M, then

$$P_g[M-f](x, \vec{\theta}) = M - P_g f((x, \vec{\theta})),$$

which is also lower semicontinuous so that $P_g f$ is also upper semicontinuous. It follows that, for any bounded continuous f, $P_g f$ is also continuous. Now we define, for any probability measure λ on Ω, the probability measure $P_g \lambda$ by

$$P_g \lambda(B) = \int P_g((x, \vec{\theta}), B)\lambda(d(x, \vec{\theta})).$$

Notice that if $\lambda_n \to \lambda$ weakly in $P(\Omega)$, then for any bounded continuous function f on Ω we have

$$\int f d(P_g \lambda_n) = \int [\int f((y, \vec{\theta}')) P_g((x, \vec{\theta}), d(y, \vec{\theta}'))] \lambda_n(d(x, \vec{\theta}))$$
$$\to \int [\int f((y, \vec{\theta}')) P_g((x, \vec{\theta}), d(y, \vec{\theta}'))] \lambda(d(x, \vec{\theta})) = \int f d(P_g \lambda),$$

so that $P_g \lambda_n \to P_g \lambda$ weakly as $n \to \infty$.

Let us now write, for any Borel subset $F \subset \Omega$,

$$\nu_n(F) \equiv \frac{1}{n} \sum_{k=1}^{n} \int P_g^k((x, \vec{\theta}), F)\beta(d(x, \vec{\theta})).$$

Notice that

$$\int f d\nu_{n_k} = \frac{1}{n_k} \sum_{i=1}^{n_k} \iint f(y, \vec{\theta}') P_g^i((x, \vec{\theta}), d(y, \vec{\theta}'))\beta(d(x, \vec{\theta})) \to \int f d\nu_0;$$

also,

$$\int f d(P_g \nu_{n_k}) = \frac{1}{n_k} \sum_{i=1}^{n_k} \iint f(y, \vec{\theta}') P_g^{i+1}((x, \vec{\theta}), d(y, \vec{\theta}'))\beta(d(x, \vec{\theta})) \to \int f d(P_g \nu_0).$$

It is now clear that, for any bounded continuous function f on Ω,

$$\int f d\nu_0 = \int f d(P_g \nu_0).$$

Since Ω is a metric space, given any compact subset $F \subset \Omega$, F is also a G_δ-set so that there is a sequence of continuous functions uniformly bounded by 1, converging pointwise to I_F. This means that, for every compact set $F \in \Omega$,

$$\nu_0(F) = P_g \nu_0(F).$$

By regularity of the measures, it follows that $\nu_0 = P_g \nu_0$. This completes the proof of Step IV.

Step V. Let λ be the invariant probability measure for P_g. Define ν on $(\Omega, \mathscr{F})$ by

$$\nu(F) = \int E_{(x,\vec{\theta})} [I_F(Y_1) + \sum_{n=2}^{\infty} \prod_{k=1}^{n-1} (1-g(Y_k)) I_F(Y_n)] \lambda(d(x, \vec{\theta})).$$

The formula for this ν is actually a formula that probably was first considered by Harris (see [7]). We have taken it from [18]. Notice that Ω is σ-compact and

$$\nu(\Omega) \geqslant \int P_g((x, \vec{\theta}), \Omega)\lambda(d(x, \vec{\theta})) = 1;$$

also, if F is a compact subset of Ω, there is a positive number M such that

$$I_F(Y_n) \leqslant Mg(Y_n),$$

since g is strictly positive and continuous on Ω, so that

$$\nu(F) \leqslant M \int P_g\big((x, \vec{\theta}), \Omega\big) \lambda\big(d(x, \vec{\theta})\big) \leqslant M.$$

This means that ν is a σ-finite, non-trivial and locally finite measure on Ω.

Let us now establish that ν is invariant for P. (Our proof here is quite different from that given in Lemma 6 of [18].) Observe that

