

FINITELY ADDITIVE MARKOV CHAINS

BY

S. RAMAKRISHNAN

ABSTRACT. In this paper we develop the theory of Markov chains with stationary transition probabilities, where the transition probabilities and the initial distribution are assumed only to be finitely additive. We prove a strong law of large numbers for recurrent chains. The problem of existence and uniqueness of finitely additive stationary initial distributions is studied and the ergodicity of recurrent chains under a stationary initial distribution is proved.

1. Introduction. In this paper we start the study of Markov chains with stationary finitely additive transition probabilities and an initial distribution that need only be a finitely additive probability. The first serious attempt to study stochastic processes in a finitely additive setting was made by Dubins and Savage [8] who studied gambling with finitely additive probabilities. Given any strategy, they obtained a finitely additive probability on all clopen subsets of the infinite product in a natural way. This enabled them to state and prove weak versions of various convergence theorems. Dubins [7] showed how this measure can be extended to the open sets. Purves and Sudderth [13] then showed that every Borel set can be squeezed between an open set and a closed set. Therefore a strategy induces a finitely additive probability on all Borel sets of the product space, unique subject to some regularity conditions. Purves and Sudderth [13] then formulated and proved a special case of the strong law of large numbers for the i.i.d. situation and a version of the martingale convergence theorem. This was followed by a systematic study of almost sure convergence for independent strategies and martingales by Chen ([2] and [3]). In the same spirit we take up here the study of Markov chains.

In §2, after setting up the basic framework, we state some of the results from earlier work in the subject, which we shall need in the sequel. In §3, we obtain the strong Markov property which is fundamental for our theory. The classification of states is studied in §4 and the notions of recurrence and transience are examined in detail. §5 contains the very important blocks theorem which enables us to use in our theory known results on almost sure convergence in an i.i.d. setting. In §6 we study positive recurrence. §7 contains the strong law of large numbers for Markov chains and a few other related results on almost sure convergence. §8 shows the existence of finitely additive stationary initial distributions for any Markov chain. It needs to be remarked that in the countably additive case, a countably additive

Received by the editors April 21, 1980.

1980 *Mathematics Subject Classification.* Primary 60G05, 60J10, Secondary 60F15.

Key words and phrases. Finitely additive probabilities, strategy, Markov strategy, Markov chain, stationary transition, stationary initial distribution.

stationary initial distribution need not always exist. In §9 the problem of uniqueness is taken up. Although we do not have uniqueness in general, in many cases, the uniqueness of the countably additive part can be shown. This section also discusses the existence of countably additive stationary initial distributions for finitely additive Markov chains. In §10 we construct a special stationary initial distribution in case the chain has a positive recurrent state. We identify the limit in the strong law as the integral with respect to this special stationary initial distribution. Finally in §11 we prove the ergodicity of recurrent chains under a stationary initial distribution.

Generally, in our proofs familiar arguments are carried through with the help of stop rule techniques. We have drawn a lot of inspiration from the books of Chung [5] and Freedman [10] which have greatly influenced our presentation.

2. Preliminaries, notations and definitions. Let I be an arbitrary nonempty set equipped with the discrete topology. Let N be the set of positive integers. Let $H = I^N$ and equip H with the product topology. Let I^* be the set of all finite sequences of elements of I , including the empty one. A *strategy* σ is a function which associates with each $p \in I^*$ a finitely additive probability measure $\sigma(p)$, defined on all subsets of I . In [13], Purves and Sudderth have shown that a strategy σ induces a finitely additive probability measure on a field $\mathcal{Q}(\sigma)$ of subsets of H , including \mathfrak{B} , the σ -field of Borel subsets of H . This measure on $\mathcal{Q}(\sigma)$ is also denoted by σ . We shall now state some properties of the measure σ which we shall need.

The following notations will appear throughout. Let $p, q \in I^*$ and $h \in H$. pq is the element of I^* whose terms consist of the terms of p followed by the terms of q and ph is the element of H whose terms consist of the terms of p followed by the terms of h . If $A \subseteq H$, $Ap = \{h \in H \mid ph \in A\}$.

If $p \in I^*$, $\sigma[p]$, the *conditional strategy*, is defined by $\sigma[p](q) = \sigma(pq) \forall q \in I^*$.

A *stop rule* s is a function on H into N such that if h, h' belong to H and $h_i = h'_i$ for $i = 1, \dots, s(h)$, then $s(h) = s(h')$, where h_i, h'_i are the i th coordinates of h, h' respectively. For $h \in H$, $n \in N$ let $p_n(h) = (h_1, \dots, h_n)$, and, if s is a stop rule, $p_s(h) = p_n(h)$ where $n = s(h)$.

The following two propositions about σ , which we shall use, are proved in [13].

PROPOSITION 1. For every $A \in \mathfrak{B}$ and every stop rule s ,

$$\sigma(A) = \int \sigma[p_s(h)](Ap_s(h)) d\sigma(h).$$

PROPOSITION 2. Let A^1, A^2, \dots be sets in \mathfrak{B} .

- (i) If s is a stop rule, then $A^s \in \mathfrak{B}$, where $A^s = \{h \in H: h \in A^{s(h)}\}$.
 - (ii) If $A^1 \supseteq A^2 \supseteq \dots$ and $A = \bigcap_{n=1}^{\infty} A^n$, then $\sigma(A) = \inf_s \sigma(A^s)$.
 - (iii) If $A^1 \subseteq A^2 \subseteq \dots$ and $A = \bigcup_{n=1}^{\infty} A^n$, then $\sigma(A) = \sup_s \sigma(A^s)$.
- (The infimum and supremum are taken over all stop rules s .)

REMARK. Let s, s' be two stop rules such that $s(h) \geq s'(h)$ for all $h \in H$. Then under the hypothesis of (ii) of Proposition 2, $A^s \subseteq A^{s'}$ and under the hypothesis of

(iii) of Proposition 2, $A^s \supseteq A^{s'}$. Therefore it is enough to take the infimum and supremum in the assertion of (ii) and (iii) respectively, over all stop rules s such that $s(h) \geq s_0(h) \forall h \in H$ where s_0 is a fixed stop rule.

DEFINITION. A strategy σ is called a Markov strategy if $\sigma(i_1, \dots, i_n)$ depends only on n and i_n for all $n \in N$ and for all $i_1, \dots, i_n \in I$. In such a case we say that the strategy satisfies the Markov property.

DEFINITION. A Markov strategy σ is called a Markov strategy with stationary transition probabilities if $\sigma(i_1, \dots, i_n) = \sigma(i_n)$ for all $i_1, \dots, i_n \in I$ and $n \in N$ (i.e. $\sigma(i_1, \dots, i_n)$ depends only on i_n and not on n).

The measure induced on $(H, \mathcal{Q}(\sigma))$ by a Markov strategy σ is called a Markov measure and the sequence of coordinate maps is said to form a Markov chain. Since we shall never have occasion to talk about a Markov strategy that does not have stationary transition probabilities, whenever we talk of a Markov strategy, it is to be understood to have stationary transition probabilities.

For any strategy, the measure σ_0 which σ associates with the empty sequence will be called the initial distribution of the strategy. A Markov strategy σ is completely specified if we specify σ_0 and $\sigma(i)$, $i \in I$. Since for many of our results it is immaterial what the initial distribution is, we shall in such cases denote the Markov strategy by $\{\sigma(i)\}_{i \in I}$. When the initial distribution σ_0 is also important, we shall denote the Markov strategy by $\{\sigma_0, \{\sigma(i)\}_{i \in I}\}$.

3. The strong Markov property. Many of the important results about Markov chains need a stronger version of the Markov property. In order to obtain that, we first need to strengthen Proposition 1 of §2.

DEFINITION. An incomplete stop rule t is a function on H into $N \cup \{\infty\}$ such that if $h, h' \in H$ and $t(h) \in N$, then $h_i = h'_i$ for $i = 1, \dots, t(h)$ implies that $t(h) = t(h')$.

LEMMA 1. Let t be an incomplete stop rule and s a stop rule. Then $s_1 = \min(t, s)$ is a stop rule.

Lemma 1 is well known and can be found in [8]. We omit its proof since it is straightforward.

LEMMA 2. Let t be an incomplete stop rule and s a stop rule. Let $A \in \mathfrak{B}$ be such that $A \subseteq \{h \in H | t(h) < s(h)\}$. Then, for any strategy σ ,

$$\sigma(A) = \int_{\{t < s\}} \sigma[p_t(h)](Ap_t(h)) d\sigma(h),$$

the integral above being taken over $\{h \in H | t(h) < s(h)\}$.

PROOF. Define s_1 as in Lemma 1. Then by Proposition 1 of §2, we know that $\sigma(A) = \int \sigma[p_{s_1}(h)](Ap_{s_1}(h)) d\sigma(h)$. Observe that since $A \subseteq \{h \in H | t(h) < s(h)\}$, $Ap_{s_1}(h) = \emptyset$ if h is such that $t(h) > s(h)$; and if $t(h) < s(h)$, $s_1(h) = t(h)$ and therefore $p_{s_1}(h) = p_t(h)$. Hence Lemma 2 follows.

PROPOSITION 3. Let t be an incomplete stop rule and $A \in \mathfrak{B}$ be such that $A \subseteq \{h \in H | t(h) < \infty\}$. Then, for any strategy σ ,

$$\sigma(A) = \int_{\{t < \infty\}} \sigma[p_t(h)](Ap_t(h)) d\sigma(h).$$

PROOF. Let $A^n = \{h \in H | h \in A \text{ and } t(h) \leq n\}$, $n \in N$. Clearly $A^1 \subseteq A^2 \subseteq \dots$ and $\bigcup_{n=1}^{\infty} A^n = A$. Further $A^s = A \cap \{t \leq s\}$ for every stop rule s . Hence by Proposition 2 of §2 and by Lemma 2 of this section,

$$\sigma(A) = \sup_s \sigma(A^s) = \sup_s \int_{\{t \leq s\}} \sigma[p_t(h)](A^s p_t(h)) d\sigma(h).$$

Observe now that, for h such that $t(h) \leq s(h)$, $A^s p_t(h) = Ap_t(h)$.

$$\begin{aligned} \therefore \sigma(A) &= \int_{\{t < \infty\}} \sigma[p_t(h)](Ap_t(h)) d\sigma(h) \\ &\quad - \inf_s \int_{\{\infty > t > s\}} \sigma[p_t(h)](Ap_t(h)) d\sigma(h). \end{aligned}$$

Further, by Proposition 2 of §2,

$$\sigma(\{t < \infty\}) = \sigma\left(\bigcup_{n=1}^{\infty} \{t \leq n\}\right) = \sup_s \sigma(\{t \leq s\}).$$

Hence $0 \leq \inf_s \int_{\{\infty > t > s\}} \sigma[p_t(h)](Ap_t(h)) d\sigma(h) \leq \inf_s \sigma(\{\infty > t > s\}) = 0$. Consequently Proposition 3 follows.

