The Annals of Probability
1990, Vol. 18, No. 2, 642–654

ON DIRECT CONVERGENCE AND PERIODICITY FOR TRANSITION PROBABILITIES OF MARKOV CHAINS IN RANDOM ENVIRONMENTS

BY ROBERT COGBURN

University of New Mexico

We study direct convergence of the products $P(\theta_0) \cdots P(\theta_n)$ of random stochastic matrices. These products can be interpreted as the transition probabilities of nonhomogeneous Markov chains selected at random by a stationary "environmental" sequence $\{\theta_n\}$, in other words, a Markov chain in a random environment. Rather than make assumptions analogous to irreducibility and aperiodicity for homogeneous Markov chains, we introduce equivalence relations that allow convergence results on the equivalence classes. The classical decomposition into a cycle of periodic sets is not possible in general, so the "periodicity" in the title is meant only to be suggestive. We also examine the frequency of times of positive probability of return to a state or set.

1. Introduction and preliminaries. Let $\{P(\theta), \theta \in \Theta\}$ be a family of stochastic matrices acting on a common, finite or denumerable space $\mathscr{X}$. Let $P(\theta; x, y)$ be the (x, y) entry of $P(\theta)$ and let $\mathscr{B}$ be a σ-field in Θ such that $P(\cdot; x, y)$ is $\mathscr{B}$ measurable for each $x, y \in \mathscr{X}$. Let $\vec{\Theta} = \Theta^{\mathbf{Z}}$ be the product space of doubly infinite sequences $\{\theta_n\}$ let $\vec{\mathscr{B}} = \mathscr{B}^{\mathbf{Z}}$ be its product σ-field and let π be a shift invariant probability on $(\vec{\Theta}, \vec{\mathscr{B}})$, so $\vec{\theta}$ is stationary under π. Now let $X_0, X_1, \ldots$ be a sequence in $\mathscr{X}$ such that

$$(1.1) \qquad P\Big(X_{n+1} = y | X_n = x,\, X_{n-1}, \ldots, X_0; \vec{\theta}\Big) = P(\theta_n : x, y) \quad \text{a.s.}$$

for all $x, y \in \mathscr{X}$ and $n \geq 0$. This two-level stochastic sequence is called a *Markov chain in a random environment.*

Given the sequence $\vec{\theta}$, the $\{X_n\}$ sequence evolves as a nonhomogeneous Markov chain and we will call these sequences the *$\vec{\theta}$-chains*. The transition probability from time m to $n > m$ for the $\vec{\theta}$-chain is $P(\theta_m) \cdots P(\theta_{n-1})$ and we will write $P(\theta_m \cdots \theta_{n-1})$ for this product. We will be studying the direct (i.e., unaveraged) convergence of these transition probabilities for the $\vec{\theta}$-chains, or, equivalently, of these products of random stochastic matrices.

To study this problem we need the space $\mathbf{S} = \mathscr{X} \times \vec{\Theta}$ with σ-field $\mathscr{F} = \mathscr{A} \times \vec{\mathscr{B}}$, where $\mathscr{A} = 2^x$, and measure $\mu = \kappa \times \pi$, where κ is counting measure on $\mathscr{X}$. Letting T denote the sequence shift on $\vec{\Theta}$, we define a transition probability P on S by

$$(1.2) \qquad P\Big((x, \vec{\theta}), \big(y, T\vec{\theta}\big)\Big) = P(\theta_0; x, y).$$

Received April 1988; revised March 1989.
AMS 1980 *subject classifications*. 60J10, 60J99.
Key words and phrases. Markov chains in random environments, products of random stochastic matrices, direct convergence, Kingman's subadditive ergodic theorem.

This formulation allows the L_1 approach to Markov processes of Hopf (see Foguel [3] for a general introduction) to be applied to our problem. For a given $\vec{\theta}$ sequence, the first coordinate of the $\{(X_n, T^n\vec{\theta})\}$ sequence evolves as the $\vec{\theta}$-chain described above. We also need the space $\hat{\mathbf{S}} = \mathscr{X} \times \mathscr{X} \times \vec{\Theta}$ with σ-field $\hat{\mathscr{F}} = \mathscr{A} \times \mathscr{A} \times \vec{\mathscr{B}}$, measure $\hat{\mu} = \kappa \times \kappa \times \pi$ and transition probability

$$\hat{P}\Big(\big(x,y,\vec{\theta}\big),\big(x',y',T\vec{\theta}\big)\Big) = P(\theta_0;x,x')P(\theta_0;y,y'). \tag{1.3}$$

Details of this application are developed in [2]. Let $\mathscr{F}_i$ denote the *σ-field of invariant sets* in S: $F \in \mathscr{F}_i$ if $P(I_F) = I_F$ a.e. Of course, F is *closed* if $P(I_F) \geq I_F$ a.e., and similar definitions hold on $\hat{S}$. In many cases in what follows, we will restrict the process on S or $\hat{S}$ to a closed subset in the usual way. Note that in the L_1 theory, relations are generally taken to apply up to null sets and these null sets are often ignored. In the present model, if a relation holds on S up to a null set N then, letting $(N)_x$ denote the section of N at x [we also use $(N)^{\vec{\theta}}$ for the section of N at $\vec{\theta}$], let $\overline{N} = \bigcup_{x \in \mathscr{X}}(N)_x$. Then $\overline{N}$ is a π-null set in $\vec{\Theta}$, and the set $\mathscr{X} \times \overline{N}^c$ is a closed set in S on which the relation holds pointwise, while $\mu(\mathscr{X} \times \overline{N}) = 0$. Thus, as long as we consider at most a countable number of relations, we can argue as though they hold everywhere.

Let $P_{x,\vec{\theta}}$ ($P_{\nu,\vec{\theta}}$) denote the distribution on the $\vec{\theta}$-chain $\{X_n\}$ when $X_0 = x$ (X_0 has distribution ν) and $\vec{\theta}$ is the environmental sequence. Similarly, $P_{x,y,\vec{\theta}}$ is the distribution on the $\vec{\theta}$-chain $\{(X_n, Y_n)\}$ when $X_0 = x$, $Y_0 = y$. Also P_φ denotes the distribution on the Markovian sequence $\{(X_n, T^n\vec{\theta})\}$ in $\mathbf{S}$ when $(X_0, \vec{\theta})$ has initial density φ. Expectations on these spaces are denoted by E with the corresponding subscript.