$$\begin{aligned}
\nu(F) &= \int E_{(x,\vec{\theta})}\, I_F(Y_1)\, d\lambda + \int E_{(x,\vec{\theta})}\Big(\sum_{n=2}^{\infty} \prod_{k=1}^{n-1} \big(1-g(Y_k)\big) I_F(Y_n)\Big)\, d\lambda \\
&= \int P\big((x, \vec{\theta}), F\big)\, d\lambda + \int E_{(x,\vec{\theta})}\Big[\sum_{n=2}^{\infty} \Big\{\prod_{k=1}^{n-2} \big(1-g(Y_k)\big) I_F(Y_n) \\
&\qquad - \prod_{k=1}^{n-2} \big(1-g(Y_k)\big) g(Y_{n-1}) I_F(Y_n)\Big\}\Big]\, d\lambda \\
&= \int P\big((x, \vec{\theta}), F\big)\, d\lambda + \int E_{(x,\vec{\theta})}\Big[\sum_{n=1}^{\infty} \prod_{k=1}^{n-1} \big(1-g(Y_k)\big) I_F(Y_{n+1})\Big]\, d\lambda \\
&\qquad - \int E_{(x,\vec{\theta})}\Big[\sum_{n=1}^{\infty} \prod_{k=1}^{n-1} \big(1-g(Y_k)\big) g(Y_n) I_F(Y_{n+1})\Big]\, d\lambda \\
&= \int P\big((x, \vec{\theta}), F\big)\, d\lambda + \int E_{(x,\vec{\theta})}\Big[\sum_{n=1}^{\infty} \prod_{k=1}^{n-1} \big(1-g(Y_k)\big) E_{Y_n}\big(I_F(\bar{Y}_1)\big)\Big]\, d\lambda \\
&\qquad - \int E_{(x,\vec{\theta})}\Big[\sum_{n=1}^{\infty} \prod_{k=1}^{n-1} \big(1-g(Y_k)\big) g(Y_n) E_{Y_n}\big(I_F(\bar{Y}_1)\big)\Big]\, d\lambda \\
&= \int P\big((x, \vec{\theta}), F\big)\, d\lambda + \int E_{(x,\vec{\theta})}\Big[\sum_{n=1}^{\infty} \prod_{k=1}^{n-1} \big(1-g(Y_k)\big) P(Y_n, F)\Big]\, d\lambda \\
&\qquad - \int E_{(x,\vec{\theta})}\Big[\sum_{n=1}^{\infty} \prod_{k=1}^{n-1} \big(1-g(Y_k)\big) g(Y_n) P(Y_n, F)\Big]\, d\lambda \\
&= \int P\big((x, \vec{\theta}), F\big)\, d\lambda + \int E_{(x,\vec{\theta})}\Big[\sum_{n=1}^{\infty} \prod_{k=1}^{n-1} \big(1-g(Y_k)\big) P(Y_n, F)\Big]\, d\lambda \\
&\qquad - \int P_g\big(P(\cdot, F)\big)(x, \vec{\theta})\, d\lambda \\
&= \int E_{(x,\vec{\theta})}\Big[\sum_{n=1}^{\infty} \prod_{k=1}^{n-1} \big(1-g(Y_k)\big) P(Y_n, F)\Big]\, d\lambda = (\nu P)(F)
\end{aligned}$$

(the first and the third expressions above cancel since λ is invariant for P_g).

This completes Step V and the proof of the theorem. ∎

Let us remark that we would often like to have the invariant measure ν (for P) to be absolutely continuous with respect to the initial distribution β for the Markov chain (Y_n). This, of course, will necessitate imposing some conditions on the transition kernels $P_\theta(x, \cdot)$, $\theta \in \Theta$. One easy condition is, of course, that

$$\beta(F) = 0 \Rightarrow P_\theta(x, F_{T\vec{\theta}}) = 0 \quad \text{for all } x \in X \text{ and } \theta \in \Theta,$$

where $F_{T\vec{\theta}} = \{y \in X \mid (y, T\vec{\theta}) \in F\}$. If this condition holds, it is obvious that the invariant measure ν obtained in Step V is absolutely continuous with respect to β. In the remainder of this section, we present two results relevant to the problem of the existence of a σ-finite invariant measure for P (in the context of Theorems A, C and D in [8], pp. 65–70). These results, even though likely known to experts, are recorded here (as we have not found them in the literature in the form we present them here).

Let us first note the following definition. A set $F \in \mathscr{F}$ is *closed* if $\beta(F) > 0$ and $I_F \leqslant PI_F$ β a.s., where β is a given measure (not necessarily finite) on $(\Omega, \mathscr{F})$ and

$$PI_F(x, \vec{\theta}) = P\big((x, \vec{\theta}), F\big).$$

The *conservative set* C is defined by

$$C = \{(x, \vec{\theta}) \mid \sum_{k=0}^{\infty} uP^k(x, \vec{\theta}) = \infty\},$$

where $u \in L_1(\beta)$, u strictly positive and $\langle uP, f\rangle = \langle u, Pf\rangle$ for $f \in L_\infty(\beta)$. The set C is independent of u, see [8]. In what follows, we assume, like in [8], p. 2, for each $(x, \vec{\theta}) \in \Omega$, $P\big((x, \vec{\theta}), \cdot\big) \ll \beta$.

PROPOSITION 3.3. *Suppose that $\Omega = C$. Then the following are equivalent:*
(1) *P is β-irreducible, that is, for any F with $\beta(F) > 0$ β-a.s.,*

$$P_{(x,\vec{\theta})}(Y_n \in F \text{ for some } n \geqslant 1) > 0.$$

(2) *$\mathscr{F}_i = \{F \in \mathscr{F} \mid I_F = PI_F \ \beta \text{ a.s.}\}$ is trivial.*
(3) *P is β-recurrent, that is, for any F with $\beta(F) > 0$ β-a.s.,*

$$P_{(x,\vec{\theta})}(Y_n \in F \text{ i.o.}) = 1.$$

(4) *Ω is minimal closed, that is, if $F \subset \Omega$ and $\beta(\Omega - F) > 0$, then F is not closed.*
(5) *If $Pf = f$, $f \in L_\infty(\beta)$, then f is a constant β-a.s.*
(6) *If $f \in L_\infty^+(\beta)$, $f \not\equiv 0$, then $\sum_{k=1}^{\infty} P^k f = \infty$ β-a.s.* ∎

We omit the proof which is not difficult. Note that when $\Omega = C$ and one of the conditions in Proposition 3.2 holds, then there exists a σ-finite P-invariant measure on Ω equivalent to β provided the condition of Theorem D (on p. 70 of [8]) holds. This then leads to the well-known fact that if P satisfies the Harris condition, then there exists a σ-finite P-invariant measure equivalent to β (see Theorem E on p. 73 in [8]). We close this section with the following result for the case when X is countable and $\beta = \kappa \times \pi$ (like in [5], κ being the counting measure on X), π being a given stationary probability measure on $(\vec{\Theta}, \vec{\mathscr{B}})$.

