The Annals of Probability
1991, Vol. 19, No. 3, 907–928

SPECIAL INVITED PAPER

MARKOV CHAINS WITH STOCHASTICALLY STATIONARY TRANSITION PROBABILITIES[1]

BY STEVEN OREY[2]

University of Minnesota

Markov chains on a countable state space are studied under the assumption that the transition probabilities $(P_n(x, y))$ constitute a stationary stochastic process. An introductory section exposing some basic results of Nawrotzki and Cogburn is followed by four sections of new results.

0. Introduction. Let X be a finite or denumerably infinite set, and consider a stationary stochastic sequence (P_n), where each P_n is a stochastic matrix $(P_n(x, y), x \in X, y \in X)$. One may consider P_n as the time n-to-$(n + 1)$ transition probabilities of a Markov chain moving on X. The study of such Markov chains with random transition probabilities has been pursued for some time, often for special cases or examples. Nawrotzki [12, 13, 14] introduced a general theory. Then Cogburn [3, 5] developed such a theory in a wider context making use of more powerful tools. It is Cogburn's theory that will be our setting. For extensive references to previous work, we refer to the discussion and bibliography in [3, 4, 5, 12, 13, 14]. There is of course no implication that the general theory subsumes all the special investigations. In this article we attempt no historical survey, nor a discussion of the applied problems which have motivated much of the literature. Our concern is primarily with questions related to the ergodic theory of the Markov chains. For example, central limit problems, treated in [6], will not be discussed.

Section 1 is intended as an introduction to the theory. There our notation will be introduced. For ease of comparison we have followed the notational setup of [3, 5], but we do not always use the same symbols to denote the same concepts as in [3, 5]. The work of [3, 5] draws heavily on the Hopf theory of L_1-Markov chains. In Section 1 we assemble the necessary facts from this theory and then outline some of the major concepts and results from Cogburn's theory. For detailed proofs of some of the results, we refer of course to the original papers [3, 5]. It is our intention, however, that except for a few facts (which the reader can take on faith, try to prove him- or herself or look up in

---

Received February 1990; revised July 1990.

[1]Supported in part by NSF Grant MCS-83-01080.

[2]Steven Orey died March 8, 1991. His friends and colleagues wish to express their deep sorrow over this untimely death.

AMS 1980 *subject classifications*. Primary 60G10; secondary 60J10.

Key words and phrases. Markov chains with random transition probabilities, products of random stochastic matrices.

the provided references) the present paper is self-contained. Thus the aim of Section 1 is expository, with perhaps some minor novelties.

Sections 2–5 contain new contributions to the theory. Each section has its own introduction and motivation. Though the sections deal with distinct problems, they are connected by a common theme.

Several open problems are mentioned. These are given as Problems 1.3.1, 1.3.3, 1.3.6, 5.1 and 5.7.

1. Background. We shall use the same setup as Cogburn [3, 4, 5] which extends that of Nawrotzki [12, 13, 14]. X is to be a finite or denumerable set, $\mathscr{A}$ the σ-field of all subsets of X. Let $(\Theta, \mathscr{B}_0)$ be a measurable space, $\Omega = \Theta^{\mathbb{Z}}$, $\mathbb{Z}$ being the integers. Let $\Theta_n: \Omega \to \Theta$ be the nth coordinate function, that is, $\Theta_n(\omega) = \omega_n$. Let $\mathscr{B}_k^l$ be the σ-field generated by the Θ_n, $k - 1 < n < l + 1$, where $-\infty \le k \le l \le \infty$. Let $\mathscr{B} = \mathscr{B}_{-\infty}^{\infty}$. The *shift* T: $\Omega \to \Omega$ is defined by $(T\omega)_n = \omega_{n+1}$, $n \in \mathbb{Z}$. Let $(\Omega, \mathscr{B}, \pi)$ be a probability space such that $\pi = \pi \circ T^{-1}$. Let P be a measurable map from Θ into the space of stochastic matrices, the latter equipped with the obvious Borel field. Then (Θ_n) is a stationary process, and so is $(P \circ \Theta_n)$. (Thus $P \circ \Theta_n$ takes the place of P_n in the Introduction. For our present purposes one could take Θ to be the space of stochastic matrices and P the identity function.) Define

$$P(\omega_m \cdots \omega_n) = P(\omega_m)P(\omega_{m+1}) \cdots P(\omega_n) \quad \text{for } -\infty < m \le n < \infty,$$

and write $P(\omega_m \cdots \omega_n; x, y) := P(\omega_m \cdots \omega_n)(x, y)$, $x, y \in X$. For a fixed sample point ω, $P(\omega_n)$ can be considered as the time n-to-$(n+1)$ transition probability of a corresponding *sample Markov chain* X_n moving on the space X. If this chain is started at time 0 at position x, then the probability that the chain visits $x_0, x_1, \dots, x_n$ at times $0, 1, \dots, n$ is given by

$$\begin{aligned} &P_x^{\omega}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n] \\ &\quad = I_{\{x\}}(x_0)P(\omega_0; x_0, x_1)P(\omega_1; x_1, x_2) \cdots P(\omega_{n-1}; x_{n-1}, x_n), \end{aligned} \tag{1.1}$$

where I_A is used throughout as the indicator function of the set A. Note that (1.1) determines the measure P_x^{ω} on $X^{\mathbb{Z}_+}$, with $\mathbb{Z}_+$ the nonnegative integers. A typical question dealt with by the theory is the behavior of

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P_x^{\omega}[X_k = y]$$

as a function of (x, ω). Cogburn had the good idea to make the connection with the well-developed theory of Hopf Markov chains. We now review some results from this theory.

1.1. *Hopf Markov chains.* This theory originated with Hopf [10] and was further developed subsequently. For present purposes the work of Foguel [7] is a particularly useful reference. We summarize some key concepts and results for later use.

A *Hopf Markov chain* is a quadruple $\Sigma = (E, \mathscr{E}, P, \Pi)$, where $(E, \mathscr{E}, \Pi)$ is a measure space with positive σ-finite measure Π and P is a positive contraction on $L_1(\Pi)$. The result of P acting on φ in $L_1(\Pi)$ will be written φP. The dual operator for P is a positive contraction on $L_\infty(\Pi)$; its action will be written as Pf. When not otherwise specified, "a.e." will mean "almost everywhere with respect to the reference measure Π"; furthermore since L_1 and L_∞ functions are defined only as equivalence classes, we will sometimes write $f = g$ or $A \subseteq B$ with the understanding that these relations hold a.e.

For $\varphi \in L_1(\Pi)$, $f \in L_\infty(\Pi)$ let $\langle \varphi, f \rangle := \int \varphi \cdot f \, d\Pi$. Then

$$\langle \varphi P, f \rangle = \langle \varphi, Pf \rangle. \tag{1.2}$$

For $A \in \mathscr{E}$, write $P(z, A) := PI_A(z)$, $z \in E$. Every $\varphi \in L_1(\Pi)$ gives rise to a signed measure Φ on $\mathscr{E}$ via

$$\Phi(B) = \int_B \varphi \, d\Pi, \qquad B \in \mathscr{E}, \tag{1.3}$$

and defining

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

one finds

$$\Phi P(B) = \int_B \psi \, d\Pi, \qquad B \in \mathscr{E}, \quad \text{with } \psi = \varphi P. \tag{1.4}$$

We note that $P(z, A)$ defines a sub-Markovian "transition probability kernel"; the quotation marks are a reminder that all the usual relations hold, but only up to Π-null sets; see [7], page 2.

The conservative part of E is defined by

$$C := \left\{ z \colon \sum_{k=0}^{\infty} \psi P^k(z) = \infty \right\}, \tag{1.5}$$

where $\psi \in L_1(\Pi)$, $\psi > 0$; the definition does not depend on the choice of ψ (see [7], page 11). Furthermore ([7], page 11) if $\psi \in L_1(\Pi)$, $\psi \geq 0$, then

$$\begin{aligned} &\sum_{k=0}^{\infty} \psi P^k(x) < \infty, \qquad x \in E \setminus C, \\ &\sum_{k=0}^{\infty} \psi P^k(x) = 0 \text{ or } \infty, \qquad x \in C. \end{aligned} \tag{1.6}$$

It is also possible to characterize C in terms of the dual operator; see [11]. The σ-ring of *invariant sets* is defined by

$$\mathscr{E}_i = \{ A \in \mathscr{E} \colon PI_A = I_A \} \qquad ([7], \text{page } 8) \tag{1.7}$$

and the larger σ-ring of *deterministic sets* is defined by

$$\mathscr{E}_d := \left\{ A \in \mathscr{E} \colon \text{there exists } B_n \in \mathscr{E} \text{ with } P^n I_A = I_{B_n}, \, n = 1, 2, \dots \right\} \tag{1.8}$$

([7], page 7). A set $A \in \mathscr{E}$ is *closed* if $\Pi(A) > 0$ and $I_A \leq PI_A$. If a closed set

A satisfies $B \subseteq A$ and B closed implies $B = A$ (up to Π-null sets of course), then A is a *minimal closed set*. An important fact is

(1.9) $0 \le f,\ f \in L_\infty(\Pi)$ and $Pf \ge f$ on C imply $Pf = f$ on C ([7], page 15).