The conservative set $\mathbf{C}$ in $\mathbf{S}$ can be characterized as follows:

$$(\mathbf{C})_x = \left\{\vec{\theta}\colon \sum_{n=1}^{\infty} P(\theta_{-n} \cdots \theta_{-1}; x, x) = \infty\right\}, \qquad x \in \mathscr{X}. \tag{1.4}$$

Similarly, the conservative set $\hat{\mathbf{C}}$ in $\hat{\mathbf{S}}$ is given by

$$(\hat{\mathbf{C}})_{x,y} = \left\{\vec{\theta}\colon \sum_{n=1}^{\infty} P(\theta_{-n} \cdots \theta_{-1}; x, x) \times P(\theta_{-n} \cdots \theta_{-1}; y, y) = \infty\right\}, \qquad x,y \subset \mathscr{X} \tag{1.5}$$

(see [2]). Of course, the conservative set of a Markov process is closed. Now for any $F \in \mathscr{F}$, let

$$F^{(2)} = \Big\{\big(x,y,\vec{\theta}\big)\colon \big(x,\vec{\theta}\big) \in F \text{ and } \big(y,\vec{\theta}\big) \in F\Big\}. \tag{1.6}$$

Then $\mathbf{C}^{(2)}$ is closed in $\hat{\mathbf{S}}$ and $\mathbf{C}^{(2)} \supset \hat{\mathbf{C}}$, as is apparent from the above characterizations.

Let $\|\cdot\|$ be total variation norm and, given two distributions λ, ν on $\mathscr{X}$, let

$$\delta_n(\lambda,\nu,\vec{\theta}) = \|(\lambda - \nu)P(\theta_0 \cdots \theta_{n-1})\|. \tag{1.7}$$

When $\lambda(x) = 1$, write $\delta_n(x, \nu, \vec{\theta})$ and, if $\nu(y) = 1$ as well, then write $\delta(x, y, \vec{\theta})$. The δ_n are nonincreasing and we let $\delta = \lim_{n\to\infty} \delta_n$. Let $f\colon \mathbf{S} \to \mathbf{R}$ and $0 \le f \le 1$. Then we have the standard inequality

$$|P^n f(x, \vec{\theta}) - P^n f(y, \vec{\theta})| \le \delta_n(x, y, \vec{\theta})/2. \tag{1.8}$$

Let

$$S(x, \vec{\theta}, n) = \{y\colon P(\theta_0 \cdots \theta_{n-1}; x, y) > 0\}. \tag{1.9}$$

The following well-known result is noted for later use.

LEMMA 1. *Let* $s = \sup\{\delta_n(x', y', T^m\vec{\theta})\colon x' \in S(x, \vec{\theta}, m)$ *and* $y' \in S(y, \vec{\theta}, m)\}$. *Then*

$$\delta_{m+n}(x, y, \vec{\theta}) \le s\delta_m(x, y, \vec{\theta})/2. \tag{1.10}$$

We also need

LEMMA 2. $\hat{P}\delta_n \ge \delta_{n+1}$, $\hat{P}\delta \ge \delta$ *and* $\{(x, y, \vec{\theta})\colon \delta(x, y, \vec{\theta}) = 2\}$ *is closed in* $\hat{S}$.

PROOF. For the first assertion, note that $\hat{P}\delta_n(x, y, \vec{\theta})$ equals

$$\begin{aligned}
&\sum_{x', y' \in \mathscr{X}} P(\theta_0; x, x') P(\theta_0; y, y') \\
&\quad \times \sum_{z \in x} |P(\theta_1 \cdots \theta_n; x', z) - P(\theta_1 \cdots \theta_n; y', z)| \\
&\quad\quad \ge \sum_z \Big| \sum_{x', y'} P(\theta_0; x, x') P(\theta_0; y, y')(P(\theta_1 \cdots \theta_n; x', z) \\
&\quad\quad\quad\quad - P(\theta_1 \cdots \theta_n; y', z))\Big| \\
&\quad\quad = \sum_z |P(\theta_0 \cdots \theta_n; x, z) - P(\theta_0 \cdots \theta_n; y, z)| = \delta_{n+1}(x, y, \vec{\theta}).
\end{aligned}$$

The second assertion follows from the first and the dominated convergence theorem. Finally, since $\delta \le 2$ and $\hat{P}\delta \ge \delta$, the set where $\delta = 2$ must be closed. □

We want to study direct convergence of the transition probabilities: in effect, the zero set of δ. For homogeneous Markov chains this leads to the assumption of aperiodicity, then to cyclic decompositions to handle the periodic case. The following examples consider analogous behaviors for $\vec{\theta}$-chains.

EXAMPLE 1. Let

$$\mathscr{X} = \mathbf{Z}, \quad \Theta = [0, 1] \quad \text{and} \quad P(\theta; x, x+1) = 1 - P(\theta; x, x-1) = \theta.$$

In effect, this is the homogeneous birth and death in a random environment considered by Torrez [6]. This case resembles the classical random walk: X_n

alternates between the odd and even integers and these sets form a two-cycle. The standard approach to periodicity for homogeneous Markov chains, based on a periodic decomposition and on comparing products of transition probabilities at multiples of the period, works well here. In particular, under mild conditions on the environmental sequence, $\delta(x, y, \vec{\theta}) = 0$ or 2 according as $x - y$ is even or odd.

EXAMPLE 2. Let $\mathscr{X} = \mathbf{Z}$, $\Theta = \{0, 1\}$ and $P(0; x, x) = 1$, $P(1; x, x+1) = P(1; x, x-1) = 1/2$. Since this process can rest (when $\theta = 0$) as well as move, we might expect no periodicity, and this is the case, provided $\vec{\theta}$ is aperiodic. But even if the θ_n are independent, we have $\delta(x, y, \vec{\theta}) = 2$ when $x - y$ is odd. It should be noted that the marginal distribution of the $\{X_n\}$ sequence is that of a birth and death chain with a zero-one tail σ-field in this case. Nevertheless, this is not true for the conditional distributions of the $\vec{\theta}$-chains, and the failure of direct convergence for these chains when $x - y$ is odd is not due to periodicity.

EXAMPLE 3. Let $\Theta = \mathscr{X}$ and $P(\theta; x, \theta) = 1$. We have called this the perfect copy process. Evidently $\delta(x, y, \vec{\theta}) = 0$ for all $x, y, \vec{\theta}$, however, we could choose π so the $\vec{\theta}$ sequence is periodic, and then the $\vec{\theta}$-chain would exhibit periodic behavior.

These examples show periodicity is not adequate to delineate conditions for direct convergence of the $\vec{\theta}$-chains. In Section 2 we will examine an equivalence relation that is useful in this study. Then in Section 3 we consider conditions under which $\delta(x, y, \vec{\theta}) = 0$ on equivalence classes. Finally, in Section 4 we consider the frequency of positive return times to a set or state.