PROPOSITION 3.4. *P is β-irreducible iff for any $y \in X$ and $\vec{B} \in \vec{\mathscr{B}}$, $\pi(\vec{B}) > 0$,*

$$\sum_{k=1}^{\infty} P(\theta_0 \ldots \theta_{k-1};\, x, y) I_{\vec{B}}(T^k \vec{\theta}) > 0 \quad \textit{for } (\beta-) \textit{ almost all } (x, \vec{\theta}),$$

where $P(\theta_0 \ldots \theta_{k-1}; x, y)$ means the entry on the x-th row and the y-th column of the product of the k stochastic matrices $P_{\theta_0}, P_{\theta_1}, \ldots, P_{\theta_{k-1}}$. ∎

4. Markov random walks, attractors, and invariant measures. In this section, we consider a special class of Markov chains with random transition probabilities involving transformations from R^d into R^d of the form

$$T(x) = A(x) + x_0, \quad x \in R^d, \tag{4.1}$$

where x_0 is a fixed element in R^d and A is linear. The reason for including this section is to provide a concrete example of Cogburn chains and at the same time to show connections between the invariant measures, left and right attractors for the Markov random walks that we get from these chains. Similar results as presented here appeared earlier in [3]; but the discussion involved only a finite number of transformations and the proofs there are not valid for an arbitrary family of such transformations.

Let Θ be a compact metric space and $\{W_\theta\colon \theta \in \Theta\}$ be a family of transformations of the form (4.1) such that

$$W_\theta(x) = A_\theta(x) + B_\theta. \tag{4.2}$$

Then, if $\theta_i \in \Theta$, $0 \leqslant i \leqslant n$, for $x \in R^d$ we have

$$\begin{cases} W_{\theta_n} W_{\theta_{n-1}} \ldots W_{\theta_0}(x) = A_{\theta_n} A_{\theta_{n-1}} \ldots A_{\theta_0}(x) \\ \quad + A_{\theta_n} \ldots A_{\theta_1} B_{\theta_0} + A_{\theta_n} \ldots A_{\theta_2} B_{\theta_1} + \ldots + A_{\theta_n} B_{\theta_{n-1}} + B_{\theta_n}, \\ W_{\theta_0} W_{\theta_1} \ldots W_{\theta_n}(x) = A_{\theta_0} A_{\theta_1} \ldots A_{\theta_n}(x) \\ \quad + A_{\theta_0} A_{\theta_1} \ldots A_{\theta_{n-1}} B_{\theta_n} + \ldots + A_{\theta_0} B_{\theta_1} + B_{\theta_0}. \end{cases} \tag{4.3}$$

Let (ξ_n) be an indecomposable Markov chain with state space Θ and transition function P such that the (initial) distribution π_0 of ξ_0 is invariant with respect to P. Then the stationary process gotten by taking π_0 as the distribution of ξ_0 is ergodic (see [4], Theorem 7.16, p. 136).

Let $\vec{\Theta} = \Theta^{Z^+}$, $\vec{\mathscr{B}} = \prod_{i \in Z^+} \mathscr{B}_i$, $\mathscr{B}_i \equiv \mathscr{B}$, where $\mathscr{B}$ is the class of Borel subsets of Θ. Let π be the probability measure on $(\vec{\Theta}, \vec{\mathscr{B}})$ induced by the stationary ergodic process (ξ_n), so that it is stationary and ergodic with respect to the left shift T on $\vec{\Theta}$ given by

$$T\vec{\theta} = (\theta_1, \theta_2, \ldots), \quad \text{where } \vec{\theta} = (\theta_0, \theta_1, \theta_2, \ldots).$$

Let π^* be the unique probability measure on $(\vec{\Theta}, \vec{\mathscr{B}})$ such that, for any finite-dimensional rectangle of the form $A^{(n)} = \{\vec{\theta}\colon \theta_i \in A_i, 0 \leqslant i \leqslant n\}$,

$$\pi^*(A^{(n)}) = \pi\{\vec{\theta}\colon \theta_i \in A_{n-i}, 0 \leqslant i \leqslant n\}.$$

Let us assume that

(4.4) π^* is stationary and ergodic.

[Let us remark that if Θ is finite and P is a strongly ergodic stochastic matrix with a stationary distribution π_0 such that $\pi_0(\theta) > 0$ for each $\theta \in \Theta$, then π induced by π_0 and P, as well as the corresponding π^*, as defined above, is stationary and ergodic.]

Let us write $E \equiv R^d$ and $\Omega \equiv E \times \vec{\Theta}$. For $(x, \vec{\theta}) \in \Omega$, let

$$\Phi(x, \vec{\theta}) = \big(W_{\theta_0}(x), T\vec{\theta}\big). \tag{4.5}$$

Let $\mathscr{A}$ be the Borel subsets of E and $\mathscr{F} = \mathscr{A} \times \vec{\mathscr{B}}$, and for $F \in \mathscr{F}$ let

$$K\big((x, \vec{\theta}), F\big) = I_F\big(\Phi(x, \vec{\theta})\big).$$

Consider the Markov chain (Y_n) on Ω with transition function K so that $Y_n = \Phi^n(Y_0)$. Let us now make the following assumptions:

(i) For some $M_1, M_2 > 0$ and all $\theta \in \Theta$,

$$\|A_\theta\| \leqslant M_1, \quad \|B_\theta\| \leqslant M_2.$$

(ii) For $n \geqslant 0$, $x \in E$, and $r \geqslant 0$, the sets

$$\{\vec{\theta} \in \vec{\Theta}\colon \|W_{\theta_n} \ldots W_{\theta_0}(x)\| \leqslant r\} \quad \text{and} \quad \{\vec{\theta} \in \vec{\Theta}\colon \|W_{\theta_0} \ldots W_{\theta_n}(x)\| \leqslant r\} \tag{4.6}$$

are in $\vec{\mathscr{B}}$.

