The Annals of Probability
1976, Vol. 4, No. 5, 850–858

PARTIAL COUPLING AND LOSS OF MEMORY FOR MARKOV CHAINS

BY DAVID GRIFFEATH
Cornell University

Coupling methods are used to obtain a structure theorem for the atomic decomposition of the tail σ-algebra of an arbitrary nonhomogeneous Markov chain. Various related results are also derived by coupling.

1. Introduction. The structure of the tail σ-algebra $\mathscr{T}$ of a (denumerable) Markov chain has been the subject of a number of recent papers. Blackwell and Freedman [2] showed that $\mathscr{T}$ is finitely atomic for irreducible recurrent homogeneous chains, proved the equivalence of trivial $\mathscr{T}$ and a mixing property for arbitrary homogeneous chains, and showed that these properties imply weak ergodicity. Jamison and Orey [10] proved the equivalence of the above three conditions, and their equivalence to the lack of nonconstant space-time harmonic functions, again for the homogeneous case. Next, Bartfay and Révész [1] extended some of these results to nonhomogeneous chains with partial mixing to obtain conditions for finiteness of $\mathscr{T}$, and generalized 0–1 laws. Iosifescu [9] gave further conditions for finiteness of $\mathscr{T}$, including a necessary and sufficient form of partial weak ergodicity. Finally, Cohn [4] has used reverse Markov chains to obtain a detailed description of $\mathscr{T}$ in the nonhomogeneous setting. Namely, he has shown that if $\mathscr{F}_0{}^m$ and $\mathscr{F}_n{}^\infty$ denote the σ-algebras of information from time 0 to m and n to ∞ respectively, then

$$\begin{aligned}
&\lim_{m\to\infty}\lim_{n\to\infty}\sup_{B\in\mathscr{F}_0{}^m}\{P(B)-P(B\,|\,\mathscr{F}_n{}^\infty)(\omega)\} \\
(1)\qquad &\quad = 1-P(A) \qquad \text{for almost all } \omega \text{ in the atom } A \text{ of } \mathscr{T}, \\
&\quad = 1 \qquad\qquad\quad \text{for almost all } \omega \text{ in no atom of } \mathscr{T}.
\end{aligned}$$

As a consequence he is able to obtain various structure results for $\mathscr{T}$, including the fact that any finite chain with n states has an atomic tail σ-algebra with at most n atoms (cf. Senchenko [14]).

In a related development, Markov chain coupling techniques have been used to study weak ergodicity of homogeneous and nonhomogeneous Markov chains [7, 8, 12]. The object of coupling is to construct two copies of a given Markov chain on a product state space, with an interdependence yielding ergodic properties. This strategy was explained in some detail in the author's paper [7], where examples are given, and a "maximal coupling" is constructed (all for the homogeneous case).

Received April 9, 1975; revised January 29, 1976.
AMS 1970 *subject classification*. Primary 60J05.
Key words and phrases. Nonhomogeneous Markov chains, coupling, tail σ-algebra, mixing, weak ergodicity, 0–1 law.

The present paper may be viewed as a continuation of [7]; our objective is to show how maximal coupling can be used to study asymptotic properties of nonhomogeneous Markov chains. In Section 3 this approach leads to a structure theorem which might be described as dual to Cohn's; namely, we show that

$$\lim_{m\to\infty} \lim_{n\to\infty} \sup_{B\in\mathscr{F}_n^\infty} \{P(B\mid\mathscr{F}_0^m) - P(B)\}$$

takes on the same values as (1) a.s. Our "forward" approach has the advantage that the main results of [1] and [9], as well as some additional ones, are easy consequences. At the same time, slight generalizations of the applications in [4] may be derived by our method. This is the subject matter of Section 4.

2. Notation. Let S be a finite or countable set of integers, $\mathbb{N} = \{0, 1, \cdots\}$, $\Omega = S^{\mathbb{N}}$. For $\omega = (\omega_0, \omega_1, \cdots) \in \Omega$, $\xi_n : \Omega \to S$ is given by $\omega \mapsto \omega_n$. Let $\mathscr{F} = \sigma\langle(\xi_n); n \in \mathbb{N}\rangle$ be the σ-algebra on Ω generated by the ξ_n. Write $\mathscr{F}_m{}^n = \sigma\langle(\xi_r; m \leqq r \leqq n)\rangle$, $\mathscr{F}_m{}^\infty = \sigma\langle(\xi_r; m \leqq r < \infty)\rangle$, and abbreviate $\mathscr{F}_m = \mathscr{F}_m{}^m$. The *tail σ-algebra* is $\mathscr{T} = \bigcap_{m\in\mathbb{N}} \mathscr{F}_m{}^\infty$. A (1-step) *transition function* $\pi = (\pi(m))_{m\in\mathbb{N}}$ is a sequence of stochastic matrices $\pi(m) = (\pi_{ij}(m))_{i,j\in S}$. If $\iota = (\iota_k)_{k\in S}$ is an initial probability distribution on S, and π a transition function, then the Markov measure $P = P(\iota, \pi)$ given by ι and π is the unique measure on $(\Omega, \mathscr{F})$ satisfying $P(\xi_0 = k) = \iota_k$, and for $n \geqq 0$,