DEFINITION. Let $A \in \mathfrak{B}$ and let t be an incomplete stop rule. Say that A is conditionally determined given t if (i) $A \subseteq \{t < \infty\}$ and (ii) there exists $B \in \mathfrak{B}$ such that $Ap_t(h) = B$ for all h such that $t(h) < \infty$.

EXAMPLES. (i) Let $i \in I$. Let

$$\begin{aligned} t(h) &= n \quad \text{if } h_m \neq i, 1 \leq m \leq n-1 \text{ and } h_n = i, \\ &= \infty \quad \text{if } h_n \neq i \text{ for all } n \in N. \end{aligned}$$

This t will be called the incomplete stop rule corresponding to the first occurrence of i . Let $A = \{h \in H | h_n = i \text{ for some } n \in N\}$. Then A is conditionally determined given t , with $B = H$.

(ii) Let t be the incomplete stop rule corresponding to the first occurrence of i and let $A = \{h \in H | h_n = i \text{ for infinitely many } n \in N\}$. Then A is conditionally determined given t , with $B = A$.

From now onwards, for the rest of the paper, σ will stand for a Markov strategy with stationary transition probabilities.

THEOREM 4 (STRONG MARKOV PROPERTY). Suppose A is conditionally determined given t . Suppose further that there exists $i \in I$ such that $h_{t(h)} = i$ for all h such that $t(h) < \infty$. Then $\sigma(A) = \sigma[i](B) \cdot \sigma(\{t < \infty\})$ where B is such that $Ap_t(h) = B$ for all $h \in \{t < \infty\}$.

PROOF. It follows from Proposition 3 that

$$\begin{aligned}\sigma(A) &= \int_{\{t < \infty\}} \sigma[p_t(h)](Ap_t(h)) \, d\sigma(h) \\ &= \int_{\{t < \infty\}} \sigma[i](B) \, d\sigma(h), \quad \text{by the hypotheses and the Markov property,} \\ &= \sigma[i](B)\sigma(\{t < \infty\}).\end{aligned}$$

4. Classification of the state space. We shall refer to I as the state space of the Markov chain and to elements of I as the states of the Markov chain.

For $j \in I$ and $n \in N$, let A_j^n denote the set of all $h \in H$ such that $h_n = j$. Let $B_j^n = \bigcup_{k=1}^n A_j^k$. Let $A_j = \bigcup_{n=1}^{\infty} A_j^n = \bigcup_{n=1}^{\infty} B_j^n$. Let $i, j \in I$.

DEFINITION. We shall say that i weakly leads to j (denoted by $i \xrightarrow{w} j$) in case $\sigma[i](A_j) > 0$. We shall denote $\sigma[i](A_j)$ by f_{ij}^* .

DEFINITION. We shall say that i strongly leads to j (denoted by $i \xrightarrow{s} j$) in case $\sigma[i](A_j^n) > 0$ for some $n \in N$.

REMARK. Always, if $i \xrightarrow{s} j$ then $i \xrightarrow{w} j$. If we assume that σ is countably additive, i.e. if $\sigma(p)$ is a countably additive probability measure for each $p \in I^*$, then $i \xrightarrow{w} j$ implies that $i \xrightarrow{s} j$.

EXAMPLE 1. Let $I = N$. Let σ be a stationary Markov strategy such that $\sigma(\emptyset) = \lambda$, where λ is an arbitrary finitely additive probability measure on I , $\sigma(1) = \gamma$, where γ is a purely finitely additive probability measure, and $\sigma(n+1) = \partial_n$, the Dirac measure at n , $n \in N$.

It is easy to see that in the above example $1 \xrightarrow{w} 1$, but $1 \not\xrightarrow{s} 1$ is not true. Thus the notions of i strongly leading to j and i weakly leading to j are not in general equivalent.

PROPOSITION 1. $i \xrightarrow{w} j$ iff, for some stop rule s , $\sigma[i](A_j^s) > 0$.

PROOF. Since $A_j^s \subseteq A_j$ for all stop rules, $\sigma[i](A_j^s) > 0$ for some stop rule s implies that $f_{ij}^* > 0$. Conversely, suppose that $f_{ij}^* > 0$. By Proposition 2 of §2,

$$f_{ij}^* = \sigma[i](A_j) = \sigma[i]\left(\bigcup_{n=1}^{\infty} B_j^n\right) = \sup_s \sigma[i](B_j^s).$$

Hence there exists a stop rule s such that $\sigma[i](B_j^s) > 0$. Let t be the incomplete stop rule corresponding to the first occurrence of j and define $s_1 = \min(t, s)$. Clearly $B_j^s \subseteq A_j^{s_1}$. Therefore $\sigma[i](A_j^{s_1}) > 0$ and thus the proof of the proposition is complete.

REMARK. $f_{ij}^* = \sup_s \sigma[i](A_j^s)$.

DEFINITION. Let s_1, s_2 be stop rules. s_1 composed with s_2 , denoted by $s_1 * s_2$, is the stop rule defined by

$$s_1 * s_2(h) = s_1(h) + s_2(h_{s_1(h)+1}, h_{s_1(h)+2}, \dots), \quad h \in H.$$

PROPOSITION 2 (CHAPMAN-KOLMOGOROV EQUATIONS). Let $i, j \in I$ and let s_1, s_2 be stop rules. Let μ be the measure on I defined by $\mu(J) = \sigma[i](\{h \in H \mid h_{s_1(h)} \in J\})$

for every $J \subseteq I$. Then

$$\sigma[i](A_j^{s_1 * s_2}) = \int \sigma[k](A_j^{s_2}) d\mu(k).$$

PROOF. By Proposition 1 of §2, we have

$$\begin{aligned} \sigma[i](A_j^{s_1 * s_2}) &= \int \sigma[ip_{s_1}(h)](A_j^{s_1 * s_2} p_{s_1}(h)) d\sigma[i](h) \\ &= \int \sigma[h_{s_1}(h)](A_j^{s_2}) d\sigma[i](h). \end{aligned}$$

The proof of the proposition is completed by applying the change of variable theorem (see Lemma 8, page 182 of [9]).

COROLLARY 3. For $i, j, k \in I$,

(a) $i \xrightarrow{w} j, j \xrightarrow{w} k$ implies $i \xrightarrow{w} k$.

(b) $i \xrightarrow{s} j, j \xrightarrow{s} k$ implies $i \xrightarrow{s} k$.

(c) $f_{ik}^* > f_{ij}^* f_{jk}^*$.

PROOF. It follows from the Chapman-Kolmogorov equations that $\sigma[i](A_k^{s_1 * s_2}) > \sigma[i](A_j^{s_1}) \cdot \sigma[j](A_k^{s_2})$ for $i, j, k \in I$ and all stop rules s_1, s_2 . Assertion (a) now follows from Proposition 1 of this section. In order to prove (b), observe that if $s_1 \equiv n_1$ and $s_2 \equiv n_2$, $n_1, n_2 \in N$, then $s_1 * s_2 \equiv n_1 + n_2$ and use it in the above inequality. Assertion (c) follows by taking supremum over all stop rules s_1, s_2 in both sides of the above inequality.

DEFINITION. A state i is said to strongly communicate with j in case $i \xrightarrow{s} j$ and $j \xrightarrow{s} i$. In such a case we write $i \leftrightarrow j$.

A state i is said to weakly communicate with j in case $i \xrightarrow{w} j$ and $j \xrightarrow{w} i$. In such a case we write $i \leftrightarrow^w j$.

On the set of all states i such that $i \leftrightarrow^w i$ ($i \leftrightarrow i$), \leftrightarrow^w (\leftrightarrow) is an equivalence relation. We thus get a partition of such states into equivalence classes. By considering $\{i\}$ for i such that $i \xrightarrow{w} i$ ($i \leftrightarrow i$) is not true, we get a partition of the state space into classes called weakly (strongly) communicating classes. Since $i \leftrightarrow^w j$ implies that $i \leftrightarrow j$, every weakly communicating class is a union of strongly communicating classes.

DEFINITION. A state i is called recurrent if $f_{ii}^* = 1$. Otherwise it is called transient. Let $G_j = \limsup_{n \rightarrow \infty} A_j^n$ and $g_{ij} = \sigma[i](G_j)$, for $i, j \in I$.

PROPOSITION 4. $g_{ij} = f_{ij}^* g_{jj}$.

PROOF. Let t be the incomplete stop rule corresponding to the first occurrence of j . Then G_j is conditionally determined given t and $G_j p_t(h) = G_j$ for all $h \in \{t < \infty\}$. Hence by the strong Markov property

$$g_{ij} = \sigma[i](G_j) = \sigma[j](G_j) \cdot \sigma[i](\{t < \infty\}) = f_{ij}^* g_{jj}.$$

PROPOSITION 5.

$$\begin{aligned} g_{ij} &= f_{ij}^* \quad \text{if } j \text{ is recurrent,} \\ &= 0 \quad \text{if } j \text{ is transient.} \end{aligned}$$

PROOF. Let $G_j^m = \{h \mid \text{at least } m \text{ coordinates of } h \text{ are } j\}$ and let $g_{ij}(m) = \sigma[i](G_j^m)$.