2. An equivalence. Let $\mathbf{M}$ be the maximal support for a finite invariant measure (see [2]).

LEMMA 3. *For μ-a.e. $(x, \vec{\theta}) \in \mathbf{M}$ and all $y, z \in S(x, \vec{\theta}, n)$ and $n \geq 1$,* $\delta(y, z, T^n\vec{\theta}) < 2$.

PROOF. Let $\varphi(x, \vec{\theta})$ be an invariant density with support $\mathbf{M}$, and set

$$d(x, \vec{\theta}) = \sup_{n \geq 1} \Big\{ \Sigma \varphi(y, T^n\vec{\theta}) \varphi(z, T^n\vec{\theta}) : y, z \in S(x, \vec{\theta}, n) \text{ such that } \delta(y, z, T^n\vec{\theta}) = 2 \Big\}.$$

If $y \in S(x, \vec{\theta}, n)$ then $S(y, T^n\vec{\theta}, k) \subset S(x, \vec{\theta}, n+k)$ for $k \geq 0$, hence $d(y, T^n\vec{\theta}) \leq d(x, \vec{\theta})$. Thus $d \geq Pd \geq 0$, and on C we have that d is harmonic ($d = Pd$) and d is $\mathscr{F}_i$ measurable (see Foguel [3]). If $d(x, \vec{\theta}) > 0$, then there exist n and $y, z \in S(x, \vec{\theta}, n)$ with $\delta(y, z, T^n\vec{\theta}) = 2$. But then $S(y, T^n\vec{\theta}, k) \cap S(z, T^n\vec{\theta}, k) = \varnothing$ for every $k \geq 0$, and it follows that

$$d(x, \vec{\theta}) \geq d(y, T^n\vec{\theta}) + d(z, T^n\vec{\theta}).$$

Since d is $\mathscr{F}_i$ measurable, the three values of d in the above inequality must be the same for μ-a.e. $(x, \vec{\theta})$. But this requires $d(x, \vec{\theta}) = 0$. Thus $d(x, \vec{\theta}) = 0$ a.e., and the lemma follows. □

We say that $(x, \vec{\theta})$ *meets* $(y, \vec{\theta})$, denoted $(x, \vec{\theta}) \leftrightarrow (y, \vec{\theta})$, if for some $n \geq 1$, $S(x, \vec{\theta}), n) \cap S(y, \vec{\theta}, n) \neq \varnothing$, and in this case we also say that $(x, \vec{\theta})$ *meets* $(y, \vec{\theta})$ *in n steps* and write $(x, \vec{\theta}) \leftrightarrow_n (y, \vec{\theta})$. Note that $(x, \vec{\theta}) \leftrightarrow_n (y, \vec{\theta})$ implies $(x, \vec{\theta}) \leftrightarrow_{n+k} (y, \vec{\theta})$ for all $k \geq 0$. Also, $(x, \vec{\theta}) \leftrightarrow (y, \vec{\theta})$ if and only if $\delta(x, y, \vec{\theta}) < 2$. It is easy to see that, while $\leftrightarrow$ is reflexive and symmetric, it need not be transitive on $\mathbf{S} - \mathbf{M}$. Let $\mathbf{C}_1$ be the set of all $(x, \vec{\theta}) \in \mathbf{C}$ for which Lemma 3 holds, that is, such that $\delta(y, z, T^n\vec{\theta}) < 2$ for all $y, z \in S(x, \vec{\theta}, n)$ and all $n \geq 1$. Then $\mathbf{C}_1$ is closed and *throughout the remainder of this section we restrict* $\mathbf{S}$ *to* $\mathbf{C}_1$.

PROPOSITION 1. $\leftrightarrow$ *is an equivalence relation for the process restricted to* $\mathbf{C}_1$.

PROOF. Let $(x, \vec{\theta}) \leftrightarrow_j (y, \vec{\theta})$ and $(y, \vec{\theta}) \leftrightarrow_k (z, \vec{\theta})$. Then for $l = \max\{j, k\}$, both pairs meet in l steps, so there exists $u \in S(x, \vec{\theta}, l) \cap s(y, \vec{\theta}, l)$ and $v \in S(y, \vec{\theta}, l) \cap S(z, \vec{\theta}, l)$. Since both $u, v \in S(y, \vec{\theta}, l)$, we have $\delta(u, v, T^l\vec{\theta}) < 2$, so u, v meet in, say, n steps. Then $(x, \vec{\theta}) \leftrightarrow_{l+n} (z, \vec{\theta})$. □

THEOREM 1. *Restrict* $\mathbf{S}$ *to* $\mathbf{C}_1$. *Let* $x' \in S(x, \vec{\theta}, n)$ *and* $y' \in S(y, \vec{\theta}, n)$. *Then* $(x, \vec{\theta}) \leftrightarrow (y, \vec{\theta})$ *if and only if* $(x', T^n\vec{\theta}) \leftrightarrow (y', T^n\vec{\theta})$.

PROOF. If $\delta(x, y, \vec{\theta}) = 2$, then since $\{(u, v, \vec{\psi}): \delta(u, v, \vec{\psi}) = 2\}$ is closed in $\hat{\mathbf{S}}$ and $\hat{P}^n((x, y, \vec{\theta}), (x', y', T^n\vec{\theta})) > 0$, we must have $\delta(x', y', T^n\vec{\theta}) = 2$. On the other hand, if $\delta(x, y, \vec{\theta}) < 2$, then $(x, \vec{\theta}) \leftrightarrow_k (y, \vec{\theta})$ for $k \geq k_0$ and some k_0. If $k_0 \leq n$, then take $k = n$ and we have $S(x, \vec{\theta}, n) \cap S(y, \vec{\theta}, n) \neq \varnothing$. By Proposition 1 all states in $S(x, \vec{\theta}, n) \cup S(y, \vec{\theta}, n)$ meet, so $\delta(x', y', T^n\vec{\theta}) < 2$. If $k_0 > n$, then take $x'' \in S(x', T^n\vec{\theta}, k_0 - n)$, $y'' \in S(y', T^n\vec{\theta}, k_0 - n)$. Now $(x, \vec{\theta}) \leftrightarrow_{k_0} (y, \vec{\theta})$ implies $S(x, \vec{\theta}, k_0) \cap S(y, \vec{\theta}, k_0)$ is nonempty; then Proposition 1 implies all states of $S(x, \vec{\theta}, k_0) \cup S(y, \vec{\theta}, k_0)$ meet. In particular $(x'', T^{k_0}\vec{\theta}) \leftrightarrow (y'', T^{k_0}\vec{\theta})$ and $\delta(x'', y'', T^{k_0}\vec{\theta}) < 2$. Since $\delta(x', y', T^n\vec{\theta}) = 2$ implies $\delta(x'', y'', T^{k_0}\vec{\theta}) = 2$ by the first part of the proof, we have $\delta(x', y', T^n\vec{\theta}) < 2$. □