$$P(\xi_{m+1} = j\mid\mathscr{F}_0{}^m)(\omega) = P(\xi_{m+1} = j\mid\mathscr{F}_m)(\omega) = \pi_{ij}(m) \qquad \text{on} \quad \{\xi_m = i\}\,.$$

For our purposes, any (denumerable) Markov chain may be identified with the coordinate process (ξ_n) on $(\Omega, \mathscr{F}, P)$ for some $P = P(\iota, \pi)$ manufactured according to the above recipe. We regard ι and π, and hence P, as *fixed* throughout the discussion.

Denote $p_j(n) = P(\xi_n = j)$, $p_{ij}(m, n) = P(\xi_n = j\mid\xi_m = i)$, $i, j \in S$, $m, n \in \mathbb{N}$. For $m \in \mathbb{N}$, define $\theta^m : \Omega \to \Omega$ by $\xi_n \circ \theta^m = \xi_{m+n}$, $n \in \mathbb{N}$. Viewing θ^m as a set map, $\theta^m B = \{\theta^m\omega; \omega \in B\}$ and $\theta^{-m}B = \{\omega : \theta^m\omega \in B\}$, $B \in \mathscr{F}$. Both of these sets are in $\mathscr{F}$, though in general $\theta^{-m}\theta^m B \neq B$ unless $B \in \mathscr{F}_m{}^\infty$. For $m \in \mathbb{N}$, let the measure P^m and the conditional measure $P^m_{\xi_m}$ be defined by

$$P^m(B) = P(\theta^{-m}B)\,, \qquad P^m_{\xi_m(\omega)}(B) = P(\theta^{-m}B\mid\mathscr{F}_0{}^m)(\omega)\,, \qquad B \in \mathscr{F}\,.$$

The following notation is necessary to formulate Markov chain coupling (cf. [7]). Let $\tilde{S} = S \times S$, $\tilde{\Omega} = \tilde{S}^{\mathbb{N}}$; $\tilde{\xi}_n = (\xi_n{}^1, \xi_n{}^2)$ is the bivariate coordinate map sending $\tilde{\omega} = ((\omega_0{}^1, \omega_0{}^2), (\omega_1{}^1, \omega_1{}^2), \cdots) \in \tilde{\Omega}$ to $\tilde{\omega}_n = (\omega_n{}^1, \omega_n{}^2)$. $\tilde{\mathscr{F}} = \sigma\langle(\tilde{\xi}_n); n \in \mathbb{N}\rangle$. It is sometimes convenient to think of $\tilde{\Omega}$ as $S^{\mathbb{N}} \times S^{\mathbb{N}}$, with $\tilde{\omega} = (\omega^1, \omega^2) \in \tilde{\Omega}$. The *diagonal* of $\tilde{S}$ is $D = \{(s, s); s \in S\}$, and $\tau_D(\tilde{\omega}) = \min\{n \in \mathbb{N} : \tilde{\omega}_n \in D\}$ ($= \infty$ if $\tilde{\omega}_n \notin D$ for every n).

Finally, if a and b are real numbers, let $a \wedge b = \min\{a, b\}$, $a \vee b = \max\{a, b\}$, $a^+ = \max\{a, 0\}$ and $[a]$ = the greatest integer in a.

3. A structure theorem for $\mathscr{T}$. Recall that the set $A \in \mathscr{T}$ is an *atom* of $\mathscr{T}$ (w.r.t. P) iff $P(A) > 0$ *and* whenever $B \subset A$ and $B \in \mathscr{T}$, then $P(B) = 0$ or

$P(B) = P(A)$. Ω admits the decomposition

(2) $$\Omega = F + \sum_{r \in I} A_r \qquad F, A_r \in \mathscr{T},$$

into a fully nonatomic set F and a disjoint sum of atoms A_r. The index set I in (2) may be empty, finite or countable. This partition is unique modulo P-null sets, and should be kept fixed in the development which follows. $\mathscr{T}$ is *atomic* iff F may be taken as $\emptyset$, in which case it is *finite* when $|I| < \infty$ and *trivial* when $|I| = 1$. Clearly all of these notions depend implicitly on P. We use the symbol $\cong$ for equality mod P-null sets among events, and equality of P-completions among σ-algebras.