Let t be the incomplete stop rule corresponding to the first occurrence of j . For $m \in N$, G_j^{m+1} is conditionally determined given t and $G_j^{m+1}p_t(h) = G_j^m$ for $h \in \{t < \infty\}$. Thus, by the strong Markov property, we have, for all $m \in N$,

$$g_{ij}(m+1) = \sigma[j](G_j^m)\sigma[i](\{t < \infty\}) = f_{ij}^*g_{ij}(m).$$

A repeated application of the above equality gives us $g_{ij}(m+1) = f_{ij}^*(f_{ij}^*)^m$ for all $m \in N$. If j is transient, by Proposition 2 of §2,

$$\begin{aligned} g_{ij} &= \inf_s \sigma[i](G_j^s) \leq \inf_{m \in N} \sigma[i](G_j^m) \\ &< \inf_{m \in N} f_{ij}^*(f_{ij}^*)^m = 0. \end{aligned}$$

Suppose now that j is recurrent. It follows from what we have proved that $g_{ij}(m) = f_{ij}^*$ for all $m \in N$. We shall prove by an induction on the structure of stop rule s that $\sigma[i](G_j^s) = f_{ij}^*$ for all stop rules s . (For an account of the hierarchy of stop rules see [8].) In view of Proposition 2 of §2, this is enough to prove our proposition. We have already shown that $\sigma[i](G_j^s) = f_{ij}^*$ for all stop rules of structure zero. Assume that for all $i \in I$ and stop rules s of structure $< \alpha$, we have shown that $\sigma[i](G_j^s) = f_{ij}^*$. Now let s be a stop rule of structure α .

Recall that t is the incomplete stop rule corresponding to the first occurrence of j . By Proposition 3 of §3,

$$\begin{aligned} \sigma[i](G_j^{s+1}) &= \int_{\{t < \infty\}} \sigma[ip_t(h)](G_j^{s+1}p_t(h)) d\sigma[i](h) \\ &= \int_{\{t < \infty\}} \sigma[j](G_j^s) d\sigma[i](h) \end{aligned}$$

where s_h is the stop rule defined by $s_h(h') = s(p_t(h)h')$. For all h such that $t(h) < \infty$, s_h is a stop rule of structure $< \alpha$. Hence by the induction hypothesis the integrand is identically equal to $f_{ij}^* = 1$. Therefore $\sigma[i](G_j^{s+1}) = \sigma[i](\{t < \infty\}) = f_{ij}^*$. But $f_{ij}^* \geq \sigma[i](G_j^s) \geq \sigma[i](G_j^{s+1}) = f_{ij}^*$. Hence the proof of the proposition is complete.

REMARK. In the course of the proof we have also proved that for any $i \in I$, if $Y(i)$ is mapping on H defined by $Y(i)(h) = \text{number of coordinates in } h \text{ which are equal to } i$, then the distribution of $Y(i)$ is geometric under $\sigma[i]$.

The following is an immediate corollary of Proposition 5.

COROLLARY 6.

$$\begin{aligned} g_{ii} &= 1 && \text{if } i \text{ is recurrent,} \\ &= 0 && \text{if } i \text{ is transient.} \end{aligned}$$

The rest of the section is mainly devoted to showing that in a weakly communicating class either all states are transient or all states recurrent.

LEMMA 7. If i is recurrent and $i \xrightarrow{w} j$, then $j \xrightarrow{w} i$.

PROOF. Suppose on the contrary that there exists a j such that $f_{ij}^* > 0$ but $f_{ji}^* = 0$. Consequently, $g_{ji} < f_{ji}^* = 0$. Then

$$g_{ii} = \sigma[i](G_i) = \sigma[i](G_i \cap A_j) + \sigma[i](G_i \cap A_j^c).$$

Now $\sigma[i](G_i \cap A_j^c) \leq \sigma[i](A_j^c) = 1 - f_{ij}^* < 1$ since $f_{ij}^* > 0$.

Also $G_i \cap A_j$ is conditionally determined given t , the time of first occurrence of j . Further $(G_i \cap A_j)p_t(h) = G_i$ for all $h \in \{t < \infty\}$. Hence by the strong Markov property

$$\sigma[i](G_i \cap A_j) = \sigma[j](G_i) \cdot \sigma[i](\{t < \infty\}) = g_{ji}f_{ij}^* = 0.$$

Hence $g_{ii} < 1$, a contradiction, since by Corollary 6, $g_{ii} = 1$ as i is recurrent. This proves the lemma.

PROPOSITION 8. If i is recurrent, $i \neq j$ and $i \xrightarrow{w} j$, then $g_{ij} = 1$.

PROOF. Let s' and s be stop rules such that $s'(h) < s(h)$ for all $h \in H$. Then

$$\begin{aligned} \sigma[i](\{h: h_n = i \text{ for some } n > s(h) \text{ and } h_n \neq j \text{ for all } n > s'(h)\}) \\ < \sigma[i](\{h: (\exists m > s(h))(h_n \neq j \text{ for } s'(h) < n \leq s(h), h_n \neq i \text{ for} \\ s(h) < n < m, h_n = i, h_n \neq j \text{ for } n > m)\}). \end{aligned}$$

Let $C_{s,s'}$ denote the event on the right side of the above inequality. Define

$$\begin{aligned} t(h) &= n && \text{if } n > s(h), h_n = i, h_m \neq j \text{ for } s'(h) < m \leq s(h) \\ &&& \text{and } h_m \neq i \text{ for } s(h) < m < n, \\ &= \infty && \text{if no such } n \text{ exists.} \end{aligned}$$

Then t is an incomplete stop rule and $C_{s,s'}$ is conditionally determined given t , with $C_{s,s'}p_t(h) = A_j^c$ for all $h \in \{t < \infty\}$. Hence by the strong Markov property,

$$\begin{aligned} \sigma[i](C_{s,s'}) &= \sigma[i](A_j^c)\sigma[i](\{t < \infty\}) \\ &= (1 - f_{ij}^*)\sigma[i](\{h: h_n = i \text{ for some } n > s(h) \text{ and } h_m \neq j, s'(h) < m \leq s(h)\}). \end{aligned}$$

Taking infimum over all stop rules s such that $s(h) > s'(h)$ for all h on both sides of the last mentioned inequality and using Proposition 2 of §2, we get

$$\begin{aligned} \sigma[i](\{h: h_n = i \text{ for infinitely many } n, h_n \neq j \text{ for } n > s'(h)\}) \\ < (1 - f_{ij}^*)\sigma[i](\{h: h_n = i \text{ for infinitely many } n, h_n \neq j \text{ for } n > s'(h)\}). \end{aligned}$$

Since $f_{ij}^* > 0$, it follows that

$$\sigma[i](\{h: h_n = i \text{ for infinitely many } n, h_n \neq j \text{ for } n > s'(h)\}) = 0.$$

This being true for every stop rule s' , by Proposition 2 of §2 again, $\sigma[i](G_i \cap G_j^c) = 0$.

But $\sigma[i](G_i) = 1$ (i is recurrent).

So $\sigma[i](G_j) = \sigma[i](G_i \cap G_j) = 1$.

THEOREM 9. Let i be recurrent. Suppose $i \xrightarrow{w} j$. Then j is recurrent and $g_{ij} = g_{ji} = 1$. Consequently, all states in the weakly communicating class containing i are recurrent.

PROOF. Under the hypotheses of the theorem, $g_{ij} = 1$ by Proposition 8. Therefore, by Proposition 4, $g_{ji} = 1$. Hence j is recurrent by Corollary 6. By Lemma 7,

$j \xrightarrow{w} i$. Thus, once again by Proposition 8, we get $g_{ji} = 1$. The proof of the theorem is therefore complete.

5. The blocks theorem. Let i be a recurrent state. We fix this i for the rest of the section. For each $n \in N$, let $t(n)$ be the incomplete stop rule corresponding to the n th occurrence of i . Let F be the set of all nonempty finite sequences of elements of I whose last coordinate is i and none of the other coordinates is i . Let $\Omega = F^N$. Equip Ω with the product of discrete topologies and let \mathcal{F} be the Borel σ -field on Ω . On G_i , we define a sequence $\{\beta_n\}$ of functions into F called the block variables as follows:

$$\beta_1(h) = p_{t(1)}(h) \quad \text{and} \quad \beta_{n+1}(h) = (h_{t(n)+1}, \dots, h_{t(n+1)})$$

for all $h \in G_i$ and $n \in N$. Let Φ be the mapping on G_i into Ω defined by $\Phi(h) = (\beta_1(h), \beta_2(h), \dots)$. Clearly Φ is 1-1, onto Ω and is in fact a (topological) homeomorphism (G_i has the relative topology). To see the last part it is enough to look at the forward image and inverse image of sets from the canonical base of clopen sets and observe that they are open.

Let X be a topological space. Let \mathcal{Q} be the σ -field of Borel subsets of X .

DEFINITION. A finitely additive probability P on (X, \mathcal{Q}) is said to be regular just in case

- (i) for every open set U in X , $P(U) = \sup\{P(K): K \text{ clopen}, K \subseteq U\}$ and
- (ii) for every Borel set A in X , $P(A) = \inf\{P(U): U \text{ open}, U \supseteq A\}$.

LEMMA 1. Let P and P' be regular measures on topological spaces X and X' , with Borel σ -fields \mathcal{Q} , \mathcal{Q}' respectively. Let Ψ be a homeomorphism of X onto X' . Assume further that for all K clopen in X' , $P'(K) = P(\Psi^{-1}(K))$. Then $P'(A) = P(\Psi^{-1}(A))$ for all $A \in \mathcal{Q}'$.

The proof is omitted as it is straightforward.

Let γ be a measure on F defined by

$$\gamma(D) = \sigma[i](\beta_1^{-1}(D)) \quad \text{for all } D \subseteq F.$$

γ is clearly finitely additive and is a probability, since i being recurrent, $\sigma[i](G_i) = 1$. Let π be the i.i.d. strategy on F induced by γ , i.e., $\pi(p) = \gamma$ for all finite sequences p of elements of F . As usual let π denote the measure induced by this i.i.d. strategy on (Ω, \mathcal{F}) .