(iii) There are real numbers $b < a < 0$ such that

$$b < \int \log \|A_\theta\| \, \pi_0(d\theta) \leqslant a < 0.$$

Taking $f(\vec{\theta}) = \log \|A_\theta\|$ and using the ergodic theorem (and (4.6)), we infer that

(4.7) the limit $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \log \|A_{\theta_k}\|$ exists and is not greater than $a < 0$

for all $\vec{\theta} \in D \subset \vec{\Theta}$, where $\pi(D) = \pi^*(D) = 1$; moreover, for $\vec{\theta} \in D$, $x \in E$,

$$\sup\{\|W_{\theta_0} W_{\theta_1} \ldots W_{\theta_n}(x)\|\colon n \geqslant 0\} < \infty. \tag{4.8}$$

Thus, it follows from (4.3) and (4.8) that, for $\vec{\theta} \in D$ and $x \in E$,

(4.9) $Z(\vec{\theta}) \equiv \lim_{n\to\infty} W_{\theta_0} W_{\theta_1} \ldots W_{\theta_n}(x)$ exists,

and this limit is independent of x because of (4.3) and (4.7).

Now suppose that Y_0 has distribution β such that

$$\beta(E \times \vec{B}) = \pi(\vec{B}), \quad \vec{B} \subset \vec{\Theta}.$$

Given $\varepsilon > 0$, let A be a compact subset of E and $r > 0$ such that

$$\beta(A \times \vec{\Theta}) > 1 - \varepsilon,$$

and $\pi^* \{\vec{\theta} \in \vec{\Theta}: \|Z(\vec{\theta})\| \leqslant r\} > 1 - \varepsilon$.

Given $x_0 \in K$, there exists $N > 0$ such that, for $n \geqslant N$,

$$\pi^* \{\vec{\theta} \in \vec{\Theta}: \|W_{\theta_0} W_{\theta_1} \dots W_{\theta_n}(x_0)\| \leqslant 2r\} > 1 - \varepsilon.$$

It follows from the definition of π^* that

$$\pi \{\vec{\theta} \in \vec{\Theta}: \|W_{\theta_n} W_{\theta_{n-1}} \dots W_{\theta_0}(x_0)\| \leqslant 2r\} > 1 - \varepsilon.$$

For n sufficiently large we obtain

$$\beta\Phi^{-n}(\{y \in E: \|y\| \leqslant 3r\} \times \vec{\Theta}) > 1 - 3\varepsilon.$$

This means that the sequence β_n, given by

$$\beta_n = \frac{1}{n} \sum_{k=0}^{n-1} \beta\Phi^{-k},$$

is tight, and converges weakly to some probability measure $\beta_0 \in P(\Omega)$ such that $\beta_0 = \beta_0 \cdot \Phi^{-1}$ and

$$\beta_0(E \times \vec{B}) = \pi(\vec{B}), \quad \vec{B} \subset \vec{\Theta}. \tag{4.10}$$

Now notice that if f_0 is any continuous function on E with compact support and $f(x, \vec{\theta}) = f_0(x)$ for $(x, \vec{\theta}) \in \Omega$, then

$$\begin{aligned}
\int f d\beta_0(x, \vec{\theta}) &= \int f d\beta_0 \Phi^{-n}(x, \vec{\theta}) \\
&= \lim_{n \to \infty} \int f(W_{\theta_n} W_{\theta_{n-1}} \dots W_{\theta_0}(x_0)) d\beta_0(x, \vec{\theta}) \\
&= \lim_{n \to \infty} \int f(W_{\theta_n} W_{\theta_{n-1}} \dots W_{\theta_0}(x_0)) \pi(d\vec{\theta}) \\
&= \lim_{n \to \infty} \int f(W_{\theta_n} W_{\theta_{n-1}} \dots W_{\theta_0}(x_0)) \pi^*(d\vec{\theta}) \\
&= \int f(Z(\vec{\theta})) \pi^*(d\vec{\theta}).
\end{aligned}$$

It follows that, for $A \subset E$,

$$\beta_0(A \times \vec{\Theta}) = \pi^* \{\vec{\theta} \in \vec{\Theta}: Z(\vec{\theta}) \in A\}. \tag{4.11}$$

Now, if we give Y_0 the initial distribution β_0, then the process $Y_n = \Phi^n(Y_0)$ becomes stationary. Let f be any continuous function with compact support on E. Using the ergodic theorem on $(\Omega, \mathscr{F}, \beta_0)$ and Φ, we infer that the limit

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f\left(W_{\theta_k} W_{\theta_{k-1}} \dots W_{\theta_0}(x)\right) = g_f(x, \vec{\theta}) \tag{4.12}$$

exists for β_0-almost all $(x, \vec{\theta})$. It is easy to see that $g_f(x, \vec{\theta})$ is independent of x, and then since π is ergodic, it is a constant, which is easily seen to be

$$\int f\left(Z(\vec{\theta})\right) \pi^*(d\vec{\theta}).$$

Since we have assumed that π and π^* are both stationary and ergodic with respect to the left shift on $\vec{\Theta}$, it is known (see [4], Corollary 6.24, p. 116) that either $\pi = \pi^*$ or π is orthogonal to π^*, that is, there is a shift invariant subset $\vec{B} \subset \vec{\Theta}$ such that $\pi(\vec{B}) = 1$ and $\pi^*(\vec{B}) = 0$. In the case when Θ is finite and the transition matrix P is symmetric or has identical rows, it is clear that $\pi = \pi^*$.