Three simple examples illustrate various possibilities in (2):

EXAMPLE 1. $S = \mathbb{Z} =$ the integers. $\pi_{ii+1}(n) = \pi_{ii-1}(n) = \frac{1}{2}$ for all $i \in S$, $n \in \mathbb{N}$. The methods of [2] show that $F \cong \emptyset$, $I = \{0, 1\}$, and $A_0 \cong \{\xi_0 \text{ is even}\}$, $A_1 \cong \{\xi_0 \text{ is odd}\}$ if ι gives positive measure to both of these sets. If ξ_0 concentrates on either even or odd integers, then $\mathscr{T}$ is trivial.

EXAMPLE 2. $S = \mathbb{N}$. $\iota_0 = 1$, $\pi_{00}(n) = \pi_{01}(n) = \frac{1}{2}$, $\pi_{ii+1}(n) = 1$ for $i \geqq 1$, $n \in \mathbb{N}$. In this case $F \cong \emptyset$, $I = \mathbb{N}$, and $A_r \cong \{\omega^r\}$, where ω^r is the path such that $\omega_n{}^r = (n - r)^+$.

EXAMPLE 3. $S = \mathbb{N}$. $\iota_0 = 1$, $\pi_{ii}(n) = \pi_{ii+2^n}(n) = \frac{1}{2}$, $i \in S$, $n \in \mathbb{N}$. It is not hard to see that $\xi_n(\omega)$ determines $\xi_0(\omega), \xi_1(\omega), \cdots, \xi_{n-1}(\omega)$ uniquely, so $\mathscr{F}_0{}^n \cong \mathscr{F}_n$. It follows that $\mathscr{T} \cong \mathscr{F}$ and $F \cong \Omega$, i.e., the tail σ-algebra is full and fully nonatomic.

Our objective in this section is a means of determining the atomic structure of $\mathscr{T}$. It will be achieved with the aid of certain "couplings":

PROPOSITION 1. *For each $m \in \mathbb{N}$ there is a conditional probability measure $\tilde{P}^m_{\xi_m(\omega)}$ on $(\tilde{\Omega}, \tilde{\mathscr{F}})$ such that*

(3) $$\tilde{P}^m_{\xi_m}(\cdot, \Omega) = P^m_{\xi_m}(\cdot), \qquad \tilde{P}^m_{\xi_m}(\Omega, \cdot) = P^m(\cdot);$$

(4) $$\tilde{P}^m_{\xi_m}(\tilde{\xi}_n \in D \ \textit{for all} \ n \geqq \tau_D) = 1;$$

(5) $$\tilde{P}^m_{\xi_m}(\tilde{\xi}_n = (j, j)) = p_{\xi_m j}(m, m+n) \wedge p_j(m+n), \qquad j \in S,\ n \in \mathbb{N}.$$

PROOF. To obtain $\tilde{P}^m_{\xi_m(\omega)}$, take $P_\mu = P^m_{\xi_m(\omega)}$ and $P_\nu = P^m$ in the following

MAXIMAL COUPLING LEMMA. *Let π be a Markov transition function on S. Let P_μ, P_ν be two Markov measures on $(\Omega, \mathscr{F})$ with the same transition function π, and initial distributions μ and ν respectively. Then there is a probability measure $\tilde{P}$ on $(\tilde{\Omega}, \tilde{\mathscr{F}})$ such that*

(i) $\tilde{P}(\cdot, \Omega) = P_\mu(\cdot)$, $\tilde{P}(\Omega, \cdot) = P_\nu(\cdot)$;
(ii) $\tilde{P}(\tilde{\xi}_n \in D \ \textit{for all} \ n \geqq \tau_D) = 1$;
(iii) $\tilde{P}(\tilde{\xi}_n = (k, k)) = P_\mu(\xi_n = k) \wedge P_\nu(\xi_n = k)$, $k \in S$, $n \in \mathbb{N}$.

This lemma is an extension of the main result of [7] to nonhomogeneous Markov chains with arbitrary initial distributions. A streamlined proof is outlined in an appendix to this paper.

Now, define

$$\delta_m{}^n(\omega) = \sup_{B\in \mathscr{F}_{m+n}^\infty}\{P(B\mid \mathscr{F}_0{}^m)(\omega) - P(B)\}\,, \qquad \delta_m(\omega) = \lim_{n\to\infty}\delta_m{}^n(\omega)$$

$(\delta_m{}^n \downarrow \delta_m)$. Also set

$$\alpha_m(\omega) = 1 - \delta_m(\omega)\,. \tag{6}$$

The following well-known facts will be needed.

LEMMA 2. *For* $A \in \mathscr{T}$,

$$1_A - P(A) \leqq \liminf_{m\to\infty}\sup_{B\in\mathscr{T}}\{P(B\mid \mathscr{F}_0{}^m) - P(B)\} \leqq \liminf_{m\to\infty}\delta_m \quad \text{a.s.};$$

(Here $1_A(\omega)$ denotes the indicator of A.)