Consider the equivalence classes $[(x, \vec{\theta})] = \{(y, \vec{\theta}) \in \mathbf{C}_1: (y, \vec{\theta}) \leftrightarrow (x, \vec{\theta})\}$. Theorem 1 implies that for each equivalence class $D = [(x, \vec{\theta})]$ there is an equivalence class $D' = [(x', T\vec{\theta})]$ such that $P(\theta_0; y, D') = I_D(y, \vec{\theta})$. In such a case we say *D maps into D' under P*. Theorem 1 also implies that distinct equivalence classes $[(x, \vec{\theta})]$ and $[(y, \vec{\theta})]$ map into distinct equivalence classes. Moreover, if no $[(x, \vec{\theta})]$, $x \in \mathscr{X}$, maps into a given $[(x', T\vec{\theta})]$, then $[(x', T\vec{\theta})]$ cannot be in $\mathbf{C}$. Thus this mapping under P of the equivalence classes in $(\mathbf{C}_1)^{\vec{\theta}}$ to equivalence classes in $(\mathbf{C}_1)^{T\vec{\theta}}$ is a bijection.

For $F \subset \mathbf{C}_1$ let $[F] = \bigcup_{(x,\vec{\theta}) \in F}[(x, \vec{\theta})]$. Note that

$$[F] = \mathbf{C}_1 \cap \left[\bigcup_{x \in \mathscr{X}} \left(\{x\} \times \bigcup_{y \in \mathscr{X}} \left((F)_y \cap \left\{ \vec{\theta} : \delta\left(x, y, \vec{\theta}\right) < 2 \right\} \right) \right) \right],$$

and since $\delta(x, y, \cdot\,)$ is measurable in $\vec{\theta}$, it follows that $[F] \in \mathscr{F} \cap \mathbf{C}_1$ whenever $F \in \mathscr{F} \cap \mathbf{C}_1$. Let $\mathscr{F}_d = \{F \in \mathscr{F} \cap \mathbf{C}_1 \colon F = [F]\}$. It is easy to see that $\mathscr{F}_d$ is a σ-field in $\mathbf{C}_1$. Note that Theorem 1 implies that P acting on $\mathscr{F}_d$ has the following property.

(P) For each $F \in \mathscr{F}_d$ and $n \geq 0$, there exist $G_n, H_n \in \mathscr{F}_d$ such that

$$I_F = P^n I_{G_n} \quad \text{and} \quad P^n I_F = I_{H_n}.$$

Also from Theorem 1 we obtain the following characterization on $\mathbf{C}_1$.

Corollary 1. *Let $F \in \mathscr{F} \cap \mathbf{C}_1$ and suppose there exist $n_k \to \infty$ and $G_k \in \mathscr{F} \cap \mathbf{C}_1$ such that $I_F = P^{n_k} I_{G_k}$, $k = 1, 2, \ldots$. Then $F \in \mathscr{F}_d$. In particular, $\mathscr{F}_i \subset \mathscr{F}_d$ and if $P^n I_F = I_F$ for any $n > 0$, then $F \in \mathscr{F}_d$.*

Proof. If $F \notin \mathscr{F}_d$, then there exist x, y and a set Γ of $\vec{\theta}$'s with $\pi(\Gamma) > 0$ such that $(x, \vec{\theta}) \in F$, $(y, \vec{\theta}) \notin F$ and $(x, \vec{\theta}) \leftrightarrow (y, \vec{\theta})$ for $\vec{\theta} \in \Gamma$. For each such $\vec{\theta}$ there exists an $n(\vec{\theta})$ such that $(x, \vec{\theta}) \leftrightarrow_n (y, \vec{\theta})$ for $n \geq n(\vec{\theta})$. Hence there must be a set $\Gamma_0 \subset \Gamma$ with $\pi(\Gamma_0) > 0$ and an n_0 such that $S(x, \vec{\theta}, n) \cap S(y, \vec{\theta}, n) \neq \varnothing$ for $n \geq n_0$ and $\vec{\theta} \in \Gamma_0$. Let $I_F = P^n I_G$ for some $n \geq n_0$. Now there must be a set $\Gamma_1 \subset \Gamma_0$ with $\pi(\Gamma_1) > 0$ and a z such that $z \in S(x, \vec{\theta}, n) \cap S(y, \vec{\theta}, n)$ for $\vec{\theta} \in \Gamma_1$. But

$$P(\theta_0 \cdots \theta_{n-1}; x, G) = 1 \Rightarrow (z, T^n \vec{\theta}) \in G,$$

$$P(\theta_0 \cdots \theta_{n-1}; y, G) = 0 \Rightarrow (z, T^n \vec{\theta}) \notin G,$$

and this contradiction shows $F \in \mathscr{F}_d$. □

For $F \in \mathscr{F}_d$ let $\rho(\vec{\theta}, F)$ be the number of distinct equivalence classes in $(F)^{\vec{\theta}}$. If $F = [\{x\} \times \Gamma]$ then $\rho(\vec{\theta}, F) = I_\Gamma$. The class of countable unions of sets of this form with $\Gamma \in \mathscr{B}$ equals $\mathscr{F}_d$, hence $\rho(\vec{\theta}, F)$ is $\mathscr{B}$ measurable in $\vec{\theta}$. Also $\rho(\vec{\theta}, F)$ is a measure in F for each $\vec{\theta}$. Let

$$\rho(F) = \int \rho(\vec{\theta}, F) \pi(d\vec{\theta}). \tag{2.1}$$

Then ρ is equivalent to μ restricted to $\mathscr{F}_d$. Moreover, since the mapping of equivalence classes under P is a bijection, $\rho P = \rho$ on $\mathscr{F}_d$.

Note also that, for $F \in \mathscr{F}_i$ with $F \subset \mathbf{C}_1$, $(F)^{\vec{\theta}}$ has the same number of equivalence classes for a.e. $\vec{\theta}$, so $\rho(\vec{\theta}, F) = \rho(F)$ a.e. when π is ergodic.