Hence if F is a closed subset of C we have, since C itself is closed (when not void) that $C \setminus F$ is also closed or void. Define now the class of invariant probabilities absolutely continuous with respect to Π by

$$\mathscr{M}_i := \{\Phi\colon \Phi \text{ is a probability on } \mathscr{E},\ \Phi \ll \Pi,\ \Phi P = \Phi\}. \tag{1.10}$$

If $\Phi \in \mathscr{M}_i$ it has a Radon–Nikodym derivative φ with respect to Π; we call φ the *density* of Φ. The class of φ obtained in this way is the class of *invariant densities*; it coincides with

$$\mathscr{D}_i := \left\{\varphi \in L_1(\Pi)\colon \varphi P = \varphi,\ \varphi \ge 0,\ \int \varphi\, d\Pi = 1\right\}. \tag{1.11}$$

From [7], page 46, we cite the important result

$$\mathscr{M}_i = \varnothing \Leftrightarrow \text{there exist } A_m \uparrow E \text{ with } \left\| \lim_n \frac{1}{n} \sum_{j=0}^{n-1} P^j I_{A_m} \right\|_\infty = 0 \quad \text{for each } m. \tag{1.12}$$

Furthermore ([7], page 45), if $\mathscr{M}_i \ne \varnothing$ then there exists a $\Phi \in \mathscr{M}_i$ such that $\Psi \ll \Phi$ for every $\Psi \in \mathscr{M}_i$. Though Φ is not unique (unless $\mathscr{M}_i = \{\Phi\}$), it is clear that the set $\{\varphi > 0\}$, with φ the density of Φ, is uniquely determined; we now define

$$M := \{\varphi > 0\} \tag{1.13}$$

with $M := \varnothing$ if $\mathscr{M}_i = \varnothing$. Then $M \subseteq C$; see [7], Theorem E, page 45. (In the case of a Markov chain with countable state space and constant transition probabilities, C will be the set of recurrent states, M the set of positively recurrent states.)

All the above definitions depend on the Markov chain Σ. When it seems necessary to indicate this dependence, we will write $\mathscr{E}_i(\Sigma)$, $\mathscr{M}_i(\Sigma)$, etc. Starting with a Markov chain Σ as above, one sometimes wishes to look at a smaller Markov chain obtained by a process of restriction. If $\mathscr{E}'$ is a σ-sub-ring of $\mathscr{E}$ with the property that $A \in \mathscr{E}'$ implies $P(\cdot, A)$ is $\mathscr{E}'$-measurable, then we may consider $\Sigma' = (E, \mathscr{E}', P', \Pi')$, where $\Pi' = \Pi|\mathscr{E}$ and $P'f = Pf$ for $f \in L_\infty(\Pi')$; then for $\varphi \in L_1(\Pi')$, $\varphi P'$ must be given by the conditional expectation $\Pi\{\varphi P|\mathscr{E}'\}$. The most common situation is $\mathscr{E}' = \mathscr{E} \cap F$, where F is a closed set. In that case we shall define $\Sigma_F := (F, \mathscr{E} \cap F, P, \Pi)$ with the understanding that Π is restricted to $\mathscr{E} \cap F$ and P is the corresponding P'. From (1.9) and the succeeding remark, one sees that $\mathscr{E}_i(\Sigma_C)$ [or $\mathscr{E}_i(\Sigma_M)$] coincide with the class of all closed subsets of C (or M) supplemented by the empty set.

The Chacon–Ornstein ergodic theorem is given in a somewhat generalized form in Revuz [15], Chapter 4. It states that for $\psi, \varphi \in L_1(\Pi)$ with $\varphi \geq 0$:

$$(1.14) \qquad \lim_n \frac{\sum_{m=0}^{n} \psi P^m}{\sum_{m=0}^{n} \varphi P^n} =: L(\psi, \varphi)$$

exists and is finite Π-a.e. on $\{\sum_{m=0}^{\infty} \varphi P_n > 0\}$,

$$(1.15) \qquad L(\Psi, \varphi) = \frac{\sum_{m=0}^{\infty} \psi P^m}{\sum_{m=0}^{\infty} \varphi P^m} \quad \text{on } (E \setminus C) \cap \left\{ \sum_{m=0}^{\infty} \varphi P^m > 0 \right\},$$

$$(1.16) \qquad L(\psi, \varphi) = \frac{\Pi\{\psi | \mathscr{E}_i(\Sigma_C)\}}{\Pi\{\varphi | \mathscr{E}_i(\Sigma_C)\}} \quad \text{on } C \cap \left\{ \sum_{m=0}^{\infty} \varphi P^m > 0 \right\}$$

if $\{|\psi| > 0\} \subseteq C$, $\{\varphi > 0\} \subseteq C$, where the expressions on the right side of (1.16) denote the obvious conditional expectations. If the proviso in (1.16) is violated the right side must be replaced by a more complicated expression; see [15].

Another important ergodic theorem gives us the following information: Let $\Phi \in \mathscr{M}_i$, with density φ, $F = \{\varphi > 0\}$. For $f \in L_1(\Phi)$,

$$(1.17) \qquad \lim_n \frac{1}{n} \sum_{k=0}^{n-1} P^k f = \Phi\{f | \mathscr{E}_i(\Sigma_F)\}, \qquad \Phi\text{-a.e. and in } L_1(\Phi).$$

This is a consequence of [7], Theorem A, page 80.

1.2. *The work of Cogburn.* In order to study the sample Markov chains introduced earlier, it is natural to introduce a skew Markov chain with constant transition probabilities on a bigger state space. This was done by Nawrotzki [12, 13, 14], but not having observed the connection with Hopf Markov chains, he was forced to obtain his results without the benefit of a well-developed theory and had to impose stringent conditions: In his work Θ is always countable and sometimes X is finite. In this section we therefore follow Cogburn [3, 5]. We outline some of the concepts and results from [3, 5] with minor variations and additions with a view to our applications in the subsequent sections. Of course, there are many results in Cogburn's work which we do not mention or give only partially.

Let $E = X \times \Omega$, $\mathscr{E} = \mathscr{A} \times \mathscr{B}$, also $\mathscr{E}_k^l = \mathscr{A} \times \mathscr{B}_k^l$. For $F \subseteq E$ let

$$(F)_x = \{\omega \in \Omega: (x, \omega) \in F\}, \qquad (F)^\omega = \{x \in X: (x, \omega) \in F\}.$$

Let κ be counting measure on X, $\Pi = \kappa \times \pi$ product measure on $\mathscr{E}$. Next $P(z, F)$ is to be a transition probability kernel on $E \times \mathscr{E}$. It is determined by requiring

$$P((x, \omega), \{y\} \times B) = P(\omega_0; x, y) I_B(T\omega), \qquad x, y \in X,\ \omega \in \Omega,\ B \in \mathscr{B}.$$

Then P is Π-nonsingular, that is, $\Pi(F) = 0$ implies $P(z, F) = 0$ for Π-a.e. z. Now

$$Pf(z) = \int P(z, du) f(u) \tag{1.18}$$

will be a positive contraction on $L_\infty(\Pi)$: If $f_1 = f_2$ as elements of $L_\infty(\Pi)$, then $Pf_1 = Pf_2$ in the same sense. Indeed,

$$\Sigma := (E, \mathscr{E}, P, \Pi)$$

is a Hopf Markov chain. From (1.18) and (1.2) one obtains

$$\begin{aligned} &P^n f(x, \omega) \\ &\quad = \sum_y P(\omega_0 \cdots \omega_{n-1}; x, y) f(y, T^n \omega), \qquad f \in L_\infty(\Pi), n \geq 1. \end{aligned} \tag{1.19}$$

$$\begin{aligned} &\varphi P^n(x, \omega) \\ &\quad = \sum_y \varphi(y, T^{-n}\omega) P(\omega_{-n} \cdots \omega_{-1}; y, x), \qquad \varphi \in L_1(\Pi), n \geq 1. \end{aligned} \tag{1.20}$$

The Markov chain moving on E according to P will be called the *skew product Markov chain*. For each initial point $(x, \omega) \in E$, there is then determined a measure $P_{(x,\omega)}$ on the product space $E^{\mathbb{Z}_+} := X^{\mathbb{Z}_+} \times \Omega^{\mathbb{Z}_+}$. Denoting by X_n or Θ_n the nth coordinate function on $X^{\mathbb{Z}_+}$ or $\Omega^{\mathbb{Z}_+}$ respectively, our Markov process is $((X_n, \Theta_n), n = 0, 1, \dots)$. As usual the initial point x may be replaced by an initial probability distribution Φ on E giving us P_Φ. Then P_Φ will be stationary if and only if $\Phi = \Phi P$. Denoting the marginal of P_Φ on $X^{\mathbb{Z}_+}$ by P'_Φ, we have $P'_{(x,\omega)} = P^\omega_x$; see (1.1).