PROOF. The first inequality follows from the martingale convergence theorem; the second is trivial since $\mathscr{T}\subset\mathscr{F}_n{}^\infty$ for every n.

The next three propositions relate the coupling measures $P^m_{\xi_m}$ to the random variables α_m.

PROPOSITION 2.

$$\begin{aligned}\tilde P^m_{\xi_m}(\tau_D<\infty) &= 1 - \lim_{n\to\infty}\tfrac12\textstyle\sum_{j\in S}|p_{\xi_m j}(m, m+n) - p_j(m+n)|\\ &= \alpha_m \quad (\textit{everywhere in } \omega \textit{ for each } m\in\mathbb{N})\,.\end{aligned}$$

PROOF. Fix ω, m, and let $i = \xi_m(\omega)$. If $E\in\mathscr{F}_n{}^\infty$, (4) implies $\{\tau_D > n\}\supset\{(E, E^c)\}$ modulo a $P_i{}^m$-null set. Hence (3) yields $\tilde P_i{}^m(\tau_D>n) \geqq P_i{}^m(E) - P^m(E)$. But using (4) and (5),

$$\begin{aligned}\tilde P_i{}^m(\tau_D > n) = \tilde P_i{}^m(\tilde\xi_n\notin D) &= 1 - \textstyle\sum_{j\in S}(p_{ij}(m, m+n)\wedge p_j(m+n))\\ &= \tfrac12\textstyle\sum_{j\in S}|p_{ij}(m, m+n) - p_j(m+n)|\\ &= P_i{}^m(E_0) - P^m(E_0)\,,\end{aligned} \tag{7}$$

where $E_0 = \{\xi_n\in H^+\}\in\mathscr{F}_n$; here H^+ is a positive Hahn set with respect to the signed measure $p_{i\cdot}(m, m+n) - p_\cdot(m+n)$ on S. This shows that $\tilde P_i{}^m(\tau_D>n) = \sup_{E\in\mathscr{F}_n{}^\infty}\{P_i{}^m(E) - P^m(E)\}$. Now $B\in\mathscr{F}_{m+n}^\infty$ if and only if $\theta^m B\in\mathscr{F}_n{}^\infty$, so by the Markov property,

$$\tilde P_i{}^m(\tau_D>n) = \sup_{B\in\mathscr{F}_{m+n}^\infty}\{P_i{}^m(\theta^m B) - P^m(\theta^m B)\} = \delta_m{}^n(\omega)\,. \tag{8}$$

When $n\to\infty$ the claim follows from (7) and (8).

Let us introduce

$$\begin{aligned}C^1_{m,i} &= \{p_{i\xi_n}(m, m+n) \leqq p_{\xi_n}(m+n)\ \text{ i.o.}\}\,,\\ C^2_{m,i} &= \{p_{i\xi_n}(m, m+n) \geqq p_{\xi_n}(m+n)\ \text{ i.o.}\}\,;\\ \tilde C^1_{m,i} &= \{p_{i\xi_n{}^1}(m, m+n) \leqq p_{\xi_n{}^1}(m+n)\ \text{ i.o.}\}\,,\\ \tilde C^2_{m,i} &= \{p_{i\xi_n{}^2}(m, m+n) \geqq p_{\xi_n{}^2}(m+n)\ \text{ i.o.}\}\,;\end{aligned}$$

$m\in\mathbb{N}$, $i\in S$, where "i.o." abbreviates "for infinitely many n." Note that $C^1_{m,i}, C^2_{m,i}\in\mathscr{T}$; $\tilde C^1_{m,i}, \tilde C^2_{m,i}\in\tilde{\mathscr{F}}$. These events play a central role in the arguments below.

PROPOSITION 3. $\tilde P_i{}^m(\tau_D<\infty) \geqq P_i{}^m(C^1_{m,i}) \vee P(\theta^{-m}C^2_{m,i})$, $i\in S$, $m\in\mathbb{N}$.