Applying the Chacon–Ornstein and Chacon identification theorems as in [2], we have

THEOREM 2. *Let F be an atom of $\mathscr{F}_i$ with $F \subset \mathbf{C}_1$ and let $f, g \in L_1(\mathbf{C}_1, \mathscr{F}_d, \rho)$ with $g \geq 0$ and $\int_F g\,dp < 0$. Then for μ-a.e. $(x, \vec{\theta}) \in F$,*

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n} P^k f(x,\vec{\theta})}{\sum_{k=0}^{n} P^k g(x,\vec{\theta})} = \frac{\int_F f\,dp}{\int_F g\,dp}. \tag{2.2}$$

COROLLARY 2. *Let F be an atom of $\mathscr{F}_i$ with $F \subset \mathbf{C}_1$ and let $G, H \in \mathscr{F}_d$ with $\rho(G \cap F) < \infty$ and $0 < \rho(H \cap F) \leq \infty$. Then for μ-a.e. $(x, \vec{\theta}) \in F$,*

$$\lim_{n\to\infty} \frac{\sum_{k=0}^{n} P^k\big((x,\vec{\theta}), G\big)}{\sum_{k=0}^{n} P^k\big((x,\vec{\theta}), H\big)} = \frac{\rho(G \cap F)}{\rho(H \cap F)}. \tag{2.3}$$

In particular, for each x and μ-a.e. $(y, \vec{\theta}) \in F$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P\Big(\theta_0 \cdots \theta_{k-1}; y, \big[(x, T^k\vec{\theta})\big]\Big) = \frac{\pi((F)x)}{\rho(F)} \tag{2.4}$$

(*interpret* $[(x, T^k\vec{\theta})] = \varnothing$ *if* $(x, T^k\vec{\theta}) \notin \mathbf{C}_1$).

PROOF. Since μ and ρ are equivalent on $\mathscr{F}_d$, the convergences hold a.e.-μ. In the corollary we can let $\rho(H \cap F) = \infty$ since in any case there exist $H_n \uparrow H$, $H_n \in \mathscr{F}_d$ and $\rho(H_n \cap F) < \infty$. For example, take finite $A_n \uparrow \mathscr{X}$ and $H_n = H \cap [A_n \times \vec{\Theta}]$. The lim sup of the ratios is at most $\inf_n \rho(G \cap F)/\rho(H_n \cap F) = 0$. For the last assertion, take $G = [(\mathbf{C}_1)_x]$ and $H = F$, so $\rho(\vec{\theta}, G \cap F) = I_{(F)x}(\vec{\Theta})$ and $\rho(G \cap F) = \pi(F)_x$. □

Now suppose $\varphi \geq 0$ is a nontrivial finite invariant element of L_1 with support F_φ. Let $\Phi(A) = \int_A \varphi\,d\mu$, so Φ is a finite invariant measure. Both Φ and ρ restricted to $\mathscr{F}_d \cap F_\varphi$ are nontrivial invariant measures. If F_φ is an atom of $\mathscr{F}_i$, then Φ and ρ can differ by at most a multiplicative constant, hence $\Phi/\Phi(F_\varphi) = \rho/\rho(F_\varphi)$. In particular $\rho(\vec{\theta}, F_\varphi) = \rho(F_\varphi)$, the number of equivalenceclasses in $(F_\varphi)^{\theta}$, is finite and

$$\sum_{y:\,(y,\vec{\theta}) \in [(x,\vec{\theta})]} \varphi(y,\vec{\theta}) = \Phi(F_\varphi)/\rho(F_\varphi) \tag{2.5}$$

is the same for μ-a.e. $(x, \vec{\theta}) \in F_\varphi$. Also, Theorem 1 implies that, if $P(\theta_0 \cdots \theta_{n-1}; x, y) > 0$, then

$$\sum_{w:\,(w,\vec{\theta}) \in [(x,\vec{\theta})]} \varphi(w,\vec{\theta}) = \sum_{z:\,(z,T^n\vec{\theta}) \in [(y,T^n\vec{\theta})]} \varphi(z, T^n\vec{\theta}) \tag{2.6}$$

for μ-a.e. $(x, \vec{\theta}) \in F_\varphi$. [For each n, x and y for which $P(\theta_0 \cdots \theta_{n-1}; x, y) > 0$, (2.6) may fail for a π-null set $N_{n,x,y}$ of $\vec{\theta}$s. Then $N = \bigcup_{n,x,y} N_{n,x,y}$ is π-null and (2.6) holds for $\vec{\theta} \notin N$.]

In many applications it happens that the equivalence classes $[(x, \vec{\theta})]$ are the same for each given x and π-a.e. $\vec{\theta}$. If $\rho(\vec{\theta}, \mathbf{C}_1) = \mathrm{d}$, then there are disjoint sets $D_1, \dots, D_d$ in $\mathscr{X}$ such that $\mathbf{C}_1 = (\bigcup_{j=1}^{d} D_j) \times \vec{\Theta}$ up to a μ-null set. Then the mapping under P from equivalence classes for $\vec{\theta}$ to those for $T\vec{\theta}$ depends on the transition probability $P(\vec{\theta}_0)$, hence on θ_0, and is a permutation among the d equivalence classes. Let $\sigma(\theta)$ denote the permutation when $\theta_0 = \theta$. We conclude this section by considering several special cases.

CASE 1. $\sigma(\theta) = \sigma$ the same for all θ. Then if $\mathscr{F}_i$ is trivial, we must have $D_j, \sigma(D_j), \sigma^2(D_j), \dots, \sigma^{d-1}(D_j)$ all distinct so the d sets move in a cycle as in the theory of homogeneous chains.

CASE 2. $\sigma(\theta_n), n \in \mathbf{Z}$, is a periodic sequence, say of period e, with cycle $\sigma_0, \dots, \sigma_{e-1}$. By considering the permutation $\bar{\sigma} = \sigma_{e-1} \circ \cdots \circ \sigma_0$ it is easy to see that there will be a periodic sequence of the equivalence classes selected by the permutations, where the period is at most ed and, if $\mathscr{F}_i$ is trivial, at least d.

CASE 3. If the $\sigma(\theta_n)$ are aperiodic, then the $\vec{\theta}$-chain will not exhibit periodicity. In this case the closest we can come to this notion is to consider the sequence of random times $\tau_1, \tau_2, \dots$ such that $\sigma(\theta_{\tau_n}) \cdots \sigma(\theta_{\tau_{n-1}+1})$ is the identity in the permutation group.

3. Direct convergence. We look for conditions under which

$$\delta_n\big(x, y, \vec{\theta}\big) = \| P(\theta_0 \cdots \theta_{n-1}; x, \cdot) - P(\sigma_0 \cdots \sigma_{n-1}; y, \cdot) \| \to 0$$

as $n \to \infty$, in other words, that $\delta(x, y, \vec{\theta}) = 0$. In the next result, $\mathscr{T}$ denotes the tail σ-field of the $\{X_n\}$ sequence.