As mentioned earlier C can be thought of as the recurrent part of the Markov chain. (See, however, [2] for another notion of recurrence and [2], the example on page 911, for further insight.) From the definition of the conservative set C, one obtains

$$C = \left\{ (x, \omega) \colon \sum_{k=1}^{\infty} P(\omega_{-k} \cdots \omega_{-1}; x, x) = \infty \right\}. \tag{1.21}$$

Indeed, define $\chi_x := I_{\{x\} \times \Omega}$. By (1.20), $P(\omega_{-k} \cdots \omega_{-1}; x, y) = \chi_x P^k(y, \omega)$ so that by (1.6) for each $x \in X$,

$$\left\{ (x, \omega) \colon \sum_{k=0}^{\infty} \chi_x P^k(x, \omega) = \infty \right\} \subseteq C$$

and also, since $\chi_x P^0(x, \omega) = 1$,

$$\left\{ (x, \omega) \colon \sum_{k=0}^{\infty} \chi_x P^k(x, \omega) < \infty \right\} \subseteq X \setminus C$$

and (1.21) follows. See also [2], Theorem 4.1.

There are simple conditions for $\mathscr{M}_i \neq \varnothing$. In fact,

Let F be closed. Then

$$\mathscr{M}_i(\Sigma_F) \neq \varnothing \tag{1.22}$$
$$\Leftrightarrow \Pi\left\{(x,\omega) \in F: \sup_y \limsup_n \frac{1}{n} \sum_{k=0}^{n-1} P^k(\omega_0 \cdots \omega_k; x, y) > 0\right\} > 0.$$

This is essentially [3], Theorem 3.1. The implication $\Leftarrow$ follows easily from (1.12); the implication $\Rightarrow$ follows from (1.17) using this time $P^k(\omega_0 \cdots \omega_k; x, y) = P^k\chi_y(x,\omega)$.

To obtain a characterization of M, note first that if $\Phi \in \mathscr{M}_i$ has density φ, then $F = \{\varphi > 0\}$ is closed and

$$(F)_x \subseteq \left\{\lim_n \frac{1}{n} \sum_{k=0}^{n-1} P(\omega_0 \cdots \omega_k; x, x) > 0\right\}.$$

To justify the inclusion, note that on applying the ergodic theorem (1.17) to $\chi_x := I_{\{x\}\times\Omega}$,

$$\frac{1}{n} \sum_{k=0}^{n-1} P(\omega_0 \cdots \omega_k; z, x) \to \Phi\{\chi_x | \mathscr{E}_i(\Sigma_C)\}(z,\omega), \qquad \Pi\text{-a.e. } (z,\omega) \in F$$

(since on F, Π and Φ are mutually absolutely continuous). Let A be the set where $\Phi(\chi_x | \mathscr{E}_i(\Sigma_C))$ vanishes. Then

$$\int_{(A)_x} \varphi(x,\omega)\pi(d\omega) = \Phi(\{x\} \times (A)_x) = \int_A \chi_x \, d\Phi = \int_A \Phi\{\chi_x | \mathscr{E}_i(\Sigma_C)\} \, d\Phi = 0.$$

Since $F = \{\varphi > 0\}$, $(A)_x \cap (F)_x = \varnothing$. This is the desired inclusion.

A characterization of M is given by

$$M = \left\{(x,\omega): \sum_{k=0}^{\infty} P(\omega_{-k} \cdots \omega_{-1}; x, x) = \infty,\right. \tag{1.23}$$
$$\left. \limsup_n \frac{1}{n} \sum_{k=0}^{n-1} P(\omega_0 \cdots \omega_k; x, x) > 0\right\}.$$

To see this, denote the right side of (1.23) by M'. The first condition in the definition of M' just ensures $(x,\omega) \in C$, according to (1.21), and as mentioned after (1.13), $M \subseteq C$. Now the preceding paragraph shows $M \subseteq M'$. If $M' \subseteq M$ fails, M' intersects the closed set $C \setminus M$, and (1.22) shows $\mathscr{M}_i(\Sigma_{C \setminus M}) \neq \varnothing$; but then there exists $\Phi \in \mathscr{M}_i$ with $\Phi(C \setminus M) = 1$, and this contradicts the definition of M. So (1.23) is established.

As in Cogburn [5] define

$$S(x,\omega,n) = \{y: P(\omega_0 \cdots \omega_{n-1}; x, y) > 0\}, \qquad (x,\omega) \in E, n = 1, 2, \ldots,$$

and

$$(x,\omega) \leftrightarrow (y,\omega') \Leftrightarrow \omega = \omega' \quad \text{and} \quad \bigcup_{n=1}^{\infty} [S(x,\omega,n) \cap S(y,\omega,n)] \neq \varnothing.$$

The relation $\leftrightarrow$ is supposed to correspond to that of two states intercommunicating in the theory of Markov chains with constant transition probabilities. The relation $\leftrightarrow$ need not be transitive. However, it follows from [5], Lemma 3 and Proposition 1 (see also Remark 1.3.5 below) that

(1.24) $\qquad \leftrightarrow$ restricted to $M \times M$ is an equivalence relation.

For $(x,\omega) \in M$, the corresponding equivalence class will be denoted by $[x,\omega]$; we emphasize $[x,\omega] \subseteq M$. (If $(x,\omega) \notin M$, $[x,\omega]$ is not defined.) For every equivalence class $[x,\omega]$, there exists a unique equivalence class $[x', T\omega]$ such that $P(\omega_0; y, [(x', T\omega)]) = I_{[x,\omega]}(y,\omega)$ for $(y,\omega) \in M$. We denote the class $[x', T\omega]$ by $S([x,\omega])$. The map S is a bijection of the set of equivalence classes onto itself. For $F \subseteq M$ let $[F] := \cup_{(x,\omega)\in F}[(x,\omega])$. Then $F \in \mathscr{E} \cap M$ implies $[F] \in \mathscr{E} \cap M$ is easily shown ([5], Section 2). Also there the following elegant characterization of $\mathscr{E}_d(\Sigma_M)$ is obtained:

$$\mathscr{E}_d(\Sigma_M) = \{[F] \colon F \in \mathscr{E} \cap M\}. \tag{1.25}$$

Hence a function on M is $\mathscr{E}_d(\Sigma_M)$-measurable if and only if it is $\mathscr{E}$-measurable and assumes a constant value on each equivalence class $[x,\omega]$. For $F \in \mathscr{E}_d(\Sigma_M)$ let $\Lambda(\omega, F)$ equal the number of distinct $\leftrightarrow$-equivalence classes contained in $(F)^{\omega}$. Then $\Lambda(\cdot, F)$ is $\mathscr{B}$-measurable for $F \in \mathscr{E}_d(\Sigma_M)$; and for fixed ω, $\Lambda(\omega, \cdot)$ is a measure on $\mathscr{E}_d(\Sigma_M)$. Let

$$\Lambda(F) := \int \Lambda(\omega, F)\pi(d\omega), \qquad F \in \mathscr{E}_d(\Sigma_M).$$

On $\mathscr{E}_d(\Sigma_M)$, Λ and Π are mutually absolutely continuous. Since S is a bijection

$$\Lambda P = \Lambda \quad \text{on } \mathscr{E}_d(\Sigma_M). \tag{1.26}$$

Recall that $\mathscr{E}_i(\Sigma_C) \subseteq \mathscr{E}_d(\Sigma_C)$. For $F \in \mathscr{E}_i(\Sigma_M)$ one sees that when π is ergodic $\Lambda(\omega, F) = \Lambda(F)$ π-a.e. Suppose now that $\Phi \in \mathscr{M}_i$ with density φ and that $F := \{\varphi > 0\}$ is a minimal closed set. Then $\Lambda(F) < \infty$ (use (1.17) with $f = \chi_x$ and $P^k\chi_x(y,\omega)P^k\chi_x(z,\omega) = 0$ if $[y,\omega] \neq [z,\omega]$) and one obtains

$$\Phi = \Lambda_F \quad \text{on } \mathscr{E}_d(\Sigma_M), \tag{1.27}$$

where Λ_F is defined by

$$\Lambda_F(A) = \frac{\Lambda(A \cap F)}{\Lambda(F)}, \qquad A \in \mathscr{E}_d(\Sigma_M).$$

To see (1.27), note that for $A \in \mathscr{E}_d(\Sigma_M)$, $PI_A = I_B$, where necessarily $B \in \mathscr{E}_d(\Sigma_M)$. So as explained in Section 1.1 we can restrict Σ_M to the σ-field $\mathscr{E}_d(\Sigma_M)$, obtaining a restricted process, say Σ'. For this process F is a minimal closed set and Λ_F and $\Phi|\mathscr{E}_d(\Sigma_M)$ both belong to $\mathscr{M}_i(\Sigma'_F)$. Because F is minimal, $\mathscr{E}_i(\Sigma'_F)$ is trivial and (1.27) follows from the ergodic theorem. Note again that for our F,

$$\Lambda(F) < \infty. \tag{1.28}$$

Consider now any $\Phi \in \mathscr{M}_i$ with density φ. Then $\varphi = \varphi P$ and by (1.20), $\Sigma_x \varphi(x, \omega) = \Sigma_x \varphi(x, T^{-1}\omega)$. Hence

$$\Sigma_x \varphi(x, \omega) = 1 \quad \text{for } \varphi \in \mathscr{D}_i, \text{ if } \pi \text{ is ergodic.} \tag{1.29}$$

If π is ergodic and F is closed, $\pi(\bigcup_x (F)_x) = 1$. It follows now from (1.27) that if $\varphi \in \mathscr{D}_i$, $F := \{\varphi > 0\}$ is a minimal closed set and π is ergodic, then

$$\sum_{y:\,[y,\omega]=[x,\omega]} \varphi(y, \omega) = (\Lambda(F))^{-1} \qquad \pi\text{-a.e. on } F. \tag{1.30}$$

We note next the surprising fact, [3], Proposition 2.1, and [13], Theorem 2, that

$$\varphi \text{ is } \mathscr{E}_{-\infty}^{-1}\text{-measurable whenever } \varphi \in \mathscr{D}_i; \tag{1.31}$$

the key to the proof is (1.20).