THEOREM 2 (BLOCKS THEOREM). For each $B \in \mathcal{F}$,

$$\pi(B) = \sigma[i](\Phi^{-1}(B)).$$

PROOF. Since all strategic measures are regular [13], π and $\sigma[i]$ are regular. Since $\sigma[i](G_i) = 1$, restriction of $\sigma[i]$ to G_i is regular. Further Φ , as defined in the beginning of this section, is a homeomorphism of G_i and Ω . Hence, in view of Lemma 1, it is enough to show that $\pi(K) = \sigma[i](\Phi^{-1}(K))$ for all K clopen in Ω . We shall prove this by induction on the structure of the clopen set. (See [8] for an account of the hierarchy of clopen sets.)

If $K = \emptyset$, $\pi(K) = 0 = \sigma[i](\Phi^{-1}(K))$ and if $K = \Omega$, $\pi(K) = 1 = \sigma[i](G_i) = \sigma[i](\Phi^{-1}(\Omega))$ since i is recurrent. So the result is true for clopen sets of structure zero.

Assume that we have proved the result for all clopen sets of structure $< \alpha$. Let K now be a clopen set of structure α . Then

$$\begin{aligned}
 \pi(K) &= \int \pi(Kx) d\gamma(x), \quad \text{by Proposition 1 of §2 and change of variable theorem,} \\
 &= \int \sigma[i](\Phi^{-1}(Kx)) d\gamma(x), \quad \text{by induction hypothesis,} \\
 &= \int_{G_i} \sigma[i](\Phi^{-1}(K\beta_1(h))) d\sigma[i](h), \quad \text{by change of variable theorem,} \\
 &= \int_{G_i} \sigma[i](\Phi^{-1}(K)) p_{i(1)}(h) d\sigma[i](h) \\
 &= \int_{\{i(1) < \infty\}} \sigma[i](\Phi^{-1}(K)) p_{i(1)}(h) d\sigma[i](h) \\
 &= \sigma[i](\Phi^{-1}(K)), \quad \text{by Proposition 3 of §3.}
 \end{aligned}$$

Hence the theorem is proved.

6. Positive recurrence. We begin with a lemma on finitely additive integration.

Let X be an arbitrary nonempty set, \mathcal{Q} a σ -field of subsets of X and P a finitely additive probability on X . A random variable β is an extended real valued measurable function on (X, \mathcal{Q}) .

Following [9], for an unbounded nonnegative random variable β , by $\int \beta dP$ we shall mean $\sup_n \int \beta \wedge n dP$ where $\beta \wedge n = \min(\beta, n)$. For a random variable which is not necessarily nonnegative, we define $\int \beta dP = \int \beta^+ dP - \int \beta^- dP$, provided the right side makes sense, where $\beta^+ = \max(\beta, 0)$ and $\beta^- = \max(-\beta, 0)$.

LEMMA 1. *If β is a nonnegative, extended integer valued random variable on (X, \mathcal{Q}, P) , then*

$$\int \beta dP = \sum_{n=1}^{\infty} P(\{\beta \geq n\}).$$

PROOF. It can easily be verified that, for $m = 1, 2, \dots$,

$$\int (\beta \wedge m) dP = \sum_{n=1}^m P(\{\beta \geq n\}).$$

The result follows by taking the supremum over m .

The next lemma is a finitely additive version of Wald's identity.

Let X be an arbitrary nonempty set. A sequence $\{Z_n\}$ of functions on X^N is called a sequence of identical coordinate mappings if there exists a function Z on X such that $Z_n(x_1, \dots, x_n, \dots) = Z(x_n)$ for all $(x_1, \dots, x_n, \dots) \in X^N$ and for all n .

LEMMA 2 (WALD'S IDENTITY). *Let ρ be an i.i.d. strategy on X and let $\{Z_n\}$ be a sequence of identical coordinate mappings which are bounded below by the constant 1. Let τ be an incomplete stop rule and $S_\tau = \sum_{n=1}^{\tau} Z_n$. Then $\int S_\tau d\rho = (\int Z_1 d\rho)(\int \tau d\rho)$.*

PROOF. We may assume that $\int Z_1 d\rho < \infty$, for otherwise the result is trivial. Suppose first that $\tau \leq m$ and let

$$\begin{aligned} Y_n &= 1 & \text{if } \tau \geq n, \\ &= 0 & \text{if } \tau < n. \end{aligned}$$

Then $S_\tau = \sum_{n=1}^m Z_n Y_n$. Consequently,

$$\begin{aligned} \int S_\tau d\rho &= \sum_{n=1}^m \int Z_n Y_n d\rho \\ &= \sum_{n=1}^m \left(\int Z_n d\rho \right) \left(\int Y_n d\rho \right), \quad \text{because } \tau \text{ is an incomplete stop rule,} \\ &= \left(\int Z_1 d\rho \right) \left(\sum_{n=1}^m \rho(\{\tau \geq n\}) \right) \\ &= \left(\int Z_1 d\rho \right) \left(\int \tau d\rho \right). \end{aligned}$$

For the case when τ is unbounded, observe that $S_\tau \wedge n \leq S_{\tau \wedge n}$ because $\tau \leq n$ implies $S_{\tau \wedge n} = S_\tau$ and $\tau > n$ implies $S_{\tau \wedge n} = S_n \geq n$. Hence,

$$\begin{aligned} \int S_\tau d\rho &= \sup_n \int (S_\tau \wedge n) d\rho = \sup_n \int S_{\tau \wedge n} d\rho \\ &= \sup_n \left(\int Z_1 d\rho \right) \left(\int (\tau \wedge n) d\rho \right) = \left(\int Z_1 d\rho \right) \left(\int \tau d\rho \right). \end{aligned}$$

This completes the proof.

REMARK. A much more general form of Wald's identity has been proved in [15]. The proofs given above of Lemma 1 and the special case of Wald's identity were suggested by the referee.

Returning to Markov chains, for $i, j \in I$, let m_{ij} stand for $\int t d\sigma[i]$ where t is the time of first occurrence of a j .

DEFINITION. A recurrent state j is called positive recurrent in case $m_{jj} < \infty$. It is called null recurrent if $m_{jj} = \infty$.

REMARK. If $m_{jj} < \infty$, by Lemma 1, $\sum_{n=1}^{\infty} \sigma[j](\{t \geq n\}) < \infty$, so $\sigma[j](\{t \geq n\}) \rightarrow 0$ as $n \rightarrow \infty$ and hence j is recurrent.

Let i be recurrent and $i \neq j$. Let F be the space of i -blocks as defined in §5. Let D be the set of elements of F having at least one j .

Define θ on G_i as follows:

$$\begin{aligned} \theta(h) &= n & \text{if } \beta_m(h) \notin D, 1 \leq m < n, \text{ and } \beta_n(h) \in D, \\ &= \infty & \text{if no such } n \text{ exists.} \end{aligned}$$

Let $t(1)$ and t be the time of first occurrences of i and j respectively.

LEMMA 3. Suppose $i \xrightarrow{w} j$. Then $\sigma[i](\{t < t(1)\}) > 0$ and $\sigma[j](\{t(1) < t\}) > 0$.

PROOF. Since i is recurrent and $i \xrightarrow{w} j$ implies that j is recurrent and $j \xrightarrow{w} i$, it is enough to show that $\sigma[i](\{t < t(1)\}) > 0$. Clearly $\beta_1^{-1}(D) = \{t < t(1)\}$. If $\sigma[i](\beta_1^{-1}(D))$ were zero, by the blocks theorem and the Borel-Cantelli lemma [13], it would follow that $g_{ij} = \sigma[i](G_j) = 0$. This is a contradiction because $g_{ij} = 1$. Hence the lemma is proved.

LEMMA 4. If $i \xrightarrow{w} j$, $\int \theta d\sigma[i] < \infty$.

PROOF. Under the hypothesis, by Lemma 3, $\sigma[i](\beta_1^{-1}(D))$ (say equal to p) is positive. Observe that

$$\begin{aligned} \int \theta d\sigma[i] &= \sum_{n=1}^{\infty} \sigma[i](\{\theta > n\}) \\ &= \sum_{n=1}^{\infty} \sigma[i]\left(\bigcap_{k=1}^{n-1} \beta_k^{-1}(D^c)\right) \\ &= \sum_{n=1}^{\infty} (1-p)^{n-1}, \text{ by the blocks theorem,} \\ &= 1/p < \infty, \text{ because } p > 0. \end{aligned}$$

Let $\{\lambda_n\}$ be the sequence of functions on G_i defined by $\lambda_1(h) = t(1)(h)$ and $\lambda_{n+1}(h) = t(n+1)(h) - t(n)(h)$, $n \in N$, where $t(n)$ is the time of n th occurrence of i . Let S_θ be defined on G_i by $S_\theta(h) = \sum_{n=1}^{\theta(h)} \lambda_n(h)$.

LEMMA 5. If $i \xrightarrow{w} j$, $\int S_\theta d\sigma[i] = m_{ij} + m_{ji}$.

PROOF. Observe that $t(h) \leq S_\theta(h)$ for all $h \in G_i$ where t is the time of first occurrence of a j . Define on G_i the function r by

$$\begin{aligned} r(h) &= S_\theta(h) - t(h) \text{ if } t(h) < \infty, \\ &= \infty \text{ otherwise.} \end{aligned}$$

Clearly r is a nonnegative extended integer valued function such that

$$S_\theta(h) = t(h) + r(h) \text{ for all } h \in G_i.$$

Therefore $\int S_\theta d\sigma[i] = \int t d\sigma[i] + \int r d\sigma[i]$. Since by definition $\int t d\sigma[i] = m_{ij}$, it is enough to show that $\int r d\sigma[i] = m_{ji}$. Now

$$\begin{aligned} \int r d\sigma[i] &= \sum_{n=1}^{\infty} \sigma[i](\{r > n\}) \\ &= \sum_{n=1}^{\infty} \sigma[i](\{t < \infty \text{ and } r > n\}), \text{ since } f_{ij}^* = 1, \\ &= \sum_{n=1}^{\infty} \sigma[i](\{t < \infty \text{ and } S_\theta > t + n\}). \end{aligned}$$

The event A_n (say) in the n th term of the summation is conditionally determined given t and $A_{np_i}(h) = \{t(1) > n\}$ for all $h \in \{t < \infty\}$ and $n \in N$. Therefore by the strong Markov property,

$$\begin{aligned} \int r d\sigma[i] &= \sum_{n=1}^{\infty} \sigma[j](\{t(1) > n\}) \sigma[i](\{t < \infty\}) \\ &= \sum_{n=1}^{\infty} \sigma[j](\{t(1) > n\}) = m_{ji}. \end{aligned}$$

PROPOSITION 6. If i is recurrent and $i \xrightarrow{w} j$, then $m_{ii} < m_{ij} + m_{ji}$ and $m_{jj} < m_{ij} + m_{ji}$.