PROOF. By properties (4) and (5) of the maximal coupling, we have

$$\begin{aligned}\tilde{P}_i^m(\tilde{C}^1_{m,i}, \tau_D = \infty) &\leqq \textstyle\sum_{n\in\mathbb{N}} \tilde{P}_i^m(p_{i\xi_n^1}(m, m+n) \leqq p_{\xi_n^1}(m+n), \tilde{\xi}_n \notin D) \\ &\leqq \textstyle\sum_{n\in\mathbb{N}} \sum_{j_n\in S: p_{ij_n}(m,m+n)\leqq p_{j_n}(m+n)} p_{ij_n}(m, m+n) \\ &\qquad - (p_{ij_n}(m, m+n) \wedge p_{j_n}(m+n)) \\ &= 0\,.\end{aligned}$$

Thus $\tilde{C}^1_{m,i} \subset \{\tau_D < \infty\}$ modulo a $\tilde{P}_i^m$-null set. Similarly, $\tilde{C}^2_{m,i} \subset \{\tau_D < \infty\}$ $\tilde{P}_i^m$-almost surely. But by (3), $\tilde{P}_i^m(\tilde{C}^1_{m,i}) = P_i^m(C^1_{m,i})$ and $\tilde{P}_i^m(\tilde{C}^2_{m,i}) = P^m(C^2_{m,i}) = P(\theta^{-m}C^2_{m,i})$. This implies the desired result.

PROPOSITION 4. *If A is an atom of $\mathscr{T}$, then*

$$\liminf_{m\to\infty} \tilde{P}^m_{\xi_m(\omega)}(\tau_D < \infty) \geqq P(A) \qquad \text{for almost all} \quad \omega \in A\,.$$

PROOF. For given A, set $\mathbb{N}_0(\omega) = \{m \in \mathbb{N} : P(A \cap \theta^{-m}C^1_{m,i})|_{i=\xi_m(\omega)} = 0\}$, $\mathbb{N}_1(\omega) = \{m \in \mathbb{N} : P(A \cap \theta^{-m}C^1_{m,i})|_{i=\xi_m(\omega)} = P(A)\}$. $\mathbb{N}_0 \cup \mathbb{N}_1 = \mathbb{N}$ since $C^1_{m,i} \in \mathscr{T}$ and A is an atom of $\mathscr{T}$. If $m \in \mathbb{N}_0(\omega)$, $P(A \cap \theta^{-m}C^2_{m,i})|_{i=\xi_m(\omega)} = P(A)$ because $\theta^{-m}C^1_{m,i} \cup \theta^{-m}C^2_{m,i} = \Omega$; hence

$$P(\theta^{-m}C^2_{m,i})|_{i=\xi_m(\omega)} \geqq P(A)\,. \tag{9}$$

If $m \in \mathbb{N}_1(\omega)$, then $P_i^m(\theta^m A \cap C^1_{m,i})|_{i=\xi_m(\omega)} = P(A \cap \theta^{-m}C^1_{m,i} \mid \mathscr{T}_0^{\,m})(\omega)|_{i=\xi_m(\omega)} = P(A \mid \mathscr{T}_0^{\,m})(\omega)$ a.s. By martingale convergence the last term tends to 1 as $m \to \infty$ for almost all $\omega \in A$. We conclude that

$$\lim_{m\to\infty;\, m\in\mathbb{N}_1(\omega)} P_i^m(C^1_{m,i})|_{i=\xi_m(\omega)} = 1 \tag{10}$$
$$\text{for almost all} \quad \omega \in A \quad \text{such that} \quad \mathbb{N}_1(\omega) \quad \text{is infinite.}$$

Together, (9) and (10) show

$$\liminf_{m\to\infty} P_i^m(C^1_{m,i}) \vee P(\theta^{-m}C^2_{m,i})|_{i=\xi_m(\omega)} \geqq P(A) \qquad \text{a.s. on} \quad A\,.$$

An application of Proposition 3 completes the proof.

We now state and prove the structure theorem for $\mathscr{T}$.

THEOREM 1. *With α_m given by* (6), *$\alpha(\omega) = \lim_{m\to\infty} \alpha_m(\omega)$ exists* a.s., *and if* (2) *is the atomic decomposition of $\mathscr{T}$, then*

$$\begin{aligned}\alpha(\omega) &= 0 \qquad &&\text{a.s. } on \quad F\,, \\ &= P(A_r) \qquad &&\text{a.s. } on \quad A_r\,.\end{aligned}$$

PROOF. Write $\bar{\alpha}(\omega) = \limsup_{m\to\infty} \alpha_m(\omega)$, $E_k = \{\omega : \bar{\alpha}(\omega) \geqq 1/k\}$ $(k = 1, 2, \cdots)$. Clearly $E_k \in \mathscr{T}$. Suppose $B \in \mathscr{T}$, $B \subset E_k$ and $P(B) > 0$. Lemma 2 implies that $1_B - P(B) \leqq 1 - 1/k$ a.s. on B, so we clearly have $P(B) \geqq 1/k$. Thus E_k contains only atoms, at most k in number. Hence $E = \{\omega : \bar{\alpha}(\omega) > 0\} = \bigcup_{k=1}^{\infty} E_k$ contains only atoms, which shows that $\lim_{m\to\infty} \alpha(\omega) = \bar{\alpha}(\omega) = 0$ a.s. on F. Now fix A_r, an atom of $\mathscr{T}$. By Lemma 2,

$$\bar{\alpha} \leqq 1 - (1_{A_r} - P(A_r)) = P(A_r) \qquad \text{a.s. on} \quad A_r\,. \tag{11}$$

Putting $\underline{\alpha}(\omega) = \liminf_{m\to\infty} \alpha_m(\omega)$, Propositions 2 and 4 yield

$$\underline{\alpha} = \liminf_{m\to\infty} \tilde{P}^m_{\xi_m}(\tau_D < \infty) \geqq P(A_r) \qquad \text{a.s. on} \quad A_r\,. \tag{12}$$

In combination, (11) and (12) yield the theorem.