Let us now define the sets A_r and $A_l(\vec{\theta})$ as follows:

1. $A_r \subset E$ and A_r is the support of the probability measure $\beta_0(\cdot \times \vec{\Theta})$, where β_0 is the unique Φ-invariant measure on $(\Omega, \mathscr{F})$, satisfying (4.10). We will call A_r the *right attractor.*

2. $A_l(\vec{\theta}) \subset E$ and it is given by $\{y \in E \mid$ given any open subset $N(y)$ containing y, for any x in E there are infinitely many n such that $W_{\theta_n} W_{\theta_{n-1}} \dots W_{\theta_0}(x) \in N(y)\}$. We will call $A_l(\vec{\theta})$ the *left attractor.*

Note that when $\pi = \pi^*$, it follows from (4.11) that (since E is second countable) there is a set $D_0 \subset \vec{\Theta}$ with $\pi(D_0) = 1$ such that, for each $\vec{\theta} \in D_0$, $A_r \subset A_l(\vec{\theta})$. Also, when $\pi = \pi^*$, if $y \notin A_r$, then there is an open set $N(y)$ containing y such that

$$\beta_0\left(N(y) \times \vec{\Theta}\right) = \beta_0 \Phi^{-n}\left(N(y) \times \vec{\Theta}\right) = 0 \quad \text{for every } n$$

so that, for any $x \in E$,

$$\pi\{\vec{\theta} \mid W_{\theta_n} W_{\theta_{n-1}} \dots W_{\theta_0}(x) \in N(y) \text{ i.o.}\} = 0,$$

and, consequently, $y \notin A_l(\vec{\theta})$ π-a.e. Thus, we have proved the following:

THEOREM 4.1. *If π and π^* are both stationary and ergodic on $\vec{\Theta}$ with respect to the left shift T, then under assumption* (4.6) *there is a unique Φ-invariant probability measure β_0 on Ω, whose $\vec{\Theta}$-marginal is π. Furthermore, when $\pi = \pi^*$, there is a set $D_0 \subset \vec{\Theta}$ with $\pi(D_0) = 1$ such that, for any $\vec{\theta} \in D_0$, $A_l(\vec{\theta}) = A_r$, A_r being the support of the E-marginal of β_0.*

Note that when Θ is finite, it is easy to find conditions for $\pi = \pi^*$, as we mentioned earlier. When Θ is not finite, it is not so clear when (4.4) holds or when $\pi = \pi^*$. Instead of (4.4), let us now assume that

(4.13) $\quad \pi(D^*) = 1$, where D^* is given by

$$D^* = \{\vec{\theta} \in \vec{\Theta}\colon \sup_{n \geqslant 0} \|W_{\theta_n} \ldots W_{\theta_0}(x)\| < \infty \text{ for } x \in E\}.$$

[Note that (4.13) holds if for each W_θ, $\theta \in \Theta$, the corresponding A_θ has its norm less than a number $p < 1$.]

We can now state and prove the following theorem:

THEOREM 4.2. *Assume* (4.6) *and* (4.13). *Suppose that every basic open set (that is, every open rectangle) in* $\vec{\Theta}$ *has* π*-measure positive. Suppose also that, for* $\theta \in \Theta$, *each entry in* W_θ *(as a matrix with respect to a fixed basis) is a real continuous function of* θ. *Then there exists* $D \subset \vec{\Theta}$ *such that* $\pi(D) = 1$ *and, for any* $\vec{\theta}$, $\vec{\theta}'$ *in* D, $A_l(\vec{\theta}) = A_l(\vec{\theta}')$; *moreover, the left attractor* $A_l(\vec{\theta})$, $\vec{\theta} \in D$, *is the support of the E-marginal of* β_0, *where* $\beta_0 \in P(\Omega)$, $\beta_0 = \beta_0 \Phi^{-1}$ *such that its* $\vec{\Theta}$*-marginal is* π.

Proof. We separate the proof into several steps.

Step I. Let $A_1, A_2, \ldots, A_n$ be open subsets of $\vec{\Theta}$ such that

$$\alpha \equiv \pi(A_1 \times A_2 \times \ldots \times A_n \times \Theta \times \Theta \times \ldots) > 0.$$

Then there exists $D_1 \subset \vec{\Theta}$, depending upon the sets $A_1, A_2, \ldots, A_n$, such that $\pi(D_1) = 1$ and, for $\vec{\theta} \in D_1$,

$$\theta_{m+i} \in A_i, \quad 1 \leqslant i \leqslant n,$$

for infinitely many m.

To see this, define the sets $E_m \subset \vec{\Theta}$ by

$$E_m = \{\vec{\theta} \in \vec{\Theta}\colon \theta_{m+i} \in A_i,\ 1 \leqslant i \leqslant n\},$$

and let

$$g(\vec{\theta}) = \limsup_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} I_{E_m}(\vec{\theta}).$$

Then we have

$$g(T\vec{\theta}) = g(\vec{\theta});$$

moreover,

$$\int g\,d\pi \geqslant \limsup_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} \pi(E_m) = \alpha > 0.$$

Since π is ergodic, g is a positive constant for π-almost all $\vec{\theta}$. This establishes Step I.