PROOF. As before it is enough to show that $m_{ii} \leq m_{ij} + m_{ji}$. The case $i = j$ being trivial, we can also assume that $i \neq j$. By Lemma 5, $\int S_\theta d\sigma[i] = m_{ij} + m_{ji}$. But $t(1)(h) \leq S_\theta(h)$, $h \in G_i$.

Therefore $m_{ii} = \int t(1) d\sigma[i] \leq \int S_\theta d\sigma[i] = m_{ij} + m_{ji}$.

REMARK. It therefore follows that if i is recurrent and $i \xrightarrow{w} j$, and further $m_{ij} < \infty$ and $m_{ji} < \infty$, then i and j are both positive recurrent.

THEOREM 7. Let i be a positive recurrent state and $i \xrightarrow{w} j$. Then j is positive recurrent. Further $m_{ij} < \infty$ and $m_{ji} < \infty$. Consequently, in a weakly communicating class either all states are positive recurrent or none is.

PROOF. We assume that $i \neq j$ since otherwise the result is trivial. In view of Proposition 6, it is enough to prove that $m_{ij} + m_{ji} < \infty$. By Lemma 5, this is equivalent to proving that $\int S_\theta d\sigma[i] < \infty$. Using the blocks theorem and the change of variable theorem it is easy to see that $\int S_\theta d\sigma[i] = \int S_\theta \circ \Phi^{-1} d\pi$, where Φ is the homeomorphism from G_i onto F^N defined in §5. Note that $\theta \circ \Phi^{-1}$ is an incomplete stop rule and $\{\lambda_n \circ \Phi^{-1}\}$ is a sequence of identical coordinate mappings. Therefore by Wald's identity and then the change of variable theorem,

$$\int S_\theta d\sigma[i] = \left(\int \lambda_1 d\sigma[i] \right) \left(\int \theta d\sigma[i] \right).$$

Now $\int \lambda_1 d\sigma[i]$ is finite since i is positive recurrent and $\int \theta d\sigma[i] < \infty$ by Lemma 4. Therefore the theorem is proved.

REMARK. If $m_{ij} < \infty$, as before we can show that $\sigma[i](\{t \geq n\}) \rightarrow 0$ as $n \rightarrow \infty$. Therefore $\lim_n \sigma[i](\{t < n\}) = 1$. Hence for some n , $\sigma[i](A_j^n) > 0$. Consequently $i \xrightarrow{s} j$. Therefore a positive recurrent weakly communicating class is necessarily strongly communicating. In particular $i \xleftrightarrow{s} i$ if i is positive recurrent.

We conclude this section by giving an example of a Markov chain corresponding to a strategy σ which is not countably additive where the state space is one positive recurrent class.

EXAMPLE. Let $I = N$. Let $\sigma(\emptyset)$ be an arbitrary finitely additive probability on I , $\sigma(1) = \sum_{n=2}^{\infty} p_n \partial_n + (1 - \sum_{n=2}^{\infty} p_n) \gamma$, where $0 < p_n < 1$ for all $n \geq 2$ and $\sum_{n=2}^{\infty} p_n < 1$ and γ is a purely finitely additive measure on I , $\sigma(n+1) = \partial_1$, $n \in N$.

Clearly I is one weakly communicating class and is positive recurrent since $m_{11} = 2$.

7. The Strong Law of Large Numbers (SLLN). In order to prove the Strong Law of Large Numbers for Markov chains, we need the following version of the Strong Law of Large Numbers for i.i.d. strategies.

SLLN for i.i.d. strategies. Let ρ be an i.i.d. strategy on X and $\{Z_n\}$ a sequence of identical coordinate mappings. If $\int Z_n d\rho = \mu$ for all $n \in N$ where μ is a real number, then $\rho(\{w: S_n(w)/n \rightarrow \mu\}) = 1$ where $S_n(w) = \sum_{k=1}^n Z_k(w)$, $w \in X^N$. If $\int Z_n d\rho$ exists and is equal to ∞ (i.e. if $\int Z_n^+ d\rho = \infty$ and $\int Z_n^- d\rho < \infty$) for all $n \in N$, then $\rho(\{w: S_n(w)/n \rightarrow \infty\}) = 1$. Similarly if $\int Z_n d\rho$ exists and is equal to $-\infty$ for all $n \in N$, then $\rho(\{w: S_n(w)/n \rightarrow -\infty\}) = 1$. (For a proof of the above theorem see [2] and [3].)

Coming back to Markov chains, let f be a real valued function on the state space I . Let i be a recurrent state. For each $n \in N$, let $l(n)$ be the random variable defined by $l(n)(h) = \sum_{k=1}^n \partial_i(h_k)$ where $\partial_i(j) = 1$ if $i = j$ and 0 if $i \neq j$.

Define $\mu_i(f) = \int [\sum_{k=1}^\infty f(h_k) 1_{\{l(1) \geq k\}}] d\sigma[i]$, provided the integral exists. The integrand is the sum of the f values of terms in the first i -block.

REMARK. Observe that if f^+ and f^- denote the positive and negative parts of f , then $\mu_i(f) = \mu_i(f^+) - \mu_i(f^-)$ provided either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$. (However $\mu_i(f)$ could be defined even if $\mu_i(f^+) = \mu_i(f^-) = \infty$.)

In what follows i is assumed to be recurrent.

LEMMA 1. If either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$,

$$\sigma[i] \left\{ \left\{ h: \frac{\sum_{k=1}^n f(h_k)}{l(n)(h)} \rightarrow \mu_i(f) \right\} \right\} = 1.$$

PROOF. We shall prove the result for a general nonnegative f . In that case we would have the required result for f^+ and f^- and that would suffice in view of the remark made above. So assume $f \geq 0$. Define a function Z on F , the set of all finite sequences with last coordinate i and none of the other coordinates i , by $Z(i_1, \dots, i_n) = \sum_{k=1}^n f(i_k)$. Now define a sequence $\{Z_n\}$ of identical coordinate mappings on (Ω, \mathcal{F}) by $Z_n(\omega) = Z(\omega_n)$, $\omega \in \Omega$, $n \in N$, ω_n being the n th coordinate of ω ($\Omega, \mathcal{F}, \Phi$ and π are as defined in §5). Since π is an i.i.d. strategy, by SLLN for i.i.d. strategies, $\pi(B) = 1$ where

$$B = \left\{ \omega: \frac{\sum_{k=1}^n Z_k(\omega)}{n} \rightarrow \mu_i(f) \right\}.$$

Hence by the blocks theorem $\sigma[i](\Phi^{-1}(B)) = 1$. Since $f \geq 0$,

$$\frac{\sum_{k=1}^{l(n)(h)} Z_k(\Phi(h))}{l(n)(h)} \leq \frac{\sum_{k=1}^n f(h_k)}{l(n)(h)} \leq \frac{\sum_{k=1}^{l(n)(h)+1} Z_k(\Phi(h))}{l(n)(h)} \quad \text{for all } h \in G_i.$$

Further $l(n)(h) \rightarrow \infty$ as $n \rightarrow \infty$ on G_i . Therefore

$$\Phi^{-1}(B) \subseteq \left\{ \left\{ h: \frac{\sum_{k=1}^n f(h_k)}{l(n)(h)} \rightarrow \mu_i(f) \right\} \right\}$$

and the proof of the lemma is complete.

COROLLARY 2. If f and g are functions on I such that (i) either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$, (ii) either $\mu_i(g^+) < \infty$ or $\mu_i(g^-) < \infty$ and (iii) either $0 < |\mu_i(f)| < \infty$ or $0 < |\mu_i(g)| < \infty$, then

$$\sigma[i] \left\{ \left\{ h: \frac{\sum_{k=1}^n f(h_k)}{\sum_{k=1}^n g(h_k)} \rightarrow \frac{\mu_i(f)}{\mu_i(g)} \right\} \right\} = 1.$$

PROOF. In view of assumptions (i) and (ii), Lemma 1 can be applied separately to f as well as g and because of (iii) division can be carried out on the common set of convergence.

COROLLARY 3. Let f be such that either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$. Assume that either i is positive recurrent or $0 < |\mu_i(f)| < \infty$ (i is already assumed to be recurrent). Then $\sigma[i](\{S_n(h)/n \rightarrow \mu_i(f)/m_{ii}\}) = 1$, where $S_n(h) = f(h_1) + \dots + f(h_n)$.

PROOF. Let g be the function defined by $g(j) = 1$ for all $j \in I$. f satisfies condition (i) of Corollary 2 by assumption and g satisfies (ii) because g is nonnegative. (iii) is satisfied because $0 < |\mu_i(g)| = m_{ii} < \infty$ if i is positive recurrent; otherwise we anyway assume $0 < |\mu_i(f)| < \infty$. Therefore by Corollary 2, the result follows.