Returning to our coordinate variables (X_k), let $\sigma(X_k\colon k \geq n)$ be the σ-field generated by $\{X_k\colon k \geq n\}$ and let $\mathscr{T} = \bigcap_{n=0}^{\infty} \sigma(X_k\colon k \geq n)$ be the tail σ-field. For a closed set $F \in \mathscr{E}$, say that the *zero–one property holds on* F if for Π-a.e. $(x, \omega) \in F$, $P_{(x,\omega)}$ restricted to $\mathscr{T}$ assumes only the values 0 and 1. In [3] various necessary and sufficient conditions for the zero–one property are derived. In particular ([3], Corollary 3), if $\mathscr{M}_i \neq \varnothing$, then

$$\text{the zero–one property holds on } M. \tag{1.32}$$

Let $\mathscr{M}(X)$ denote the set of all probability measures on X, topologized as a subset of $\overline{\mathscr{M}}(X)$, the Banach space of all signed measures on X with finite total variation, using the total variation, denoted by $\|\cdot\|$, as the norm. For $\alpha, \beta \in \mathscr{M}(X)$ let

$$\delta_n(\alpha, \beta, \omega) = \|(\alpha - \beta)P(\omega_0 \cdots \omega_{n-1})\| \tag{1.33}$$

and define

$$\delta(\alpha, \beta, \omega) = \lim_n \delta_n(\alpha, \beta, \omega). \tag{1.34}$$

Using (1.24), one can obtain a useful reformulation of (1.32). Consider $\varphi \in \mathscr{D}_i$ and suppose $F := \{\varphi > 0\}$ is a minimal closed set. Define

$$\varphi_{[x,\omega]}(y) := \Lambda(F)\varphi(y, \omega)I_{[x,\omega]}(y, \omega), \qquad (x, \omega) \in M, y \in X. \tag{1.35}$$

Then

$$\varphi_{[x,\omega]} \in \mathscr{M}(X) \quad \text{for } \Pi\text{-a.e. } (x, \omega) \in F, \text{ provided } \pi \text{ is ergodic,} \tag{1.36}$$

by (1.29) and (1.30). The reformulation of (1.32) ([5], Corollary 5) now becomes: Let $\varphi \in \mathscr{D}_i$ with $F := \{\varphi > 0\}$ a minimal closed set; then

$$\begin{aligned} &I_F(x, \omega)\delta(\varepsilon_y, \varphi_{[x,\omega]}, \omega) = 0 \quad \text{a.e.} \\ &\qquad\qquad \text{if } [x, \omega] = [y, \omega], \text{ provided } \pi \text{ is ergodic,} \end{aligned} \tag{1.37}$$

where ε_y is the unit mass at y. Note also, following Cogburn ([5], proof of Theorem 5), that regardless of whether π is ergodic, for π-a.e. ω and all x, y such that (y, ω) and $(x, T^n\omega)$ belong to M,

$$\sum_z \varphi_{[y,\omega]}(z) P(\omega_0 \cdots \omega_{n-1}; z, x) = \begin{cases} \Lambda(F)\varphi(x, T^n\omega), & P(\omega_0 \cdots \omega_{n-1}; z, [x, T^n\omega]) > 0, \\ 0, & \text{otherwise}. \end{cases} \tag{1.38}$$

1.3. *Open problems and remarks.* We present some problems and remarks related directly to the material of the previous section.

PROBLEM 1.3.1. Find conditions for the existence of a σ-finite measure Φ on $(E, \mathscr{E})$ with $\Phi \ll \Pi$ and $\Phi = \Phi P$.

REMARK 1.3.2. The corresponding problem for general Hopf Markov chains is studied by Foguel ([7], Chapter 6). At present no nice conditions applying specifically to the chains Σ discussed in Section 1.2 are known. In the nonrandom case (i.e., countable state space Markov chains with constant transition probabilities) $C \neq \varnothing$ is according to Derman's theorem a sufficient condition ([1], Theorems 7 and 8).

PROBLEM 1.3.3. When does the zero–one property hold on C?

REMARK 1.3.4. In the nonrandom case the answer is: always; see, for example, [8], Section 1.13. Presently no counterexamples are known in the general situation, nor are any nontrivial sufficient conditions known. Even under the assumption that the underlying coordinate process (Θ_n) is a two-state Bernoulli process the problem is open.

REMARK 1.3.5. In [5] a stronger result than (1.24) is proved. With ε_x representing the unit mass at x, let $C_1 = \{(x, \omega) \in C: \delta(\varepsilon_y, \varepsilon_z, T^n\omega) < 2$ whenever $y, z \in S(x, \omega, n),\ n \geq 1\}$. Then C_1 is closed, $M \subseteq C_1$ and $\leftrightarrow$ is an equivalence relation on $C_1 \times C_1$.

PROBLEM 1.3.6. When does $C_1 = C$?

REMARK 1.3.7. As can be seen from the preceding discussion and our subsequent results, that part of the ergodic theory of Markov chains with constant transition probabilities dealing with positive recurrence by and large extends nicely to the present situation. Just the opposite is the case for the null recurrent theory: appropriate generalizations are not known, nor are counterexamples.

2. Minimal closed sets. Closed and minimal closed sets were defined in Section 1.1 for the general Hopf Markov chains. One cannot expect that closed sets always contain minimal closed sets. For a simple example consider the closed unit disk in $\mathbb{R}^2$, with Lebesgue measure, and the measure preserving transformation which sends $re^{i\theta}$ into $re^{i(\theta+\alpha)}$, with α an irrational multiple of π. In an obvious way (see [7], (c), page 6) such a transformation may be considered as a Hopf Markov chain. In this example there are many closed sets, for example, any annulus centered at the origin. However, the only candidate for minimal closed sets are circles of radius r, $0 \le r \le 1$, and center at the origin, and these are null sets; so there are no minimal closed sets. Note that in this example $X = M$.

We now discuss the existence of minimal closed sets for the systems Σ introduced in Section 2.1. To avoid the kind of counterexample of the previous paragraph, it is reasonable to assume that π is ergodic. Then for any closed set F, $\bigcup_x \pi(F)_x = 1$, and using this Cogburn ([3], Proposition 2.3) proved that if F is closed and $\Pi(F) < \infty$, then F will contain a minimal closed set.

The main result of this section is that if π is ergodic, then every closed subset of M contains a minimal closed subset. This gives immediately (Corollary 2.2) an ergodic decomposition for elements of $\mathscr{M}_i$.

For $\Phi \in \mathscr{M}_i$, F any closed set with $\Phi(F) > 0$, define Φ_F by

$$\Phi_F(A) = \frac{1}{\Phi(F)}\Phi(F \cap A). \tag{2.1}$$

Clearly $\Phi_F \in \mathscr{M}_i$. If F is a minimal closed set, Φ_F is *ergodic* in the sense that it assigns measure 0 or 1 to each element of $\mathscr{E}_i(\Sigma_M)$.

THEOREM 2.1. *Assume π is ergodic. Then every closed subset of M contains a minimal closed subset.*

PROOF. Put $\mathscr{E}' = \mathscr{E}_i(\Sigma_M)$. What must be shown is that every $\Phi \in \mathscr{M}_i$ is atomic on $\mathscr{E}'$. So let $\Phi \in \mathscr{M}_i$ with density φ. Then

$$0 \le \varphi(z, \omega) \le 1 \tag{2.2}$$

follows from (1.29). Let F be closed, $\Phi(F) > 0$. Then Φ_F defined in (2.1) has a density φ_F given by

$$\varphi_F = \frac{\varphi \cdot I_F}{\Phi(F)} \tag{2.3}$$

and applying (2.2) to this density, we have

$$\Phi(F) \ge \varphi(z, \omega) I_F(z, \omega). \tag{2.4}$$

We wish to show that Φ is atomic on $\mathscr{E}'$. One easily sees that if that is not the case there is an element of $\mathscr{M}_i$ which is nonatomic (i.e., it has no atoms) on $\mathscr{E}'$. So we assume now (in order to obtain a contradiction) that Φ is nonatomic on $\mathscr{E}'$.