Step II. There exists $D \subset \vec{\Theta}$ with $\pi(D) = 1$ such that $D \subset D^*$, D^* as in (4.13), and for $\vec{\theta} \in D$, (4.8) holds, and also, given any basic open rectangle in $\vec{\Theta}$ of the form $A_1 \times A_2 \times \ldots \times A_n \times \Theta \times \Theta \times \ldots$,

$$\theta_{m+1} \in A_i, \quad 1 \leqslant i \leqslant n,$$

for infinitely many m.

The proof of this step follows from that of Step I if we observe that $\vec{\Theta}$ is a compact metric space and, as such, it is second countable.

Step III. We show in this step that for $\vec{\theta}$, $\vec{\theta}'$ in D, D as in Step II, $A_l(\vec{\theta}) = A_l(\vec{\theta}')$.

With no loss of generality, we can assume that, for all $\vec{\theta} \in D$,

$$\prod_{i=0}^{n} \|A_{\theta_i}\| \to 0 \quad \text{as } n \to \infty.$$

Let $h \in A_l(\vec{\theta})$. Then there is a subsequence (n_k) of positive integers such that, for $x \in E$,

$$W_{\theta_{n_k}} \ldots W_{\theta_0}(x) \to h \quad \text{as } k \to \infty.$$

Let $\vec{\theta}' \in D$ and $\delta > 0$. Fix an $x \in E$. Let p be so large that $\|x\| < p$ and

$$\sup\{W_{\theta'_n} \ldots W_{\theta'_0}(x)\colon\ n \geqslant 0\} \leqslant p.$$

Let U_δ be the open set $\{h' \in E\colon\ \|h' - h\| < \delta\}$. Choose k so large that

$$W_{\theta_{n_k}} \ldots W_{\theta_0}(x) \in U_\delta \quad \text{and} \quad \prod_{i=0}^{n_k} \|A_{\theta_i}\| < \delta/(2p).$$

Let this k be now fixed. Because of the continuity assumption in the proposition, there is a basic open set

$$V = B_0 \times \ldots \times B_{n_k} \times \Theta \times \Theta \times \ldots$$

such that for any $\vec{\theta}'' \in D$ such that $\theta''_i \in \mathscr{B}_i$, $0 \leqslant i \leqslant n_k$,

$$W_{\theta''_{n_k}} \ldots W_{\theta''_0}(x) \in U_\delta \quad \text{and} \quad \prod_{i=0}^{n_k} \|A_{\theta''_i}\| < \delta/(2p).$$

By Step I, $\theta'_{m+i} \in B_i$, $0 \leqslant i \leqslant n_k$, for infinitely many m. This means that there exists a subsequence (m_s) such that, for each $s \geqslant 1$,

$$W_{\theta'_{m_s+n_k}} W_{\theta'_{m_s+n_k-1}} \ldots W_{\theta'_{m_s}}(x) \in U_\delta \quad \text{and} \quad \prod_{i=m_s}^{m_s+n_k} \|A_{\theta'_i}\| < \delta/(2p).$$

Notice that, for all $s \geqslant 1$, if $y_s = W_{\theta'_{m_s-1}} \ldots W_{\theta'_0}(x)$, then $\|y_s\| < p$ and

$$\begin{aligned} &\|W_{\theta'_{m_s+n_k}} \ldots W_{\theta'_0}(x) - W_{\theta'_{m_s+n_k}} \ldots W_{\theta'_{m_s}}(x)\| \\ &\quad \leqslant \|A_{\theta'_{m_s+n_k}} \ldots A_{\theta'_{m_s}}(y_s)\| + \|A_{\theta'_{m_s+n_k}} \ldots A_{\theta'_{m_s}}(x)\| < \frac{\delta}{2p} \cdot p + \frac{\delta}{2p} \cdot p = \delta. \end{aligned}$$

This means that, for $s \geq 1$,

$$W_{\theta'_{m_s+n_k}} \dots W_{\theta'_0}(x) \in U_{2\delta},$$

so that $A_l(\vec{\theta}') \cap \bar{U}_{2\delta} \neq \emptyset$. Consequently, we have $h \in A_l(\vec{\theta}')$. Thus, for $\vec{\theta}, \vec{\theta}' \in D$, $A_l(\vec{\theta}) \subset A_l(\vec{\theta}')$.

Step IV. Let $\beta_0 \in P(\Omega)$ such that $\beta_0 \Phi^{-1} = \beta_0$ and $\beta_0(E \times \vec{\mathscr{B}}) = \pi(\vec{\mathscr{B}})$ for $\vec{\mathscr{B}} \subset \vec{\Theta}$. (Such a β_0 exists by the same arguments used earlier.) Let $\mu(A) = \beta_0(A \times \vec{\Theta})$, $A \subset E$. Then, for $\vec{\theta} \in D$, $A_l(\vec{\theta}) = \text{supp}(\mu)$.