PROPOSITION 4. *Let f be such that either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$ and let $i \xrightarrow{w} j$. Assume that either i is positive recurrent or $0 < |\mu_i(f)| < \infty$. Then*

$$\sigma[j] \left(\left\{ h: \frac{S_n}{n} \rightarrow \frac{\mu_i(f)}{m_{ii}} \right\} \right) = 1.$$

PROOF. Observe that if $A = \{h: S_n/n \rightarrow \mu_i(f)/m_{ii}\}$, then $A \cap \{t < \infty\}$, where t is the time of first occurrence of a j , is conditionally determined given t and further $(A \cap \{t < \infty\})p_t(h) = A$ for all h such that $t(h) < \infty$. Therefore by the strong Markov property, $\sigma[i](A) = \sigma[i](A \cap \{t < \infty\}) = \sigma[j](A) \cdot \sigma[i](\{t < \infty\}) = \sigma[j](A)$. The result now follows from Corollary 3.

COROLLARY 5. *Let i be positive recurrent and $i \xrightarrow{w} j$. Let f be such that (i) either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$, and (ii) either $\mu_j(f^+) < \infty$ or $\mu_j(f^-) < \infty$. Then*

$$\mu_i(f)/m_{ii} = \mu_j(f)/m_{jj}.$$

PROOF. Since i is positive recurrent and $i \xrightarrow{w} j$, j is also positive recurrent. Therefore by Corollary 3 applied to j , S_n/n converges to $\mu_j(f)/m_{jj}$ on a set of $\sigma[j]$ -measure one. Also by Proposition 4, S_n/n converges to $\mu_i(f)/m_{ii}$ on a set of $\sigma[j]$ -measure one. Therefore the result follows.

COROLLARY 6. *If i is positive recurrent and $i \xrightarrow{w} j$, for each subset E of the state space,*

$$\mu_i(1_E)/m_{ii} = \mu_j(1_E)/m_{jj}.$$

PROOF. Immediate from Corollary 5 by taking $f = 1_E$.

THEOREM 7 (SLLN). *Let I be a recurrent class under a Markov strategy σ . Let f be a function on I such that either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$ for some $i \in I$. Assume that either I is a positive recurrent class or $0 < |\mu_i(f)| < \infty$. Then*

$$\sigma \left(\left\{ h: \frac{S_n}{n} \rightarrow \frac{\mu_i(f)}{m_{ii}} \right\} \right) = 1.$$

PROOF. By Proposition 5, $\sigma[j](\{h: S_n/n \rightarrow \mu_i(f)/m_{ii}\}) = 1$ for all $j \in I$. Further if σ_0 is the initial distribution of σ ,

$$\sigma \left(\left\{ h: \frac{S_n}{n} \rightarrow \frac{\mu_i(f)}{m_{ii}} \right\} \right) = \int \sigma[j] \left(\left\{ h: \frac{S_n}{n} \rightarrow \frac{\mu_i(f)}{m_{ii}} \right\} \right) d\sigma_0(j)$$

by Proposition 1 of §2. Thus the result follows.

Many important limit theorems for Markov chains follow as special cases of the above theorem. We give below two such as examples.

COROLLARY 8. If I is a recurrent class,

$$\sigma\left(\left\{h: \frac{l(n)}{n} \rightarrow \frac{1}{m_{ii}}\right\}\right) = 1.$$

PROOF. In Theorem 7, take f to be the function defined by

$$\begin{aligned} f(j) &= 1 \quad \text{if } j = i, \\ &= 0 \quad \text{if } j \neq i. \end{aligned}$$

COROLLARY 9. If I is a positive recurrent class and $i \in I$, then

$$\sigma(\{h: t(l(n))/n \rightarrow 1\}) = 1,$$

where, as before, for each $k \in N$, $t(k)$ is the time of k th occurrence of i .

PROOF. In view of Corollary 8, it is enough to show that

$$\sigma(\{h: t(l(n))/l(n) \rightarrow m_{ii}\}) = 1.$$

This would follow if we show that $\sigma[j](\{h: t(l(n))/l(n) \rightarrow m_{ii}\}) = 1$ for each $j \in I$.

By applying the blocks theorem and SLLN for i.i.d. strategies to the block length functions $\{\lambda_n\}$ defined as in §6, we can show that $\sigma[i](\{t(l(n))/l(n) \rightarrow m_{ii}\}) = 1$. Now an argument similar to the one used in the proof of Proposition 4 shows that

$$\sigma[j]\left(\left\{\frac{t(l(n))}{l(n)} \rightarrow m_{ii}\right\} \cap G_i\right) = \sigma[i]\left(\left\{\frac{t(l(n))}{l(n)} \rightarrow m_{ii}\right\} \cap G_i\right) \quad \text{for all } j \in I.$$

Therefore

$$\sigma[j]\left(\left\{\frac{t(l(n))}{l(n)} \rightarrow m_{ii}\right\}\right) = \sigma[i]\left(\left\{\frac{t(l(n))}{l(n)} \rightarrow m_{ii}\right\}\right) = 1 \quad \text{for all } j \in I.$$

This completes the proof of the corollary.

8. Existence of stationary initial distributions. Let $\{\sigma(i)\}_{i \in I}$ be a Markov strategy.

DEFINITION. A stationary initial distribution λ for $\{\sigma(i)\}_{i \in I}$ is a finitely additive probability measure defined on all subsets of I such that

$$\int \sigma(j)(E) d\lambda(j) = \lambda(E) \quad \text{for all } E \subseteq I. \quad (1)$$

Suppose l is a Banach limit (see page 73 of [9]).

LEMMA 1. Let $\{\lambda_n\}$ be a sequence of finitely additive nonnegative measures defined on all subsets of I such that there is a positive real number K with $\lambda_n(I) \leq K$ for all $n \in N$. Define the finitely additive measure λ by $\lambda(E) = l(\{\lambda_n(E)\})$ for all $E \subseteq I$. Then for every bounded real valued function f on I , $l(\{f d\lambda_n\}) = \int f d\lambda$.

(In particular if $\{\lambda_n\}$ and λ are finitely additive measures such that (i) $\lambda_n(I) \leq K$ for all $n \in N$ and some positive real number K and (ii) $\text{Lim}_n \lambda_n(E) = \lambda(E)$ for all $E \subseteq I$, then for every bounded real valued function f on I

$$\text{Lim}_n \int f d\lambda_n = \int f d\lambda.)$$

PROOF. If $f = 1_E$, $E \subseteq I$, the result is true by definition. Hence by linearity of the Banach limit, the result is true for simple functions. The general formula follows after approximation by simple functions.

For each $n \in N$ and $i \in I$ we shall now define inductively a measure $\sigma^n(i)(\cdot)$ on all subsets of I . Set $\sigma^1(i) = \sigma(i)$, $i \in I$ and $\sigma^{n+1}(i)(E) = \int \sigma(j)(E) d\sigma^n(i)(j)$, $E \subseteq I$, $n \in N$.

It is easy to check inductively that $\sigma^n(i)$ is a finitely additive probability for all $n \in N$ and all $i \in I$.

THEOREM 2. *Every Markov strategy $\{\sigma(i)\}_{i \in I}$ has a stationary initial distribution.*

PROOF. Fix an $i \in I$ and a Banach limit l . Let $\lambda_n = \sigma^n(i)$, $n \in N$. Then we claim that λ defined as in Lemma 1 is a stationary initial distribution. This is because Lemma 1 applied to each of the functions f_E defined by $f_E(j) = \sigma(j)(E)$ for all $j \in I$, $E \subseteq I$, gives us (1).

As is evident from the proof of Theorem 2, the stationary initial distribution for a Markov strategy $\{\sigma(i)\}_{i \in I}$ is not in general unique. We give an example to show that even in the case $\sigma(i)$ is countably additive for each $i \in I$, there can be many stationary initial distributions.

EXAMPLE. $I = N$. $\sigma(1) = \sum_{n=2}^{\infty} p_n \partial_n$ where p_n is a sequence of nonnegative real numbers such that (i) $\sum_{n=2}^{\infty} p_n = 1$, (ii) $\sum_{n=2}^{\infty} n p_n < \infty$, ∂_n being the Dirac measure at n , and $\sigma(n+1) = \partial_n$, $n \in N$.

In the above example it is easy to see that I is a positive recurrent class. It is known in such a case that there is a unique countably additive stationary initial distribution. (See [5] or [10].) However every finitely additive probability λ on all subsets of N such that $\lambda(E) = \lambda(E+1)$ for all $E \subseteq N$ where $E+1 = \{n+1: n \in E\}$, is a stationary initial distribution. Every Banach limit induces one such measure and such measures are necessarily purely finitely additive. Therefore in the above example, the convex set of stationary initial distributions is quite large.

9. The countably additive part. Even though we cannot assert uniqueness of the stationary initial distribution, it is possible in many cases to prove uniqueness of the countably additive part up to multiplication by a constant.

LEMMA 1. *If $i \in I$ is transient then $\int Y(i)(h) d\sigma(h) < \infty$ where $Y(i)(h) = \sum_{n=1}^{\infty} \partial_i(h_n)$, ∂_i being the function on I defined by*

$$\begin{aligned} \partial_i(j) &= 1 \quad \text{if } i = j, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

PROOF. Since $Y(i)$ is nonnegative and extended integer valued,

$$\int Y(i)(h) d\sigma(h) = \sum_{n=1}^{\infty} \sigma(\{Y(i) \geq n\}).$$

If $t(1)$ is the time of first occurrence of i , then by the strong Markov property applied to each term in the summation,

$$\begin{aligned} \int Y(i)(h) d\sigma(h) &= \sum_{n=1}^{\infty} \sigma(\{t(1) < \infty\}) \sigma[i](\{Y(i) \geq n-1\}) \\ &= \sum_{n=1}^{\infty} \sigma(\{t(1) < \infty\}) (f_{ii}^*)^{n-1} \end{aligned}$$

since $Y(i)$ is geometric under $\sigma[i]$ (see remark after Proposition 5 of §4),

$$= \frac{\sigma(\{t(1) < \infty\})}{1 - \hat{f}_{ii}^*} < \infty, \quad \text{because } i \text{ is transient.}$$

THEOREM 2. *Let λ be a stationary initial distribution for a Markov strategy $\{\sigma(i)\}_{i \in I}$. Suppose $i \in I$ is transient. Then $\lambda(\{i\}) = 0$.*

PROOF. Consider the Markov strategy $\sigma = \{\lambda, \{\sigma(i)\}_{i \in I}\}$. It can easily be checked by induction that

$$\int 1_E(h_n) d\sigma(h) = \lambda(E) \quad \text{for all } E \subseteq I \quad (2)$$

for each $n \in N$.