Since Φ is nonatomic on $\mathscr{E}'$ it is possible, for each positive integer n, to partition M into $\{F_i, 1 \le i \le n\}$ so that $F_i \in \mathscr{E}'$ and $\Phi(F_i) = n^{-1}$, $1 \le i \le n$. By (2.4), $\varphi \le n^{-1}$ Φ-a.e. Hence $\varphi = 0$ Φ-a.e., and this is impossible. □

COROLLARY 2.2. *Assume π is ergodic. Then for each $\Phi \in \mathscr{M}_i$,*

$$\Phi = \sum \Phi(F)\Phi_F,$$

where the sum extends over all minimal closed sets F with $\Phi(F) > 0$; the Φ_F are ergodic.

PROOF. Follows immediately from Theorem 2.1. □

PROBLEM 2.3. Do closed subsets of C always contain minimal closed subsets?

REMARK 2.4. In the nonrandom case, C decomposes into minimal closed sets; these are just the recurrent classes ([1], Chapter 1).

3. Σ and Σ^+. Instead of beginning with the bilateral sequence space Ω introduced in Section 1, one can begin with $\Omega^{(+)} = \Theta^{Z_+}$, where Z_+ is the set of nonnegative integers and proceed to the corresponding skew product Markov chain moving on $E^{(+)} := X \times \Omega^{(+)}$, giving rise to a Hopf Markov chain $\Sigma^{(+)}$. Σ and $\Sigma^{(+)}$ each have their own advantages, and it is useful to consider both, and the relations between them.

In place of $\Sigma^{(+)}$ we will study an isomorphic Hopf Markov chain Σ^+, which we now specify. Starting with Σ as defined in Section 1.2, Σ^+ is obtained by restricting to the σ-field $\mathscr{E}^+ := \mathscr{E}_0^\infty$. Note that if $A \in \mathscr{E}^+$, $P(\cdot, A)$ is $\mathscr{E}^+$-measurable, so the restriction is legitimate. We write $\Sigma^+ = (E, \mathscr{E}^+, P^+, \Pi^+)$, where $\Pi^+ = \Pi|\mathscr{E}^+$ and P^+ is the restriction of P to $L_1(\Pi^+)$. Observe that if ψ is a $\mathscr{E}^+$-measurable function, then $\psi(x, \omega)$ depends on ω only through ω^+, where ω^+ is the restriction of ω to Z_+, and we will write $\psi(x, \omega) = \psi(x, \omega^+)$.

We turn to the formulas corresponding to (1.19) and (1.20). Choosing $f \in L_\infty(\Pi^+)$ and replacing P^n by $(P^+)^n$ on the left side of (1.19), the right side requires no change. Now consider the analogue of (1.20) with $n = 1$. For $\psi \in L_1(\Pi^+)$ one finds now from (1.2) that

$$\psi P^+(x, \omega^+) = \sum_y \int \psi(y, \theta\omega^+) P(\theta; y, x) \pi_0(d\theta|\omega^+), \tag{3.1}$$

where $\theta\omega^+ = (\theta, \omega_0, \omega_1, \dots)$ and

$$\pi_0(d\theta|\omega^+) = \pi^+\{\Theta_0 \in d\theta|(\Theta_1, \Theta_2, \dots) = (\omega_0^+, \omega_1^+, \dots)\}.$$

If one does not want to assume the existence of this conditional probability distribution, the right side of (3.1) can be rewritten as a conditional expectation.

Most results about Σ reported in Section 1.2 carry over to Σ^+; some evidently have no analogue, for example, (1.31). Since (1.23) also holds for Σ_F^+, $\mathscr{M}_i(\Sigma) \neq \varnothing$ if and only if $\mathscr{M}_i(\Sigma^+) \neq \varnothing$. If $\Phi \in \mathscr{M}_i(\Sigma)$, then Φ^+ defined by

$$\Phi^+ := \Phi|\mathscr{E}^+ \tag{3.2}$$

will belong to $\mathscr{M}_i(\Sigma^+)$; if furthermore Φ has density φ, the density of Φ^+ will be given by

$$\varphi^+ := \Pi\{\varphi|\mathscr{E}^+\}. \tag{3.3}$$

Since the right member of (3.3) also equals $\Pi\{\varphi I_{\{\varphi^+ > 0\}}|\mathscr{E}^+\}$, one obtains

$$\{\varphi > 0\} \subseteq \{\varphi^+ > 0\}. \tag{3.4}$$

Frequently the inclusion in (3.4) is strict, and then knowing that something holds [e.g., the existence of the limit in (1.17)] Φ^+-a.e. is better than just knowing it holds Φ-a.e. In particular, when $\mathscr{M}_i(\Sigma^+) \neq \varnothing$, $\mathscr{M}_i(\Sigma^+)$ will contain a probability with density having maximal support, as guaranteed by (1.13). We will write $M^+ = M(\Sigma^+)$, and continue to write M for $M(\Sigma)$. By (3.4) then $M \subseteq M^+$. If $\Phi \in \mathscr{M}_i$ has density φ with $\{\varphi > 0\} = M$ and F is any closed subset of M, then $\varphi_F := (\Phi(F))^{-1}\varphi I_F$ belongs to $\mathscr{D}_i$, and by (1.31), F is $\mathscr{E}_{-\infty}^{-1}$-measurable. That is,

$$F \text{ is } \mathscr{E}_{-\infty}^{-1}\text{-measurable if } F \subseteq M,\ F \text{ closed}; \tag{3.5}$$

of course, we do not distinguish between sets differing by Π-null sets.

Under the assumption that Θ is countable, the following theorem was proved by Nawrotzki [12, 13, 14].

THEOREM 3.1. *The restriction map taking $\Phi \in \mathscr{M}_i(\Sigma)$ into $\Phi^+ \in \mathscr{M}_i(\Sigma^+)$ is a bijection $\mathscr{M}_i(\Sigma) \to \mathscr{M}_i(\Sigma^+)$.*

PROOF. 1. Let $\Phi_j \in \mathscr{M}_i(\Sigma)$, $j = 1, 2$, and suppose $\Phi_1^+ = \Phi_2^+$. It must be shown that $\Phi_1 = \Phi_2$. Let the density of Φ_j be φ_j, $j = 1, 2$. By assumption $\varphi_1^+ = \varphi_2^+$, so that for any bounded $\mathscr{E}^+$-measurable function f,

$$\sum_x \int \varphi_1(x, \omega) f(x, \omega)\pi(d\omega) = \sum_x \int \varphi_2(x, \omega) f(x, \omega)\pi(d\omega). \tag{3.6}$$

For such an f and any positive integer n, we claim

$$\begin{aligned}
&\sum_x \int \varphi_1(x,\omega) f(x, T^{-n}\omega)\pi(d\omega)\\
&\quad = \sum_x \int \varphi_1(x, T^n\omega) f(x,\omega)\pi(d\omega)\\
&\quad = \sum_x \int (\varphi_1 P^n)(x, T^n\omega) f(x,\omega)\pi(d\omega)\\
&\quad = \sum_x \int \sum_y \varphi_1(y,\omega) P(\omega_0 \cdots \omega_{n-1}; y, x) f(x,\omega)\pi(d\omega) \qquad (3.7)\\
&\quad = \sum_y \int \varphi_1(y,\omega)\Big(\sum_x P(\omega_0 \cdots \omega_{n-1}; y, x) f(x,\omega)\Big)\pi(d\omega)\\
&\quad = \sum_y \int \varphi_2(y,\omega)\Big(\sum_x P(\omega_0 \cdots \omega_{n-1}; y, x) f(x,\omega)\Big)\pi(d\omega)\\
&\quad = \sum_x \int \varphi_2(x,\omega) f(x, T^{-n}\omega)\pi(d\omega),
\end{aligned}$$

where the second equality comes from $\varphi_1 \in \mathscr{D}_i$, the third equality follows from (1.20) and the fifth equality is (3.6) applied to the expression in parentheses in place of f; finally the sixth equality comes from the equality of the first and fifth members by replacing φ_1 by φ_2. It follows now that (3.6) holds for any bounded $\mathscr{E}_{-n}^{\infty}$-measurable f and hence, since n is arbitrary, also for any bounded $\mathscr{E}$-measurable f. This shows that $\varphi_1 = \varphi_2$. Thus the restriction map is injective.

2. Let $\Phi \in \mathscr{M}_i(\Sigma)$ be ergodic, with density φ. We will show Φ^+ is ergodic. By assumption $\{\varphi > 0\}$ is a minimal closed set, and it must be shown that the closed set $F := \{\varphi^+ > 0\}$ is also minimal closed in Σ^+. Suppose otherwise: $F = F_1 \cup F_2$ with F_1 and F_2 disjoint, closed sets in $\mathscr{E}^+$. Then $F_j \cap \{\varphi > 0\}$ is either a null set or a closed subset of $\{\varphi > 0\}$, $j = 1, 2$. Since $\{\varphi > 0\}$ is minimal closed, the first alternative must hold for $j = 1$ or $j = 2$. Say $F_1 \cap \{\varphi > 0\} = \varnothing$. Then $\Pi\{I_{F_1}\varphi | \mathscr{E}^+\} = I_{F_1}\varphi^+ = 0$ a.e., contradicting the choice of F_1.