THEOREM 3. *Let F be a closed set in $\mathbf{S}$ and restrict the process to F. Then the following four conditions are equivalent.*

(i) *T is zero-one under $P_{x,\vec{\theta}}$ for μ-a.e. $(x, \vec{\theta}) \in F$.*

(ii) *For μ-a.e. $(x, \vec{\theta}) \in F$, every $n \geq 1$ and $y, z \in S(x, \vec{\theta}, n)$, $\delta(y, z, T^n\vec{\theta}) = 0$.*

(iii) *$\{(x, y, \vec{\theta}) \in F^{(2)}: \delta(x, y, \vec{\theta}) = 0\}$ is closed for $\hat{P}$ on $F^{(2)}$.*

(iv) *$\hat{P}^n \delta = \delta$ for $n \geq 1$ on $F^{(2)} \cap \Delta$, where Δ is the diagonal in $\hat{\mathbf{S}}$.*

Under any of these conditions we have $\hat{P}\delta = \delta$ and $\delta = 0$ or 2 a.e. on $F^{(2)}$.

PROOF. (a) The equivalence of (i) and (ii) is easily derived from the Blackwell–Freedman theorem [1] and close to a result of Iosifescu [4] that $\mathscr{T}$ is zero-one under $P_{x,\vec{\theta}}$ if and only if (in our notation)

$$\delta\big(y, P(\theta_0 \cdots \theta_{n-1}; x, \cdot), T^n\vec{\theta}\big) = 0$$

for each $y \in S(x, \vec{\theta}, n)$ and each $n \geq 1$. Since, for $y, z \in S(x, \vec{\theta}, n)$,

$$\delta\big(y, z, T^n\vec{\theta}\big) \leq \delta\big(y, P(\theta_0 \cdots \theta_{n-1}; x, \cdot), T^n\vec{\theta}\big) + \delta\big(z, P(\theta_0 \cdots \theta_{n-1}; x, \cdot), T^n\vec{\theta}\big),$$

it is clear that (i) implies (ii). But if (ii) holds, then for $y \in S(x, \vec{\theta}, n)$,

$$\begin{aligned}
&\delta\big(y, P(\theta_0 \cdots \theta_{n-1}; x, \cdot), T^n\vec{\theta}\big) \\
&= \lim_{m\to\infty} \sum_{z\in\mathscr{X}} |P(\theta_n \cdots \theta_{n+m}; y, z) - P(\theta_0 \cdots \theta_{n+m}; x, z)| \\
&\leq \lim_{m\to\infty} \sum_{w\in S(x,\vec{\theta},n)} P(\theta_0 \cdots \theta_{n-1}; x, w) \\
&\quad \times \sum_{z\in\mathscr{X}} |P(\theta_n \cdots \theta_{n+m}; y, z) - P(\theta_n \cdots \theta_{n+m}; w, z)| \\
&= \lim_{m\to\infty} \sum_{w\in S(x,\vec{\theta},n)} P(\theta_0 \cdots \theta_{n-1}; x, w)\delta_m\big(y, w, T^n\vec{\theta}\big) = 0
\end{aligned}$$

by the bounded convergence theorem when (ii) holds, so (ii) implies (i).

(b) Now suppose (ii) holds and that $\delta(x, y, \vec{\theta}) < 2$. Then $(S(x, \vec{\theta}, n) \cap S(y, \vec{\theta}, n) \neq \varnothing$ for some n, and (ii) implies $\delta(x', y', T^n\vec{\theta}) = 0$ for all $x', y' \in S(x, \vec{\theta}, n) \cup S(y, \vec{\theta}, n)$. Then $\delta(x, y, \vec{\theta}) = 0$ by Lemma 1. Thus $\delta = 0$ or 2 on $F^{(2)}$. If $P(\theta_0; x, x') > 0$ and $P(\theta_0; y, y') = p > 0$ and $\delta(x', y', T\vec{\theta}) = 2$, then $S(x', T\vec{\theta}, n) \cap S(y', T\vec{\theta}, n) = \varnothing$ for all n. If $x'' \in S(x', T\vec{\theta}, n)$ and $y'' \in S(y', T\vec{\theta}, n)$, then $\delta(x'', y'', T^{n+1}\vec{\theta}) = 2$, and since $S(x', T\vec{\theta}, n) \subset S(x, \vec{\theta}, n + 1)$, condition (ii) requires $S(x, \vec{\theta}, n + 1) \cap S(y', T\vec{\theta}, n) = \varnothing$ for all n. Thus $P(\theta_0 \cdots \theta_n; x, S(y', T\vec{\theta}, n)) = 0$ and $P(\theta_0 \cdots \theta_n; y, S(y', T\vec{\theta}, n)) \geq p$ for each n. But then $\delta(x, y, \vec{\theta}) \geq p$, hence $\delta(x, y, \vec{\theta}) = 0$ implies $\delta(x', y', T\vec{\theta}) = 0$ in this case, and (iii) follows.

(c) Conversely, if (iii) holds, then since $\delta(x, x, \vec{\theta}) = 0$, clearly (ii) holds. That (iii) implies $\hat{P}\delta = \delta$ on $F^{(2)}$ follows since $\delta = 0$ or 2 and $\{(x, y, \vec{\theta})$: $\delta(x, y, \vec{\theta}) = k\}$ is closed for $k = 0$ or 2. In particular, (iii) implies (iv). Finally, (iv) requires $\hat{P}^n\delta(x, x, \vec{\theta}) = \delta(x, x, \vec{\theta}) = 0$ for $n \geq 1$, which implies (ii). □

COROLLARY 3. *If $\{(x, x, \vec{\theta})$: $(x, \vec{\theta}) \in F\} \subset \hat{\mathbf{C}}$, then the hypotheses of Theorem 3 hold.*

COROLLARY 4. *If F is the support of a finite invariant measure then the hypotheses of Theorem 3 hold.*

PROOF. By Lemma 2, $\hat{P}\delta \geq \delta$, hence $\hat{P}\delta = \delta$ on $\hat{\mathbf{C}}$ and condition (iv) of Theorem 3 holds under the hypothesis of Corollary 3. Under the hypothesis of Corollary 4, $F^{(2)}$ is also the support of a finite invariant measure on $\hat{\mathbf{S}}$: If $\varphi(x, \vec{\theta})$ is the invariant density for P, then $\hat{\varphi}(x, y, \vec{\theta}) = \varphi(x, \vec{\theta})\varphi(y, \vec{\theta})$ is invariant for $\hat{P}$. Necessarily, $F^{(2)} \subset \hat{\mathbf{C}}$ in the case. □

Note, it is established in [2] that a nontrivial finite invariant measure exists on S if and only if, the π measure of the set

$$\Delta = \left\{ \vec{\theta} \colon \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P(\theta_0 \cdots \theta_k; x, y) = 0 \text{ for all } x, y \in \mathscr{X} \right\}$$

is less than one. In particular, such a measure always exists when $\mathscr{X}$ is finite.