In particular from (2) it follows that $\int \partial_i(h_n) d\sigma(h) = \lambda(\{i\})$ for every $n \in N$. Consequently $\sum_{n=1}^{\infty} \int \partial_i(h_n) d\sigma(h) < \infty$ if and only if $\lambda(\{i\}) = 0$. Since i is transient, and since

$$\int Y(i)(h) d\sigma(h) = \int \left[\sum_{n=1}^{\infty} \partial_i(h_n) \right] d\sigma(h) > \sum_{n=1}^{\infty} \int \partial_i(h_n) d\sigma(h),$$

an application of Lemma 1 completes the proof of the theorem.

REMARK. Equation (2) holds for any stationary initial distribution λ for any Markov strategy $\sigma = \{\lambda, \{\sigma(i)\}_{i \in I}\}$.

COROLLARY 3. *Let I be a countable transient class under $\{\sigma(i)\}_{i \in I}$. Then every stationary initial distribution is purely finitely additive.*

PROOF. Immediate from Theorem 2.

We shall next study the stationary initial distribution on recurrent states. For this we shall need to prove a ratio limit theorem. We first prove a few lemmas needed for the purpose.

For each of the Lemmas 4–8, we assume that I is a recurrent class and i, j are states in I such that $i \neq j$. Let t be the incomplete stop rule corresponding to the first occurrence of j . For every $k \in N$ let $t(k)$ be the incomplete stop rule corresponding to the k th occurrence of i . Let $Z_1(h) = \sum_{m=1}^{t(1)(h)} \partial_j(h_m)$, $Z_{k+1}(h) = \sum_{m=t(k)(h)+1}^{t(k+1)(h)} \partial_j(h_m)$, $k \in N$ for all $h \in G_i$. If σ is a Markov strategy (under which I is a recurrent class),

LEMMA 4. $\int Z_1 d\sigma[j] < \infty$.

PROOF.

$$\int Z_1 d\sigma[j] = \sum_{m=1}^{\infty} \sigma[j](\{Z_1 > m\}) = \sum_{m=1}^{\infty} \theta^m,$$

where $\theta = \sigma[j](\{t < t(1)\})$. The last equality follows by repeated application of the strong Markov property to the incomplete stop rule t' defined by

$$\begin{aligned} t' &= t \quad \text{if } t < t(1), \\ &= \infty \quad \text{otherwise.} \end{aligned}$$

Observe that $1 - \theta = 1 - \sigma[j](\{t < t(1)\}) = \sigma[j](\{t(1) < t\})$ which is positive by Lemma 3 of §6. Therefore $\theta < 1$ and the lemma follows.

LEMMA 5. $\int Z_1 d\sigma \leq \int Z_1 d\sigma[j] + 1$. Consequently $\int Z_1 d\sigma < \infty$.

PROOF. $\int Z_1 d\sigma = \sum_{m=1}^{\infty} \sigma(\{Z_1 \geq m\}) = \sum_{m=1}^{\infty} \sigma(\{t' < \infty\})\sigma[j](\{Z_1 \geq m-1\})$ by the strong Markov property where t' is the incomplete stop rule defined in proof of Lemma 4. Therefore

$$\begin{aligned} \int Z_1 d\sigma &= \sigma(\{t' < \infty\}) \left\{ 1 + \sum_{m=1}^{\infty} \sigma[j](\{Z_1 \geq m\}) \right\} \\ &\leq \int Z_1 d\sigma[j] + 1. \end{aligned}$$

The previous lemma then implies that $\int Z_1 d\sigma < \infty$.

LEMMA 6. For $k \geq 2$, $\int Z_k d\sigma = \int Z_1 d\sigma[i]$.

PROOF.

$$\int Z_k d\sigma = \sum_{m=1}^{\infty} \sigma(\{Z_k \geq m\}) = \sum_{m=1}^{\infty} \sigma(\{t(k-1) < \infty\}) \cdot \sigma[i](\{Z_1 \geq m\})$$

by the strong Markov property. Since I is a recurrent class, it follows that $\sigma(\{t(k-1) < \infty\}) = 1$. Hence $\int Z_k d\sigma = \sum_{m=1}^{\infty} \sigma[i](\{Z_1 \geq m\}) = \int Z_1 d\sigma[i]$. This completes the proof of the lemma.

For each $n \in N$, let $l(n)(h) = \sum_{k=1}^n \partial_i(h_k)$, $h \in G_i$. Clearly $l(n)(h) \leq n$ for all $h \in G_i$. For each $n \in N$, let $Z'_n(h) = \sum_{m=n+1}^{l(n)+1} \partial_j(h_m)$, $h \in G_i$.

LEMMA 7. For each $n \in N$, $\int Z'_n(h) d\sigma(h) \leq \int Z_1 d\sigma[j] + 1$.

PROOF.

$$\int Z'_n(h) d\sigma(h) = \sum_{m=1}^{\infty} \sigma(\{Z'_n \geq m\}) = \sum_{m=1}^{\infty} \sigma(\{t''_n < \infty\})\sigma[j](\{Z_1 \geq m-1\})$$

by the strong Markov property where t''_n is the incomplete stop rule defined by

$$\begin{aligned} t''_n(h) &= t(h_{n+1}, h_{n+2}, \dots) + n \quad \text{if } t(h_{n+1}, h_{n+2}, \dots) + n < t(l(n) + 1), \\ &= \infty \quad \text{otherwise.} \end{aligned}$$

Therefore $\int Z'_n(h) d\sigma(h) = \sigma(\{t''_n < \infty\}) \sum_{m=1}^{\infty} \sigma[j](\{Z_1 \geq m-1\}) \leq \int Z_1 d\sigma[j] + 1$. Hence the lemma is proved.

LEMMA 8. If $k \in N$ is such that $2 \leq k \leq n+1$, then

$$\int Z_k 1_{\{l(n) \geq k-1\}} d\sigma = \sigma(\{l(n) \geq k-1\}) \cdot \int Z_1 d\sigma[i].$$

PROOF.

$$\begin{aligned} \int Z_k 1_{\{l(n) \geq k-1\}} d\sigma &= \sum_{m=1}^{\infty} \sigma(\{Z_k 1_{\{l(n) \geq k-1\}} \geq m\}) \\ &= \sum_{m=1}^{\infty} \sigma(\{Z_k \geq m \text{ and } l(n) \geq k-1\}) \\ &= \sum_{m=1}^{\infty} \sigma(\{l(n) \geq k-1\})\sigma[i](\{Z_1 \geq m\}) \end{aligned}$$

by the strong Markov property applied to the incomplete stop rule t''' defined by

$$\begin{aligned} t'''(h) &= t(k-1)(h) \quad \text{if } l(n)(h) \geq k-1, \\ &= \infty \quad \text{otherwise.} \\ &= \sigma(\{l(n) \geq k-1\}) \sum_{m=1}^{\infty} \sigma[i](\{Z_1 \geq m\}) \\ &= \sigma(\{l(n) \geq k-1\}) \int Z_1 d\sigma[i]. \end{aligned}$$

THEOREM 9. Let σ be a Markov strategy such that I is a recurrent class. Let $i, j \in I$. Assume that $\int l(n) d\sigma = \int [\sum_{k=1}^n \partial_j(h_k)] d\sigma(h)$ tends to ∞ as $n \rightarrow \infty$. Then

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \int \partial_j(h_k) d\sigma(h)}{\sum_{k=1}^n \int \partial_i(h_k) d\sigma(h)} = \int Z_1 d\sigma[i].$$

(Note that $\int Z_1 d\sigma[i] = \mu_i(\partial_j)$ according to the notation introduced in §7).

PROOF. If $i = j$, both sides in the above assertion are trivially equal to 1. Therefore assume that $i \neq j$. For each $h \in G_j$, and $n \in N$,

$$\sum_{k=1}^n \partial_j(h_k) = Z_1 + \sum_{k=2}^{l(n)+1} Z_k - Z'_n,$$

where the middle term is zero if $l(n) = 0$. Also $\sigma(G_j) = 1$. Therefore

$$\begin{aligned} \sum_{k=1}^n \int \partial_j(h_k) d\sigma &= \int \left[\sum_{k=1}^n \partial_j(h_k) \right] d\sigma \\ &= \int Z_1 d\sigma + \int \left[\sum_{k=2}^{l(n)+1} Z_k \right] d\sigma - \int Z'_n d\sigma. \end{aligned} \quad (3)$$

Now the middle term in the right of (3),

$$\begin{aligned} \int \left[\sum_{k=2}^{l(n)+1} Z_k \right] d\sigma &= \int \left[\sum_{k=2}^{n+1} Z_k 1_{\{l(n) \geq k-1\}} \right] d\sigma, \quad \text{since } l(n)(h) \leq n \text{ for all } h, \\ &= \sum_{k=2}^{n+1} \int Z_k 1_{\{l(n) \geq k-1\}} d\sigma \\ &= \sum_{k=2}^{n+1} \sigma(\{l(n) \geq k-1\}) \cdot \int Z_1 d\sigma[i], \quad \text{by Lemmas 6 and 8,} \\ &= \left(\int l(n) d\sigma \right) \left(\int Z_1 d\sigma[i] \right). \end{aligned}$$

Therefore from (3) we have, for each $n \in N$,

$$\frac{\sum_{k=1}^n \int \partial_j(h_k) d\sigma(h)}{\sum_{k=1}^n \int \partial_i(h_k) d\sigma(h)} = \frac{\int Z_1 d\sigma}{\int l(n) d\sigma} + \int Z_1 d\sigma[i] - \frac{\int Z'_n d\sigma}{\int l(n) d\sigma}.$$

By Lemmas 4, 5 and 7 and the assumption that $\int l(n) d\sigma \rightarrow \infty$ as $n \rightarrow \infty$, the theorem follows.