As an immediate consequence we obtain

COROLLARY 1. *$\mathscr{T}$ is*

(a) *trivial* iff $\alpha(\omega) = 1$ a.s.;
(b) *finite* iff $\alpha(\omega) \geqq \lambda$ a.s. *for some* $\lambda > 0$;
(c) *atomic* iff $\alpha(\omega) > 0$ a.s.;
(d) *fully nonatomic* iff $\alpha(\omega) = 0$ a.s.

REMARK 1 (added in revision). H. Cohn [6] has recently obtained a proof of Theorem 1 which does not rely on coupling methods. Note that Propositions 2 through 4 are only used to obtain the inequality (12); Cohn's alternative derivation of (12) employs various martingale arguments.

4. Applications. We next prove a theorem relating various expressions of partial loss of memory for a Markov chain. These conditions will be indexed by a parameter λ, $0 < \lambda \leqq 1$; intuitively: the larger λ, the more forgetful is (ξ_n). Thus the leading case, $\lambda = 1$, expresses total loss of memory of the chain's past history.

Say that *a* $0 - \lambda$ *law holds for* (ξ_n) if

$$P(B) = 0 \qquad \textit{or} \qquad P(B) \geqq \lambda \quad \text{whenever} \quad B \in \mathscr{T} \tag{13}$$

$(0 < \lambda \leqq 1)$. This condition clearly implies that $\mathscr{T}$ is finite, with at most $[1/\lambda]$ atoms, while any Markov chain with finite tail σ-algebra satisfies (13) with $\lambda = \min_{r\in I} P(A_r)$.

THEOREM 2. *Let (ξ_n) be a Markov chain on $(\Omega, \mathscr{F}, P)$, $P = P(\iota, \pi)$. Consider the following four conditions for fixed $\lambda \in (0, 1]$:*

(i) $P_i^m(C^1_{m,i}) \vee P(\theta^{-m}C^2_{m,i})|_{i=\xi_m(\omega)} \geqq \lambda$ i.o. a.s.;
(ii) $\alpha_m \geqq \lambda$ i.o. a.s.;
(iii) $\lim_{n\to\infty} \frac{1}{2} \sum_{j\in S} |p_{\xi_m j}(m, m+n) - p_j(m+n)| \leqq 1 - \lambda$ i.o. a.s.;
(iv) *If* $h : S \times \mathbb{N} \to \mathbb{R}$, $0 \leqq h \leqq 1$, *satisfies* $h(i, 0) = \rho_i$ $(i \in S)$, *and* $h(i, n) = \sum_{j\in S} \pi_{ij}(n)h(j, n+1)$ $(i \in S, n \in \mathbb{N})$, *then*

$$|h(\xi_m, m) - \textstyle\sum_{j\in S} \iota_j \rho_j| \leqq 1 - \lambda \quad \text{i.o.} \quad \text{a.s.}$$

(In each case "i.o." abbreviates "for infinitely many m.") *If any of* (i)—(iv) *holds, then (ξ_n) satisfies a 0–λ law. More generally, if $(\sigma_m)_{m\in\mathbb{N}}$ is a strictly increasing sequence of* a.s. *finite stopping times, and if any of* (i)—(iv) *holds with m replaced by σ_m, then* (13) *follows. Conversely, if a 0–λ law holds for (ξ_n), then all of* (i)—(iv) *are satisfied with* "i.o." *replaced by "for every m."*