We say that $(x, \vec{\theta})$ is *convergence equivalent* to $(y, \vec{\theta})$, denoted $(x, \vec{\theta}) \sim (y, \vec{\theta})$, if $\delta(x, y, \vec{\theta}) = 0$. Clearly this *is* an equivalence relation and is well-defined on $\mathbf{S}$, while on $\mathbf{C}_1$, $\sim$ is stronger than $\leftrightarrow$. Now suppose the hypotheses of Theorem 3 hold on F. Note, Theorem 3 does not require $F \subset \mathbf{C}_1$, but when $F \cap \mathbf{C}_1$ is non-null, the two equivalence relations coincide on this set.

COROLLARY 5. *Let φ be an invariant probability density whose support is F_φ. Suppose F_φ is an atom of F_i (equivalently, suppose P_φ is ergodic). For each $\vec{\theta}$ and $\leftrightarrow$ equivalence class $D = [(x, \vec{\theta})]$, define $\theta_D(y)$ as a probability on $\mathscr{X}$ with support $\{y \colon (y, \vec{\theta}) \in D\}$ by*

$$\varphi_D(y) = \rho(F_\varphi)\varphi(y, \vec{\theta}) I_D(y, \vec{\theta}).$$

Then for μ-a.e. $(x, \vec{\theta}) \in F_\varphi$ and every y such that $(y, \vec{\theta}) \in D = [(x, \vec{\theta})]$, $\delta(y, \Phi_D, \vec{\theta}) = 0$.

PROOF. By Corollary 4, Theorem 3 applies, so $\leftrightarrow$ and $\sim$ coincide on F_φ (of course, the support of a finite invariant measure is contained in $\mathbf{M}$). Thus, for $(x, \vec{\theta})$, $(y, \vec{\theta}) \in D$, $\delta(x, y, \vec{\theta}) = 0$, (2.5) implies $\sum_y \varphi_D(y) = 1$, hence

$$\begin{aligned} \delta_n(x, \varphi_D, \vec{\theta}) &= \| P(\theta_0 \cdots \theta_{n-1}; x, \cdot) - \Phi_D P(\theta_0 \cdots \theta_{n-1})(\cdot) \| \\ &\le \rho(F_\varphi) \sum_{y \colon (y, \vec{\theta}) \in D} \varphi(y, \vec{\theta}) \delta_n(x, y, \vec{\theta}) \to 0 \end{aligned}$$

as $n \to \infty$, by the dominated convergence theorem. □

By a similar argument we have

COROLLARY 6. *Let the hypotheses of Theorem 3 hold on a set F. Let λ be a distribution on $[(x, \vec{\theta})]$ and ν be a distribution on $[(y, T^n\vec{\theta})]$ for some $n \ge 0$. Assume $P(\theta_0 \cdots \theta_{n-1}; x, y) > 0$ (or that $y = x$ when $n = 0$). Then $\delta(\lambda P(\theta_0 \cdots \theta_{n-1}), \nu, T^n\vec{\theta}) = 0$..*

4. The frequency of positive return times. In this section we will examine sets of positive return times for the $\vec{\theta}$-chains. Given any $A \subset \mathscr{X}$ and

$n \geq 1$, let

$$N_A(x, \vec{\theta}, n) = \{k\colon 1 \leq k \leq n \text{ and } P(\theta_0 \cdots \theta_{k-1}; x, A) > 0\}.$$

Note that, for $m \geq 1$, $y \in S(x, \vec{\theta}, n)$ implies

$$\left(N_A(x, \vec{\theta}, n) \cup \left\{N_A\left(y, T^n\vec{\theta}, m\right) + n\right\}\right) \subset N_A(x, \vec{\theta}, n + m).$$

Hence, if $\{X_n\}$ is realization of the $\vec{\theta}$-chain and $i < j < k$, then

$$\begin{aligned}(4.1)\qquad &\left(N_A\left(X_i, T^i\vec{\theta}, j - i\right) \cup \left\{N_A\left(X_j, T^j\vec{\theta}, k - j\right) + j\right\}\right)\\ &\qquad \subset N_A\left(X_i, T^i\vec{\theta}, k - i\right) \quad \text{a.s.}\end{aligned}$$

Now let $z_A(x, \vec{\theta}, n)$ be the cardinal number of $N_A(x, \vec{\theta}, n)$ and, given the realization $\{X_n\}$ of the $\vec{\theta}$-chain, for $0 \leq j < k$ let

$$Z_A(j, k) = z_A\left(X_j, T^j\vec{\theta}, k - j\right).$$

Then the relation (4.1) implies

$$(4.2)\qquad Z_A(i, j) + Z_A(j, k) \leq Z_A(i, k).$$

Taking $i = 0$, $j = 1$, $k = n + 1$ and starting the $\vec{\theta}$-chain at $(x, \vec{\theta})$, we obtain

$$z_A(x, \vec{\theta}, 1) + z_A\left(X_1, T\vec{\theta}, n\right) \leq z_A(x, \vec{\theta}, n + 1),$$

and dropping the first term on the left and applying $E_{x,\theta'}$ to the second term yields

$$(4.3)\qquad Pz_A(x, \vec{\theta}, n) \leq z_A(x, \vec{\theta}, n + 1).$$

THEOREM 4. *Suppose there is a finite invariant measure Φ. Then for Φ-a.e. $(x, \vec{\theta})$ the limit*

$$r_A(x, \vec{\theta}) = \lim_{n \to \infty} \frac{1}{n} z_A(x, \vec{\theta}, n)$$

exists. Restricting **S** *to the support of Φ, r_A is harmonic and $\mathscr{F}_i$ measurable.*

PROOF. For the distribution P_φ, the sequence $\{(X_n, T^n\vec{\theta})\}$ is stationary, hence the distribution of the $Z_A(j, k)$, both singly and for the joint distribution of n terms, depends only on $k - j$. This, combined with relation (4.2), shows that $-Z_A(j, k)$ is a subadditive process in the sense of Kingman [5]. Moreover, $-Z_A(j, k) \geq -(k - j)$. Then Kingman's subadditive ergodic theorem asserts that

$$\lim_{n \to \infty} \frac{1}{n} Z_A(0, n) = \lim_{n \to \infty} \frac{1}{n} z_A\left(X_0, \vec{\theta}, n\right)$$

exists a.s. $-P_\varphi$. The P_φ null set on which convergence fails must be $P_{x,\theta'}$ null for φ-a.e. $(x, \vec{\theta})$, and the convergence assertion follows. Furthermore, under $P_{x,\theta'}$ each $z_A(x, \vec{\theta}, n)$ is a constant, so $r_A(x, \vec{\theta})$ is a (nonrandom) function of