3. By paragraph 2 the ergodic elements of $\mathscr{M}_i(\Sigma)$ map into ergodic elements of $\mathscr{M}_i(\Sigma^+)$, and using the ergodic decomposition of Corollary 2.2, it follows that to conclude that the restriction map is surjective it only remains to show that $\{(\varphi_M)^+ > 0\} = M^+$, that is, $F := M^+ \setminus \{(\varphi_M)^+ > 0\} = \varnothing$. Suppose $F \neq \varnothing$; then F is a closed subset of M^+, $F \in \mathscr{E}^+$ and there will exist $\Psi \in \mathscr{M}_i(\Sigma^+)$ with $\Psi(F) = 1$. From this one obtains, by applying the ergodic theorem,

$$\Pi\left\{(x,\omega) \in F\colon \sup_y \limsup_n \frac{1}{n} \sum_{k=0}^{n-1} P^k(\omega_0 \cdots \omega_k; x, y) > 0\right\} > 0$$

and so, by (1.23), $\mathscr{M}_i(\Sigma_F) \neq \varnothing$. Hence $F \cap M \neq \varnothing$, so that $\varphi_M I_F$ does not vanish a.e., which contradicts the assumption that $(\varphi_M)^+$ vanishes a.e. on F. □

COROLLARY 3.2. *The ergodic decomposition in Corollary 2.2 holds for* $\Phi^+ \in \mathscr{M}_i(\Sigma^+)$.

PROOF. Use Corollary 2.2 and Theorem 3.1. □

We follow Nawrotzki [14], Theorem 5, and now apply Theorem 3.1 to the case that the coordinate random variables (Θ_n) are independent under π. For $\varphi \in L_1(\pi)$ define $\bar{\varphi}(y) := \int \varphi(y,\omega)\pi(d\omega)$; also set $\bar{P}(x,y) := \int P(\omega_0; x, y)\pi(d\omega)$ and $\bar{\mathscr{D}}_i = \{\psi \in \mathscr{M}(X): \psi = \psi\bar{P}\}$.

THEOREM 3.3. *Assume the* (Θ_k) *are independent. Then the densities* $\mathscr{D}_i(\Sigma^+)$ *are nonrandom* (*i.e., they do not depend on* ω) *and* $\mathscr{D}_i(\Sigma^+) = \bar{\mathscr{D}}_i$.

PROOF. Let $\varphi \in \mathscr{D}_i(\Sigma)$. Then by (1.20),

$$\varphi(x,\omega) = \sum_y \varphi(y, T^{-1}\omega)P(\omega_{-1}; y, x). \tag{3.8}$$

According to (1.31), φ is $\mathscr{E}_{-\infty}^{-1}$-measurable, so the two factors inside the sum on the right are independent. Integrating, one obtains

$$\bar{\varphi}(x) = \sum_y \bar{\varphi}(y)\bar{P}(y,x). \tag{3.9}$$

Since φ belongs to $\mathscr{D}_i(\Sigma)$, $\varphi^+ := \Pi\{\varphi | \mathscr{E}^+\}$ belongs to $\mathscr{D}_i(\Sigma^+)$, and again using that φ is $\mathscr{E}_{-\infty}^{-1}$-measurable and that π is product measure, one sees that $\varphi^+ = \bar{\varphi}$. By Theorem 3.1 every element of $\mathscr{D}_i(\Sigma^+)$ is of the form φ^+ for some φ in $\mathscr{D}_i(\Sigma)$. So $\mathscr{D}_i(\Sigma^+) \subseteq \bar{\mathscr{D}}_i$ has been established.

Now assume $\psi \in \bar{\mathscr{D}}_i$. Then $\psi \in L_1(\Pi^+)$ and ψ does not depend on ω. Using (3.1) and the fact that π is product measure, one obtains $\psi P^+(x,\omega^+) = \psi\bar{P}(x)$ and by assumption $\psi\bar{P} = \psi$. This proves $\bar{\mathscr{D}}_i \subseteq \mathscr{D}_i(\Sigma^+)$. □

COROLLARY 3.4. *Assume the* (Θ_k) *are independent. If* F *and* G *are disjoint closed subsets of* M, *then for every* $x \in X$ *either* $(F)_x = \varnothing$, *or* $(G)_x = \varnothing$.

PROOF. If the corollary is false there must exist disjoint minimal closed subsets F and G of M and $x \in X$ with $(F)_x \neq \varnothing$, $(G)_x \neq \varnothing$. Then there will exist Φ and Ψ in $\mathscr{M}_1$ with densities φ and ψ satisfying $F = \{\varphi > 0\}$, $G = \{\psi > 0\}$. It follows from Theorem 3.1 that Φ^+ and Ψ^+ must be distinct ergodic members of $\mathscr{M}_i(\Sigma^+)$ and hence $\{\varphi^+ > 0\} \cap \{\psi^+ > 0\} = \varnothing$. By Theorem 3.2, $\varphi^+ = \bar{\varphi}$ and $\psi^+ = \bar{\psi}$ and so $\bar{\varphi}(x)\bar{\psi}(x) = 0$. On the other hand $(F)_x \neq \varnothing$ and $(G)_x \neq \varnothing$ imply $\bar{\varphi}(x) > 0$, $\bar{\psi}(x) > 0$. This contradiction proves the corollary. □

4. A weak convergence theorem. In [5] Cogburn investigates

$$\lim_n \frac{1}{n} \sum_{k=0}^{n-1} I_{(0,1]}(P(\omega_0 \cdots \omega_k; y, x)),$$

showing in [5], Theorem 5, that if $\Phi \in \mathscr{M}_i$ and Φ is ergodic and $x \in X$, then the limit exists Φ-a.e. and giving the value of the limit. We will obtain this result again in Corollary 4.2. Theorem 4.1 is the main result of this section; it was found in the process of trying to get a better understanding of Cogburn's result.

A vector $(P(x))_{x \in X}$ will be denoted simply by $P(\cdot)$. Note that

$$P(\omega_0 \cdots \omega_{n-1}; z, \cdot) \in \mathscr{M}(X).$$

For $(y, \omega) \in E$ and N a positive integer define:

(4.1) $m_N^{(y,\omega)}$ is the probability on $\mathscr{M}(X)$ assigning mass $1/N$ to $P(\omega_0 \cdots \omega_{n-1}; y, \cdot)$, $n = 1, \dots, N$.

In Theorem 4.1 it is shown that if π is ergodic and $\Phi \in \mathscr{M}_i$ is ergodic, then there exists a probability m on $\mathscr{M}(X)$ with $m_N^{(y,\omega)}$ converging to m weakly as $N \to \infty$ for Φ-a.e. (y, ω), and m is given explicitly.

If π is ergodic and $\mathscr{M}_i \neq \varnothing$, the assumption that we are dealing with an ergodic element of $\mathscr{M}_i$ is not an essential restriction, in view of Corollary 2.2. On the other hand, if π is not ergodic, and if it is assumed that there exists an ergodic $\Phi \in \mathscr{M}_i$ with density φ and $F = \{\varphi > 0\}$, then π restricted to $\bigcup_x (F)_x$ would be an ergodic measure with total mass less than 1. So to avoid introducing and carrying along another constant, one may as well assume that π is ergodic.

So we assume now that π is ergodic, Φ is ergodic, φ, F as above. Then (1.36) applies. Relation (1.38) can be written

$$\sum_z \varphi_{[y,\omega]}(z) P(\omega_0 \cdots \omega_{n-1}; z, \cdot) = \Lambda(F)\varphi(\cdot, T^n\omega) I_{S^n[y,\omega]}(\cdot, T^n\omega)$$

and this holds for π-a.e. ω and all y such that $[y, \omega] \in M$; the dot signifies that this is a vector equality, that is, if the dot is replaced by x the two sides are to be equal for all $x \in X$. Let

$$B(y, x, n) = \{\omega\colon (x, \omega) \in M, (y, T^{-n}\omega) \in M, S^n[y, T^{-n}\omega] = [x, \omega]\}.$$

The preceding equality becomes

$$(4.2)\quad \sum_z \varphi_{[y,\omega]}(z) P(\omega_0 \cdots \omega_{n-1}; z, \cdot) = \Lambda(F)\varphi(\cdot, T^n\omega) I_{B(y,\cdot,n)}(T^n\omega).$$

Recall the Markov chain $((X_n, \Theta_n), n = 0, 1, \dots)$ introduced near the beginning of Section 1.2. Of course,

$$(4.3)\qquad P_{(y,\omega)}[\Theta_n = T^n\omega, n = 0, 1, \dots] = 1, \qquad (y, \omega) \in E.$$

Let

$$A(x) = \{(z, \omega) \in M\colon (z, \omega) \leftrightarrow (x, \omega), (x, \omega) \in C\}.$$

Then

$$P_{(y,\omega)}[[(X_n, \Theta_n) \in A(\cdot)] = [\Theta_n \in B(y, \cdot, n)]] = 1, \quad (y, \omega) \in M, n \geq 0. \tag{4.4}$$

Relations (4.2), (4.3) and (4.4) give

$$\sum_z \varphi_{[y,\omega]}(z) P(\omega_0 \cdots \omega_{n-1}; z, \cdot) = \Lambda(F)\varphi(\cdot, \Theta_n) I_{A(\cdot)}(X_n, \Theta_n) \tag{4.5}$$

with $P_{(y,\omega)}$-probability 1, for $(y, \omega) \in M$.