PROOF. By Propositions 2 and 3, (i)$\Rightarrow$(ii)$\Leftrightarrow$(iii). To see that (iii) implies (iv), let $g(i, n) = 2h(i, n) - 1$, so $|g| \leqq 1$. For any m and n, $|g(\xi_m, m) - \sum_{j \in S} \iota_j(2\rho_j - 1)| = \sum_{j \in S} [p_{\xi_m j}(m, m+n) - p_j(m+n)]g(j, m+n) \leqq \sum_{j \in S} |p_{\xi_m j}(m, m+n) - p_j(m+n)|$. Let $n \to \infty$ and use (iii) to obtain an equivalent form of (iv). Next, fix $B \in \mathscr{T}$ and set $h(i, n) = P(B \mid \xi_n = i)$. Then h satisfies the hypotheses of (iv), and the conclusion becomes $|P(B \mid \mathscr{F}_0^m) - P(B)| \leqq 1 - \lambda$ i.o. a.s. Now (13) follows from Lemma 2. We have shown that any of (i)—(iv) implies (13); the proofs with m replaced by σ_m are analogous. Let (i′)—(iv′) denote the conditions of the theorem when "i.o." is replaced by "for every m." In the same way, (i′) $\Rightarrow$ (ii′) $\Leftrightarrow$ (iii′) $\Rightarrow$ (iv′). Thus it remains only to show (13) $\Rightarrow$ (i′). For each $m \in \mathbb{N}$ and $i \in S$ such that $p_i(m) > 0$, either $P_i^m(C^1_{m,i}) = 1$ or $P_i^m(C^2_{m,i}) > 0$. In the latter case $P(\theta^{-m}C^2_{m,i}) > 0$, and since $\theta^{-m}C^2_{m,i} \in \mathscr{T}$, (13) implies that $P(\theta^{-m}C^2_{m,i}) \geqq \lambda$. As $\{\omega : p_{\xi_m(\omega)} > 0\}$ is a P-full set, we have derived (i′).

REMARK 2. Condition (i) of Theorem 2 generalizes the one in Theorem 5 of [7], and says roughly that (ξ_n) should often visit states which are more likely to be visited from other starting positions. It does not seem straightforward to prove that (i) is equivalent to (ii)—(iv) and (13) without coupling; the equivalence of all conditions except (i) can be obtained by martingale methods.

We mention here a well-known curiosity, which seems impossible to prove directly:

PROPOSITION 5. *If* any *of* (i)—(iv) *holds with* $\lambda > \frac{1}{2}$, *then* all *of* (i′)—(iv′) *hold with* $\lambda = 1$.

PROOF. It suffices to show that (13) for $\lambda > \frac{1}{2}$ implies (13) with $\lambda = 1$. This is clear since $B, B^c \in \mathscr{T}$ cannot satisfy $P(B) > \frac{1}{2}$ *and* $P(B^c) > \frac{1}{2}$.

The next estimate leads to simple sufficient conditions for trivial, finite or atomic $\mathscr{T}$.

LEMMA 3. $\alpha_m(\omega) \geqq p_{\xi_m}(m)$.

PROOF. Using Proposition 2 and (4),

$$\alpha_m(\omega) = \tilde{P}^m_{\xi_m}(\tau_D < \infty) \geqq \tilde{P}^m_{\xi_m}(\tilde{\xi}_0 = (\xi_m, \xi_m)) = p_{\xi_m}(m) .$$

PROPOSITION 6. *If there is a sequence* $(\sigma_m)_{m \in \mathbb{N}}$ *of strictly increasing* a.s. *finite stopping times such that* $P(p_{\xi_{\sigma_m}}(\sigma_m) \geqq \lambda$ i.o.$) = 1$ *for a fixed* λ, $0 < \lambda \leqq 1$, *then a* 0–λ *law holds for* (ξ_n). $\mathscr{T}$ *is atomic if* $P(p_{\xi_{\sigma_m}}(\sigma_m) \geqq \lambda$ i.o. *for some* $\lambda > 0) = 1$. (Here "i.o." abbreviates "for infinitely many m.")

PROOF. The first claim is immediate from Lemma 3 and the stopping time version of (iii) in Theorem 2. Also, by Theorem 1 and Lemma 3, $P(\alpha > 0) = P(\limsup_{m \to \infty} \alpha_{\sigma_m} \geqq \lambda$ for some $\lambda > 0) \geqq P(p_{\xi_{\sigma_m}}(\sigma_m) \geqq \lambda$ i.o. for some $\lambda > 0)$. The hypothesis and Corollary 1 imply the second claim.

As corollaries to this last proposition, we derive results explicit and implicit in [4]:

COROLLARY 2. *If* $\limsup_{m\to\infty} P(p_{\xi_m}(m) \geqq \lambda) = 1$ *for fixed* $\lambda > 0$, *then a* 0–λ *law holds for* (ξ_n). *If* $\lim_{\lambda\to 0} \limsup_{m\to\infty} P(p_{\xi_m}(m) \geqq \lambda) = 1$, *then* $\mathscr{T}$ *is atomic.*

PROOF. The expressions in the hypotheses are majorized by $P(p_{\xi_m}(m) \geqq \lambda$ i.o.) and $P(p_{\xi_m}(m) \geqq \lambda$ i.o. for some $\lambda > 0)$ respectively. Apply Proposition 6 with $\sigma_m(\omega) \equiv m$.