$(x, \vec{\theta})$. Now the dominated convergence theorem and relation (4.3) imply, for **S** restricted to the support of Φ, that

$$Pr_A(x, \vec{\theta}) = P \lim_n \frac{1}{n} z_A(x, \vec{\theta}, n)$$

$$= \lim_n \frac{1}{n} P z_A(x, \vec{\theta}, n) \leq \lim_n \frac{1}{n} z_A(x, \vec{\theta}, n+1) = r_A(x, \vec{\theta}).$$

Thus r_A is superharmonic and, since the support of any finite invariant measure is contained in **C**, r_A is harmonic and $\mathscr{F}_i$ measurable. (See [3].) □

We call $r_A(x, \vec{\theta})$ *the relative frequency of positive return times to A*. Note that when P_φ is ergodic, the support of φ is an atom of $\mathscr{F}_i$, hence $r_A = r_A(x, \vec{\theta})$ is constant. In particular, when P_φ is ergodic there is a constant relative frequency $r_x := r_{\{x\}}$ of positive return times to any state x. The next result identifies the limit r_x in this case and should be compared to the result for a homogeneous Markov chain that $r_x = 1/c$, where c is the period of x.

THEOREM 5. *Let φ be an invariant probability density with support F_φ. Let F_φ be an atom of $\mathscr{F}_i$ (equivalently, let P_φ be ergodic). Then for each x and μ-a.e. $(y, \vec{\theta}) \in F_\varphi$,*

$$r_x = r_x(y, \vec{\theta}) = \pi((F_\varphi)_x) / \rho(F_\varphi). \tag{4.4}$$

PROOF. Since $r_x(y, \vec{\theta})$ is the frequency of positive terms in the sequence $\{P(\theta_0 \cdots \theta_{n-1}; y, x)\}$ and since $P(\theta_0 \cdots \theta_{n-1}; y, [(x, T^n\vec{\theta})]) = 0$ or 1, Corollary 2 implies $r_x \leq \pi((F_\varphi)_x)/\rho(F_\varphi)$ a.e. This proves the theorem if $\pi((F_\varphi)_x) = 0$. Assume $\pi((F_\varphi)_x) > 0$ and let $D = [(y, \vec{\theta})]$ and $\varphi_D(z) = \rho(F_\varphi)\varphi(z, \vec{\theta})I_D(z, \vec{\theta})$. By stationarity of φ,

$$\sum_{z \in \mathscr{X}} \varphi(z, \vec{\theta}) P(\theta_0 \cdots \theta_{n-1}; z, x) = \varphi(x, T^n \vec{\theta})$$

and by (2.6), $\varphi_D(z) = \rho(F_\varphi)\varphi(z, \vec{\theta})$ for all states $(z, \vec{\theta})$ mapping into $[(x, T^n\vec{\theta})]$ when $P(\theta_0 \cdots \theta_{n-1}; y, [(x, T^n\vec{\theta})]) > 0$. Hence

$$\sum_{z \in \mathscr{X}} \varphi_D(z) P(\theta_0 \cdots \theta_{n-1}; z, x)$$

$$= \begin{cases} \rho(F_\varphi)\varphi(x, T^n\vec{\theta}) & \text{if } P\left(\theta_0 \cdots \theta_{n-1}; y, \left[(x, T^n\vec{\theta})\right]\right) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\varphi(x, \vec{\eta}) > 0$ for $\vec{\eta} \in (F_\varphi)_x$, given $\varepsilon > 0$ there exist $\varepsilon' > 0$ and $\Gamma \in \vec{\mathscr{B}}$ such that $\pi(\Gamma) > 1 - \varepsilon$ and $\rho(F_\varphi)\varphi(x, \vec{\eta}) > \varepsilon'$ for $\vec{\eta} \in \Gamma \cap (F_\varphi)_x$. Corollary 5 implies there exists an $n_0 = n_0(\varepsilon')$ such that for $n \geq n_0$

$$\left| P(\theta_0 \cdots \theta_{n-1}; y, x) - \sum_z \varphi_D(z) P(\theta_0 \cdots \theta_{n-1}; z, x) \right| < \varepsilon'.$$

Thus $P(\theta_0 \cdots \theta_{n-1}; y, x) > 0$ whenever

$$n \geq n_0, T^n\vec{\theta} \in \Gamma \quad \text{and} \quad P\Big(\theta_0 \cdots \theta_{n-1}; y, \big[(x, T^n\vec{\theta})\big]\Big) > 0$$

since F_φ is closed, so $T^n\vec{\theta} \in (F_\varphi)_x$ as well. Also, since F_φ is an atom of $\mathscr{F}_i$, the ergodic theorem implies

$$\lim \frac{1}{n} \sum_{k=1}^{n} I_T(T^k\vec{\theta}) \geq \pi(\Gamma) > 1 - \varepsilon$$

for π-a.e. $\vec{\theta} \in \bigcup_y (F_\varphi)_y$. Applying Corollary 2 again,

$$r_x \geq \lim \frac{1}{n} \sum_{k=1}^{n} P\Big(\theta_0 \cdots \theta_k; y, \big[(x, T^k\vec{\theta})\big]\Big) - \lim \frac{1}{n} \sum_{k=1}^{n} I_{\Gamma^c}(T^k\vec{\theta}) \geq \frac{\pi(F_\varphi)_x}{\rho(F_\varphi)} - \varepsilon$$

and the theorem follows since $\varepsilon > 0$ is arbitrary. □

REFERENCES

[1] Blackwell, D. and Freedman, D. (1964). The tail σ-field of a Markov chain and a theorem of Orey. *Ann. Math. Statist.* **35** 1291–1295.
[2] Cogburn, R. (1984). The ergodic theory of Markov chains in random environments. *Z. Wahrsch. verw. Gebiete* **66** 109–128.
[3] Foguel, S. R. (1969). *The Ergodic Theory of Markov Processes*. Van Nostrand, Princeton, N.J.
[4] Iosifescu, M. (1972). On finite tail σ-algebras. *Z. Wahrsch. verw. Gebiete* **24** 159–166.
[5] Kingman, J. F. C. (1973). Subadditive ergodic theory. *Ann. Probab.* **1** 883–909.
[6] Torrez, W. C. (1978). The birth and death chain in a random environment: Instability and extinction theorems. *Ann. Probab.* **6** 1026–1043.

Department of Mathematics and Statistics
University of New Mexico
Albuquerque, New Mexico 87131