To prove this let $y \notin A_l(\vec{\theta})$. Then there is an open ball N_y containing y such that $N_y \cap A_l(\vec{\theta}) = \emptyset$, and $\pi(C) = 0$, where for $x_0 \in E$

$$C = \{\vec{\theta}\colon\ W_{\theta_n} \dots W_{\theta_0}(x_0) \in N_y \ \text{ i.o.}\}.$$

Then $\mu(N_y) = \beta_0(N_y \times \vec{\Theta}) = \beta_0 \Phi^{-n}(N_y \times \vec{\Theta})$ for $n \geq 1$; also

$$E \times C = \{(x, \vec{\theta})\colon\ W_{\theta_n} \dots W_{\theta_0}(x_0) \in N_y \ \text{ i.o.}\}$$

$$\supset \{(x, \vec{\theta})\colon\ W_{\theta_n} \dots W_{\theta_0}(x) \in N'_y \ \text{ i.o.}\} = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \Phi^{-n}(N'_y \times \vec{\Theta}),$$

where N'_y is an open ball containing y, but with its diameter half that of N_y. This means that

$$\mu(N'_y) > 0 \Rightarrow \beta_0(E \times C) = \pi(C) > 0,$$

which is a contradiction. Thus, $y \notin \text{supp}(\mu)$. Thus, the support of μ is contained in $A_l(\vec{\theta})$ for $\vec{\theta} \in D$. The converse is also clear by similar arguments. ∎

We conclude this section with a few computer-generated pictures of left and right attractors (for a finite Θ).

Example 1. Here we take three maps W_1, W_2 and W_3 from R^2 into R^2, where $W_i(x) = A_i(x) + B_i$, $x \in R^2$, given by

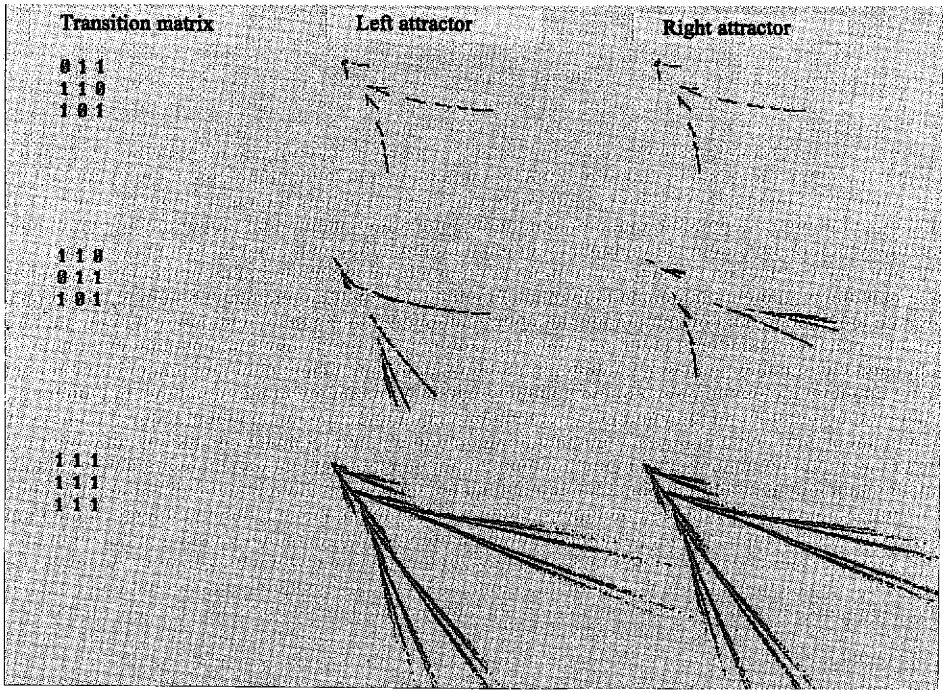
$$A_1 = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0.5 & 1 \\ 0 & 0.5 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0.5 & 0 \\ 0.5 & 0.5 \end{pmatrix};$$

$$B_1 = \begin{pmatrix} 50 \\ 80 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 50 \\ 200 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 170 \\ 200 \end{pmatrix}.$$

Here we take three different transition matrices P, one with identical rows (and symmetric) and the other two non-symmetric. In the picture, instead of giving P, we take the matrix Z, where $Z_{ij} = 0$ or 1 according as $P_{ij} = 0$ or > 0. Note that if

$$M_{ij} = Z_{ij} / \Big(\sum_{k=1}^{3} Z_{ik}\Big),$$

then the left attractors corresponding to P and M, but using the same maps W_1, W_2 and W_3, are the same, and the same is true for the right attractors.



[Let us remark that when Θ is finite, then if we define P^* by

$$P^*_{ij} = P_{ji}(\pi_0)_j/(\pi_0)_i,$$

where $\pi_0 P = \pi_0$ and $(\pi_0)_i > 0$ for all i, then $A_l(P) = A_r(P^*)$. Also, if $P_1 \approx P_2$ (that is, $(P_1)_{ij} > 0$ iff $(P_2)_{ij} > 0$), then $A_l(P_1) = A_l(P_2)$ and $A_r(P_1) = A_r(P_2)$.]