COROLLARY 3. *For each* n, *let* $E_n \subset S$, *and suppose that* $P(\xi_n \in E_n$ i.o.$) = 1$. *If* $\liminf_{n\to\infty} \inf_{k\in E_n} p_k(n) \geqq \lambda > 0$, *then a* 0–$\lambda$ *law holds for* (ξ_n).

(Note: $E_n = \varnothing$ for some n is allowed; the empty infimum is $+\infty$.)

PROOF. Apply Proposition 6 with $\sigma_m =$ the mth time that $\xi_n \in E_n$.

REMARK 3. Using Theorem 1 and Corollary 2 one can derive the Cohn–Senchenko theorem [3, 14] mentioned in the introduction. But a much shorter and more elegant proof was obtained recently by Cohn [5].

5. Appendix: Proof of the maximal coupling lemma (a sketch). Using methods from [11], Pitman has produced a much more palatable construction of the maximal coupling than the one in [7]. We outline his approach.

Denote $p_{\mu i}^{(n)} = P_\mu(\xi_n = i)$, $p_{\nu i}^{(n)} = P_\nu(\xi_n = i)$. Start by defining random variables N and Y to have joint distribution $\Pr(N \leqq n,\ Y = k) = p_{\mu k}^{(n)} \wedge p_{\nu k}^{(n)}$, $n \in \mathbb{N}$, $k \in S$, $\Pr(N = \infty) = 1 - \Pr(N < \infty)$, on some probability space with measure Pr. Enlarging the space if necessary, form a stochastic process $\tilde{X}_n = (X_n^1, X_n^2)$ taking on values in $S \times S$, and satisfying the following key properties:

(a) $\Pr(X_N^1 = X_N^2 = Y) = 1$, $N < \infty$;

(b) If $N \leqq n < \infty$, $X_n^1 = X_n^2 =$ a single Markov chain starting in Y at time N, and with transition from time n to time $n+1$ governed by $\pi(n)$;

(c) For $N < \infty$, the reverse processes $(X_N^1, X_{N-1}^1, \cdots, X_0^1)$ and $(X_N^2, X_{N-1}^2, \cdots, X_0^2)$ are independent Markov chains starting in Y at time N, with transition probabilities from j at time n to i at time $n-1$ given by

$$c^+(p_{\mu i}^{(n-1)} - p_{\nu i}^{(n-1)})^+ \pi_{ij}(n) \qquad \text{and} \qquad c^-(p_{\mu i}^{(n-1)} - p_{\nu i}^{(n-1)})^- \pi_{ij}(n)$$

respectively, where c^+ and c^- are the appropriate normalizing constants;

(d) The pre-N and post-N processes are independent;

(e) Conditional on $N = \infty$, the processes X_n^1 and X_n^2 are independent.

Properties (a)—(e) uniquely determine the probability law for the stochastic process $\tilde{X}_n$. Its coordinate representation is the desired $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}, (\tilde{\xi}_n))$. For details, the reader is referred to [13].

Acknowledgments. I would like to thank H. Kesten and S. Orey, who drew my attention to [4]. Thanks also to H. Cohn and M. Iosifescu for their helpful suggestions. Finally, I am grateful to W. Pruitt for pointing out some errors in the original manuscript.

REFERENCES

[1] BARTFAY, P. and REVESZ, P. (1967). On a zero-one law. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **7** 43-47.

[2] 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.

[3] COHN, H. (1970). On the tail σ-algebra of the finite inhomogeneous Markov chains. *Ann. Math. Statist.* **41** 2175-2176.

[4] COHN, H. (1974). On the tail events of a Markov chain. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **29** 65-72.

[5] COHN, H. (1974). A ratio limit theorem for the finite nonhomogeneous Markov chains. *Israel J. Math.* **19** 329-344.

[6] COHN, H. (1976). On the tail σ-field of the countable Markov chains. To appear.

[7] GRIFFEATH, D. (1975). A maximal coupling for Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **31** 95-106.

[8] GRIFFEATH, D. (1975). Uniform coupling of nonhomogeneous Markov chains. *J. Appl. Probability* **12** 753-762.

[9] IOSIFESCU, M. (1972). On finite tail σ-algebras. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **24** 159-166.

[10] JAMISON, B. and OREY, S. (1967). Markov chains recurrent in the sense of Harris. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **8** 41-48.

[11] PITMAN, J. W. (1974). Stopping time identities and limit theorems for Markov chains. Ph.D. thesis, Univ. of Sheffield.

[12] PITMAN, J. W. (1974). Uniform rates of convergence for Markov chain transition probabilities. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **29** 193-227.

[13] PITMAN, J. W. (1976). On coupling of Markov chains. To appear.

[14] SENCHENKO, D. V. (1972). The final σ-algebra of an inhomogeneous Markov chain with a finite number of states. *Math. Notes* **12** 610-613.

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853