Note that

$$\Lambda(F)\varphi(x, \omega) I_{A(x)}(y, \omega) = \varphi_{[y,\omega]}(x) \tag{4.6}$$

and by (1.36), $\varphi_{[y,\omega]}(\cdot) \in \mathscr{M}(x)$ if $(y, \omega) \in F$. Now let $g: \mathscr{M}(X) \to \mathbb{R}$ be bounded and measurable. Then

$$\begin{aligned} \lim_N \frac{1}{N} \sum_{n=0}^{N-1} g(\varphi_{[X_n, \Theta_n]}(\cdot)) &= \int g(\varphi_{[z,\omega]}(\cdot)) \Phi(d(z \times \omega)) \\ &= \int g(\varphi_{[z,\omega]}(\cdot)) \Lambda_F(d(z \times \omega)), \qquad P_\Phi\text{-a.e.}, \end{aligned} \tag{4.7}$$

where the first equality comes from Birkhoff's ergodic theorem, and since the integrand in the middle member is $\mathscr{F}_d$-measurable, the second equality holds by (1.27). Since the equations in (4.7) hold P_Φ-a.e., they also hold $P_{(y,\omega)}$-a.e. for Φ-a.e. (y, ω). Combining (4.7) with (4.5) and (4.6), one obtains: For Φ-a.e. (y, ω),

$$\begin{aligned} &\lim_N \frac{1}{N} \sum_{n=1}^{N} g\Big(\sum_z \varphi_{[y,\omega]}(z) P(\omega_0 \cdots \omega_{n-1}, z, \cdot)\Big) \\ &\quad = \int g(v) m(dv), \qquad P_{y,\omega}\text{-a.e.}, \end{aligned} \tag{4.8}$$

where the probability m on $\mathscr{M}(X)$ is defined by

$$m(dv) = \Lambda_F\{(z, \omega) : \varphi_{[z,\omega]}(\cdot) \in dv\}. \tag{4.9}$$

If one now makes the additional assumption about g that it is continuous in the total variation norm, it follows from (1.37) and (4.8) that

$$\lim_N \frac{1}{N} \sum_{n=1}^{N} g(P(\omega_0 \cdots \omega_{n-1}; y, \cdot)) = \int g\, dm, \qquad \Phi\text{-a.e. } (y, \omega). \tag{4.10}$$

This proves the following result.

THEOREM 4.1. *Let π be ergodic and also $\Phi \in \mathscr{M}_i$ be ergodic with density φ, $F = \{\varphi > 0\}$. Then as $N \to \infty$, $m_N^{(y,\omega)} \to m$ weakly for Φ-a.e. (y, ω), where $m_N^{(y,\omega)}$ and m are defined in* (4.1) *and* (4.9), *respectively.*

REMARK 4.2. As usual in the theory of weak convergence, it suffices for the validity of (4.10) that g is bounded and $m\{v\colon g \text{ is discontinuous at } v\} = 0$.

The measure m is the distribution under Λ_F of the $\mathscr{M}(X)$-valued random variable defined on E with value $\varphi_{[z,\omega]}(\cdot)$ at (z, ω).

If, say, g depends only on $v(x)$ and $v(z)$ where $x, z \in X$ and $x = z$ is allowed, so that $g(v) = g_0(v(x), v(z))$, the right side of (4.10) reduces to $\int g_0\, dm^{x,z}$, where $m^{x,z}$ is a probability on $[0,1]^2$, $m^{x,z}(U) = \Lambda_F\{(y,\omega)\colon (\varphi_{[y,\omega]}(x), \varphi_{[y,\omega]}(z)) \in U\}$ and the proviso in Remark 4.2 becomes $m^{x,z}\{(s,t)\colon g_0$ is discontinuous at $(s,t)\} = 0$.

COROLLARY 4.3. *Assume π is ergodic, $\Phi \in \mathscr{M}_i$ is ergodic with density φ, $F := \{\varphi > 0\}$. For $(s,t) \in [0,1]^2$, $x \in X$,*

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} & I_{(s,1]}(P(\omega_0 \cdots \omega_{n-1}; y, x)) I_{(t,1]}(P(\omega_0 \cdots \omega_{n-1}; y, z)) \\ &= \frac{1}{\Lambda(F)} \pi\{\omega \in (F)_x \cap (F)_z\colon (x,\omega) \leftrightarrow (z,\omega), \\ &\qquad \Lambda(F)\varphi(x,\omega) \in [s,1], \Lambda(F)\varphi(z,\omega) \in (t,1)\}\end{aligned} \tag{4.11}$$

for Φ-a.e. (y,ω), provided $\pi\{\omega\colon \Lambda(F)\varphi(x,\omega) = s\} = \pi\{\omega\colon \Lambda(F)\varphi(z,\omega) = t\} = 0$. In any case, for Φ-a.e. (y,ω),

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} & I_{(0,1]}(P(\omega_0 \cdots \omega_{n-1}; y, \omega)) I_{(0,1]}(P(\omega_0 \cdots \omega_{n-1}; y, z)) \\ &= \frac{1}{\Lambda(F)} \pi\{\omega \in (F)_x \cap (F)_z\colon (x,\omega) \leftrightarrow (z,\omega)\}.\end{aligned} \tag{4.12}$$

PROOF. Assertion (4.11) follows from Theorem 4.1 and Remark 4.2 by choosing $g(v) = g_{s,t}(v(x), v(z))$ with $g_{s,t} = I_{(s,1]} I_{(t,1]}$. Assertion (4.12) would be a special case of (4.11), but now no proviso is imposed. Applying (4.11) with a sequence of (s,t) satisfying the proviso and converging to $(0,0)$, one obtains (4.12) with lim inf in place of lim and $\geq$ in place of $=$. For the opposite inequality use (4.8) with $g_{0,0}$ in place of g, and note that the resulting sum in (4.8) majorizes the sum in (4.12). □

REMARK 4.4. Since the measures $m_N^{(y,\omega)}$ introduced in (4.1) depend on ω only through the future and present, it seems curious that the limiting measure m of (4.9) is defined in terms of $\varphi \in \mathscr{D}_i$, which according to (1.31) depend on the past.

5. Asymptotic behavior of $P(\omega_{-n} \cdots \omega_{-1}; x, y)$. In the theory of Markov chains with constant transition probabilities, one can do much better than (1.17). The Cesaro convergence can be replaced by convergence, except that in the presence of periodicity one must go to ∞ along λn, $n = 0, 1, \ldots,$ where λ is the period; see, for example, [1], Chapter 1.6. There is no hope of such an improvement in (1.25); but as will be shown the situation is much better for products $P(\omega_{-n} \cdots \omega_{-1}; x, y)$.

Assume $\phi \in \mathscr{M}_i$ with density φ. Information about Cesaro convergence comes directly from the Chacon–Ornstein theorem (1.14)–(1.16). Again let $\chi_y = I_{\{y\}\times\Omega}$. Using (1.20), one obtains then

$$\begin{aligned}(5.1)\quad \lim_n \frac{\sum_{k=0}^{n-1}\chi_y P^k(x,\omega)}{\sum_{k=0}^{n-1}\varphi P^k(x,\omega)} &= \lim_n \frac{\sum_{k=1}^{n-1}P(\omega_{-k}\cdots\omega_{-1}; y, x)}{n\varphi(x,\omega)}\\ &= L(\chi_y, \varphi)(x,\omega), \qquad \Pi\text{-a.e. on } \{\varphi > 0\}.\end{aligned}$$

Assume furthermore that $\{\varphi > 0\} = M$. It follows from Theorem 2.1 that M can be partitioned into minimal closed sets F_j, $0 \le j < p$, where $1 \le p \le \infty$. Let φ_j be the element in $\mathscr{D}_i$ with $\{\varphi_j > 0\} = F_j$, $0 \le j < p$. Now define

$$(5.2)\quad \begin{gathered}\varphi^*(x,\omega) := \varphi_j(x,\omega) \quad \text{and} \quad F^*(x,\omega) := F_j,\\ (x,\omega) \in F_j,\ 0 \le j < p.\end{gathered}$$

If $X = C$ the identification (1.16) may be used. One then obtains

$$\begin{aligned}(5.3)\quad \lim_n \frac{1}{n}\sum_{k=1}^{n-1} P(\omega_{-k}\cdots\omega_{-1}; y, x) &= \varphi^*(\omega, x)\pi\big((F^*(x,\omega))_y\big),\\ &\Pi\text{-a.e. } (x,\omega) \in M,\ y \in X, \text{ if } X = C.\end{aligned}$$

The last relation is similar to [5], Theorem 5.2, where however the hypothesis $X = C$ should be added.

PROBLEM 5.1. How does one obtain from the knowledge of $(P(\omega_n), -\infty < k \le -1)$ the value of p and the values $\varphi_j(x,\omega)$, $0 \le j < p$, $x \in X$? More refined problems involve estimating these values from a knowledge of $(P(\omega_k), -n \le k \le -1)$.