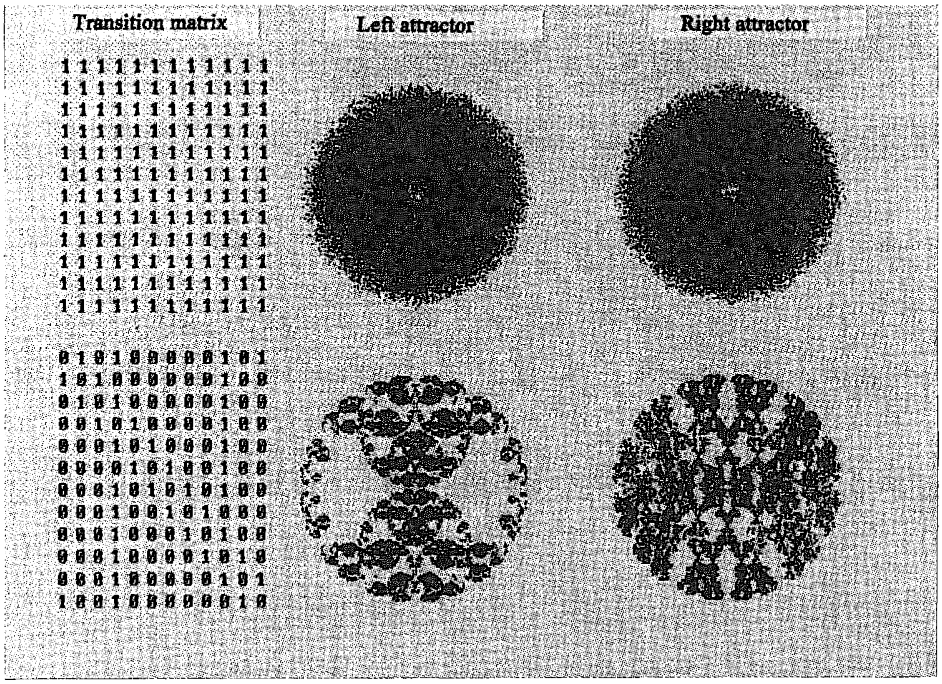
EXAMPLE 2. Here we take 12 maps W_i, $1 \leqslant i \leqslant 12$, such that the corresponding A_i's and B_i's are given by

$$A_i = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}, \quad 1 \leqslant i \leqslant 12;$$

$$B_1 = \begin{pmatrix} -100 \\ 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} -87 \\ 50 \end{pmatrix}, \quad B_3 = \begin{pmatrix} -50 \\ 87 \end{pmatrix}, \quad B_4 = \begin{pmatrix} 0 \\ 100 \end{pmatrix},$$

$$B_5 = \begin{pmatrix} 50 \\ 87 \end{pmatrix}, \quad B_6 = \begin{pmatrix} 87 \\ 50 \end{pmatrix}, \quad B_7 = \begin{pmatrix} 100 \\ 0 \end{pmatrix}, \quad B_8 = \begin{pmatrix} 87 \\ -50 \end{pmatrix},$$

$$B_9 = \begin{pmatrix} 50 \\ -87 \end{pmatrix}, \quad B_{10} = \begin{pmatrix} 0 \\ -100 \end{pmatrix}, \quad B_{11} = \begin{pmatrix} -50 \\ -87 \end{pmatrix}, \quad B_{12} = \begin{pmatrix} -87 \\ -50 \end{pmatrix}.$$



REFERENCES

[1] K. B. Athreya and P. Ney, *A new approach to the limit theory of recurrent Markov chains*, Trans. Amer. Math. Soc. 245 (1978), pp. 493–501.

[2] M. Barnsley, *Fractals Everywhere*, Academic Press, Inc. 1988.

[3] M. Barnsley et al., *Recurrent iterated function systems*, Constr. Approx. 5 (1989), pp. 3–31.

[4] L. Breiman, *Probability*, Addison-Wesley, Reading, Massachusetts, 1968.

[5] R. Cogburn, *The ergodic theory of Markov chains in random environments*, Z. Wahrsch. verw. Gebiete 66 (1984), pp. 109–128.

[6] K. J. Falconer, *Random fractals*, Math. Proc. Cambridge Philos. Soc. 100 (1986), pp. 559–582.

[7] S. R. Foguel, *Existence of a σ-finite invariant measure for a Markov process on a locally compact space*, Israel J. Math. 6 (1968), pp. 1–4.

[8] – *The Ergodic Theory of Markov Processes*, Van Nostrand, New York 1969.

[9] S. Graf, R. D. Mauldin and S. C. Williams, *The exact Hausdorff dimension in random recursive constructions*, Mem. Amer. Math. Soc. 381 (1988).

[10] T. Hattori and K. Hattori, *Self-avoiding process on the Sierpiński gasket*, Probab. Theory Related Fields 88 (1991), pp. 405–428.

[11] R. D. Mauldin and S. C. Williams, *On the Hausdorff dimension of some graphs*, Trans. Amer. Math. Soc. 298, No. 2 (1986), pp. 793–803.

[12] – *Hausdorff dimension in graph directed constructions*, ibidem 309, No. 2 (1988), pp. 811–829.

[13] A. Mukherjea, *Recurrent random walks in nonnegative matrices. II*, Probab. Theory Related Fields 96 (1993), pp. 415–434.

[14] M. Nakamura, *Invariant measures and entropies of random dynamical systems and the variational principle for random Bernoulli shifts*, Hiroshima Math. J. 21 (1991), pp. 187–216.

[15] E. Nummelin, *General irreducible Markov chains and non-negative operators*, Cambridge Univ. Press, London 1984.

[16] S. Orey, *Markov chains with stochastically stationary transition probabilities*, Ann. Probab. 19, No. 3 (1991), pp. 907–928.

[17] K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York 1967.

[18] A. V. Skorokhod, *Topologically recurrent Markov chains. Ergodic properties*, Theory Probab. Appl. 31, No. 4 (1986), pp. 563–571.

Department of Mathematics
University of South Florida
Tampa, Florida 33620-5700, U.S.A.

Received on 10.3.1996