REMARK 5.2. The results of this section have some bearing on this problem, but do not solve it. For example, (5.3) is unsatisfactory because it assumes $X = C$ and because of the presence of the factors $\pi((F^*(x,\omega))_y)$. When X is finite Theorem 5.9 provides a lot of information but still does not quite solve the problem.

The Cesaro convergence guaranteed by (5.1) will now be strengthened.

THEOREM 5.3. *Suppose π is ergodic, $X = M$, $\mathscr{M}_i = \{\Phi\}$, $\Lambda(M) = 1$; let φ be the density of Φ. For $\alpha \in \mathscr{M}(X)$,*

$$\lim_n \frac{1}{n}\sum_{k=0}^{n-1} \|\alpha P(\omega_{-k}\cdots\omega_{-1}) - \varphi(\cdot,\omega)\| = 0, \qquad \pi\text{-a.e.}$$

Proof. Let $\alpha \in \mathscr{M}(X)$. The hypotheses of the theorem and (1.37) imply

$$\delta(\alpha, \varphi(\cdot, \omega), \omega) = 0, \qquad \pi\text{-a.e.} \tag{5.4}$$

Let $\varepsilon > 0$. By Birkhoff's ergodic theorem and (5.4), there exists N_ε such that

$$\lim_n \frac{1}{n} \sum_{k=0}^{n-1} \delta_N(\alpha, \varphi(\cdot, T^{-k}\omega), T^{-k}\omega)$$

$$= \int \delta_N(\alpha, \varphi(\cdot, \omega), \omega)\pi(d\omega) \le \varepsilon, \qquad N \ge N_\varepsilon.$$

In the first member replace δ_N in the kth summand by δ_k, $0 \le k \le n - 1$. For $k \ge N$ this causes the term to decrease or remain the same. Hence

$$\limsup_n \frac{1}{n} \sum_{k=0}^{n-1} \delta_k(\alpha, \varphi(\cdot, T^{-k}\omega), T^{-k}\omega) < \varepsilon$$

and so in fact

$$\lim_n \frac{1}{n} \sum_{k=0}^{n-1} \delta_k(\alpha, \varphi(\cdot, T^{-k}\omega), T^{-k}\omega) = 0. \tag{5.5}$$

The left side of (5.5) agrees with the first member of (5.6) below, and remembering $\varphi \in \mathscr{D}_i$ and (1.20) one obtains

$$\begin{aligned} &\lim_n \frac{1}{n} \sum_{k=0}^{n-1} \|\alpha P(\omega_{-k} \cdots \omega_{-1}) - \varphi(\cdot, T^{-k}\omega) P(\omega_{-k} \cdots \omega_{-1})\| \\ &\quad = \lim_n \frac{1}{n} \sum_{k=0}^{n-1} \|\alpha P(\omega_{-k} \cdots \omega_{-1}) - \varphi(\cdot, \omega)\| = 0. \end{aligned} \tag{5.6}$$

□

An immediate consequence is the following corollary.

Corollary 5.4. *Under the hypotheses of Theorem* 5.3,

$$\frac{1}{n} \sum_{k=0}^{n-1} |P(\omega_k \cdots \omega_{-1}; y, x) - \varphi(x, \omega)| \to 0 \quad \pi\text{-a.e. } \omega, x, y \in X. \tag{5.7}$$

For $\alpha, \beta \in \mathscr{M}(X)$, $\delta_n(\alpha, \beta, \omega)$ is nonincreasing in n; hence $\delta_n^*(\omega) := \sup\{\delta_n(\alpha, \beta, \omega): \alpha, \beta \in \mathscr{M}(X)\}$ is also nonincreasing and $\delta^*(\omega) = \lim_n \delta_n^*(\omega)$ exists.

Theorem 5.5. *Suppose π is ergodic, $X = M$, $\mathscr{M}_i = \{\Phi\}$, $\Lambda(M) = 1$ and $\delta^* = 0$ π-a.e. Let φ be the density of Φ. Then*

$$\lim_n \|\alpha P(\omega_{-n} \cdots \omega_{-1}) - \varphi(\cdot, \omega)\| = 0, \qquad \alpha \in \mathscr{M}(X). \tag{5.8}$$

Proof. It must be shown that in the proof of Theorem 5.3 the Cesaro convergence in (5.5) can be replaced by ordinary convergence. To see this, note $\delta_n(\alpha, \varphi(\cdot, T^{-n}\omega), T^{-n}\omega) \le \delta_n^*(T^{-n}\omega)$. Now $\delta_n^*(T^{-n}\omega)$ is nonincreasing in n, so has a limit $\delta_*(\omega)$. Since δ_n^* and $\delta_n^* \circ T^{-n}$ have the same distribution under π and $\delta^* = 0$ π-a.e. by assumption, it follows that $\delta_* = 0$ π-a.e. □

REMARK 5.6. If X is finite (in which case the hypothesis $\delta^* = 0$ is redundant), Theorem 5.5 is known; see Nawrotzki [14], Section 6, and Cogburn [4], Theorem 6. If X is finite and the (Θ_n) are independent, a closely related result was established by Takahaski ([16], Theorem 5). A number of other related results appear in Nawrotzki ([14], Section 6).

PROBLEM 5.7. How can one weaken the hypothesis $\delta^* = 0$ in Theorem 5.5? (In this connection the discussion in Nawrotzki ([14], Section 6), is again relevant.)

REMARK 5.8. Condition (5.7) is known to imply that $P(\omega_{-k} \cdots \omega_{-1}; y, x) \to \varphi(x, \omega)$ as $k \to \infty$, avoiding a subset $K_0(\omega)$ of density 0; see the discussion in [9], page 38.

One more result will be stated, without proof, for the case X is finite. The proof is easily assembled from the basic results of Cogburn outlined in Section 1.2. Supplement definition (5.2) by

$$\tilde{\varphi}(x, \omega) := \Lambda(F^*(x, \omega))\varphi^*(x, \omega), \qquad (x, \omega) \in M. \tag{5.9}$$

THEOREM 5.9. *Let X be finite. Then $M = C \neq \varnothing$. If furthermore π is ergodic, then for π-a.e. ω and all $\varepsilon > 0$ there exists n^* such that*

$$|P(\omega_{-n} \cdots \omega_{-1}; y, x) - \psi_n(\omega, y, x)| < \varepsilon, \qquad n \geq n^*, x, y \in X,$$

where

$$\psi_n(\omega, y, x) = \begin{cases} \tilde{\varphi}(x, \omega) & (y, T^{-n}\omega) \in C, \quad S^n[y, T^{-n}\omega] = [x, \omega], \\ 0, & (y, T^{-n}\omega) \in C, \quad S^n[y, T^{-n}\omega] \neq [x, \omega], \\ 0, & (x, \omega) \notin C, \end{cases}$$

and no assertion is made when ψ_n is not defined.

REFERENCES

[1] CHUNG, K. L. (1967). *Markov Chains with Stationary Transition Probabilities*, 2nd ed. Springer, New York.

[2] COGBURN, R. (1980). Markov chains in random environments: The case of Markovian environments. *Ann. Probab.* **8** 908–916.

[3] COGBURN, R. (1984). The ergodic theory of Markov chains in random environments. *Z. Wahrsch. Verw. Gebiete* **66** 109–128.

[4] COGBURN, R. (1986). On products of random stochastic matrices. *Contemp. Math.* **50** 199–213.

[5] COGBURN, R. (1990). On direct convergence and periodicity for transition probabilities of Markov chains in random environments. *Ann. Probab.* **18** 642–654.

[6] COGBURN, R. (1989). On the central limit theorem for Markov chains in a random environment. Preprint.

[7] FOGUEL, S. R. (1969). *The Ergodic Theory of Markov Processes*. Van Nostrand, Princeton.

[8] FREEDMAN, D. (1971). *Markov Chains*. Holden-Day, San Francisco.

[9] Halmos, P. R. (1956). *Lectures on Ergodic Theory*. Math. Soc. Japan, Tokyo.
[10] Hopf, E. (1954). The general temporally discrete Markov process. *J. Reine Angew. Math.* **205** 101–106.
[11] Krengel, U. (1985). *Ergodic Theorems*. de Gruyter, Berlin.
[12] Nawrotzki, K. (1975). Ein zufälliger Ergodensatz für eine Familie stochastischer Matritzen ohne gemeinsames invariantes verteilungsgesetz. *Math. Nachr.* **70** 15–28.
[13] Nawrotzki, K. (1977). Ein Grentzwertsatz für stationäre zufällige Folgen stochastischer Matrizen. *Math. Nachr.* **80** 133–150.
[14] Nawrotzki, K. (1981–1982). Discrete open systems or Markov chains in a random environment. I, II. *J. Inform. Process. Cybernet.* **17** 569–599; **18** 83–98.
[15] Revuz, D. (1975). *Markov Chains*. North-Holland, Amsterdam.
[16] Takahashi, H. (1969). Markov chains with random transition matrices. *Kodai Math. Sem. Rep.* **21** 426–447.

Department of Mathematics
University of Minnesota
Minneapolis, Minnesota 55455