THEOREM 10. Let I be a recurrent class under $\{\sigma(i)\}_{i \in I}$. Let λ be a stationary initial distribution. If $\lambda(\{i\}) > 0$ for some i , then $\lambda(\{j\}) = \mu_i(\partial_j) \lambda(\{i\})$ for every $j \in I$. Consequently $\lambda(\{j\}) > 0$ for all $j \in I$.

PROOF. For the Markov strategy $\{\lambda, \{\sigma(i)\}_{i \in I}\}$, by the remark after proof of Theorem 2, $\int \partial_i(h_k) d\sigma(h) = \lambda(\{j\})$ for each $k \in N$ and $\int \partial_i(h_k) d\sigma(h) = \lambda(\{i\})$ for each $k \in N$. Since $\lambda(\{i\}) > 0$, $\lim_n \int l(n) d\sigma = \lim_n n\lambda(\{i\}) = \infty$. Hence by the previous theorem,

$$\lim_n \frac{n\lambda(\{j\})}{n\lambda(\{i\})} = \mu_i(\partial_j).$$

Hence $\lambda(\{j\}) = \mu_i(\partial_j) \cdot \lambda(\{i\})$.

We now only need to remark that $\mu_i(\partial_j) > 0$. This is because

$$\mu_i(\partial_j) = \int Z_i d\sigma[i] \geq \int_{\{t < t(1)\}} Z_1 d\sigma[i] \geq \sigma[i](\{t < t(1)\}) > 0$$

by Lemma 3 of §6. The proof of the theorem is complete.

COROLLARY 11. Let I be a recurrent class. Let λ_1 and λ_2 be two stationary initial distributions. If $\lambda_1(\{i\}) > 0$ for some $i \in I$, then there exists a real number c such that $\lambda_2(\{j\}) = c \cdot \lambda_1(\{j\})$ for all $j \in I$.

PROOF. If $\lambda_2(\{j\}) = 0$ for all $j \in I$, we can choose c to be zero. Otherwise by assumption and Theorem 10, $\lambda_1(\{j\}) > 0$ for all $j \in I$ and $\lambda_2(\{j\}) > 0$ for all $j \in I$.

Further $\lambda_1(\{j\}) = \mu_i(\partial_j) \cdot \lambda_1(\{i\})$ and $\lambda_2(\{j\}) = \mu_i(\partial_j) \cdot \lambda_2(\{i\})$, $i, j \in I$. Therefore

$$\frac{\lambda_2(\{j\})}{\lambda_1(\{j\})} = \frac{\lambda_2(\{i\})}{\lambda_1(\{i\})}, \quad i, j \in I.$$

Choose c to be this common value and we are done.

COROLLARY 12. Let I be a countable recurrent class such that $\sum_{j \in I} \mu_i(\partial_j) = \infty$ for some $i \in I$. Then every stationary initial distribution is purely finitely additive.

PROOF. Suppose on the contrary that there is a stationary initial distribution λ such that $\lambda(\{i\}) > 0$. Then by Theorem 10, $\lambda(\{j\}) = \mu_i(\partial_j)\lambda(\{i\})$.

Consequently $\sum_{j \in I} \lambda(\{j\}) = [\sum_{j \in I} \mu_i(\partial_j)]\lambda(\{i\}) = \infty$.

However, since λ is a finitely additive probability, $\sum_{j \in I} \lambda(\{j\}) \leq 1$, a contradiction.

Therefore $\lambda(\{i\}) = 0$ and hence $\lambda(\{j\}) = 0$ for all j . As I is countable, it follows that λ is purely finitely additive.

REMARK. Suppose I is countable. The condition that $\sum_{j \in I} \mu_i(\partial_j) = \infty$ always implies that i is null recurrent because $m_{ii} \geq \sum_{j \in I} \mu_i(\partial_j)$. In fact the condition is equivalent to null recurrence if $\sigma(i)$ is countably additive for each i , because then, $m_{ii} = \sum_{j \in I} \mu_i(\partial_j)$. We do not know if there are null recurrent chains with a stationary initial distribution which has a nontrivial countably additive part.

We now prove a theorem on the existence of stationary initial distributions which are countably additive.

THEOREM 13. Let I be a countable, weakly communicating class. If there exists a countably additive stationary initial distribution, then, for each $i \in I$, $\sigma(i)$ is a countably additive probability on I .

PROOF. Since I is countable, if λ is a countably additive stationary initial distribution, $\sum_{i \in I} \lambda(\{i\}) = 1$. This implies that $\lambda(\{i\}) > 0$ for some $i \in I$. This by Theorem 10 implies that $\lambda(\{j\}) > 0$ for all $j \in I$ if I is recurrent, but I has to be recurrent, since, by Theorem 2, I cannot be transient. Since λ is a stationary initial distribution, $\lambda(\{j\}) = \int \sigma(i)(\{j\}) d\lambda(i) = \sum_{i \in I} \sigma(i)(\{j\})\lambda(\{i\})$ for all $j \in I$, as λ is countably additive. Summing over all $j \in I$, we get

$$\begin{aligned} \sum_{j \in I} \lambda(\{j\}) &= \sum_{j \in I} \sum_{i \in I} \sigma(i)(\{j\})\lambda(\{i\}) \\ &= \sum_{i \in I} \left[\sum_{j \in I} \sigma(i)(\{j\}) \right] \lambda(\{i\}). \end{aligned}$$

Since $\sum_{j \in I} \lambda(\{j\}) = 1$ and $\sum_{j \in I} \sigma(i)(\{j\}) \leq 1$ for each $i \in I$ and $\lambda(\{j\}) > 0$ for all $j \in I$, it follows that $\sum_{j \in I} \sigma(i)(\{j\}) = 1$ for all $i \in I$. This means that $\sigma(i)$ is countably additive for $i \in I$ and the proof of the theorem is thus complete.

The problem of existence of countably additive stationary initial distributions for a Markov strategy $\{\sigma(i)\}_{i \in I}$, where $\sigma(i)$ is countably additive for each $i \in I$, has been extensively studied. Some of the references for this are [5], [6], [10], [11] and [12].

10. A canonical stationary initial distribution. In §8, we proved the existence of stationary initial distributions for Markov chains using Banach limits. However, if the chain has a positive recurrent state, it is possible to actually exhibit a stationary initial distribution with a nontrivial countably additive part. This stationary initial distribution has some desirable properties as we shall see later in this section.

Let $\{\sigma(i)\}_{i \in I}$ be a Markov strategy. Let i be a positive recurrent state in I . Define a function λ_i on all subsets of I by

$$\lambda_i(E) = \frac{1}{m_{ii}} \cdot \mu_i(1_E), \quad E \subseteq I.$$

Recall that

$$\mu_i(1_E) = \int \left(\sum_{n=1}^{\infty} 1_E(h_n) 1_{\{t(1) > n\}}(h) \right) d\sigma[i](h).$$

THEOREM 1. λ_i is a stationary initial distribution for $\{\sigma(i)\}_{i \in I}$.

PROOF. It is easy to see from definition that λ_i is a nonnegative finitely additive measure on all subsets of I . Also $\lambda_i(I) = m_{ii}^{-1} \cdot \int t(1) d\sigma[i] = 1$.

Therefore λ_i is a probability. Let

$$\alpha_{n,i}(E) = \frac{1}{m_{ii}} \int 1_E(h_n) 1_{\{t(1) > n\}} d\sigma[i](h), \quad n \in N, E \subseteq I.$$

It is easy to see that for each $n \in N$, $\alpha_{n,i}$ is a nonnegative finitely additive measure on all subsets of I (not necessarily a probability).

For each $m \in N$ let $\lambda_{m,i}$ be the finitely additive measure defined by $\lambda_{m,i} = \sum_{n=1}^m \alpha_{n,i}$. We now claim that

$$\lambda_i = \sum_{n=1}^{\infty} \alpha_{n,i} = \lim_{m \rightarrow \infty} \lambda_{m,i}. \quad (4)$$

Clearly for each $E \subseteq I$ and $m \in N$

$$\begin{aligned}\lambda_i(E) &= \frac{1}{m_{ii}} \int \left(\sum_{n=1}^{\infty} 1_E(h_n) 1_{\{t(1) > n\}}(h) \right) d\sigma[i](h) \\ &\geq \frac{1}{m_{ii}} \int \left(\sum_{n=1}^m 1_E(h_n) 1_{\{t(1) > n\}}(h) \right) d\sigma[i](h) \\ &= \frac{1}{m_{ii}} \sum_{n=1}^m \int 1_E(h_n) 1_{\{t(1) > n\}}(h) d\sigma[i](h) \\ &= \lambda_{m,i}(E).\end{aligned}$$

Therefore

$$\lambda_i(E) \geq \left(\sum_{n=1}^{\infty} \alpha_{n,i} \right)(E) \quad \text{for all } E \subseteq I. \quad (5)$$

Further

$$\begin{aligned}\left(\sum_{n=1}^{\infty} \alpha_{n,i} \right)(I) &= \frac{1}{m_{ii}} \sum_{n=1}^{\infty} \sigma[i](\{t(1) > n\}) \\ &= \frac{1}{m_{ii}} \cdot m_{ii} = 1.\end{aligned}$$

Therefore $\sum_{n=1}^{\infty} \alpha_{n,i}$ is a probability. (4) now follows from (5) and the fact that both λ_i and $\sum_{n=1}^{\infty} \alpha_{n,i}$ are probabilities.

From (4) and Lemma 1 of §8, it follows that

$$\int f d\lambda_i = \lim_{m \rightarrow \infty} \int f d\lambda_{m,i} \quad \text{for every bounded function } f \text{ on } I. \quad (6)$$

Now for each $m \in N$ and $E \subseteq I$

$$\begin{aligned}m_{ii}\lambda_{m+1,i}(E) &= \sum_{n=1}^{m+1} \int 1_E(h_n) 1_{\{t(1) > n\}}(h) d\sigma[i](h) \\ &= \sigma(i)(E) + \sum_{n=1}^m \int 1_E(h_{n+1}) 1_{\{t(1) > n+1\}}(h) d\sigma[i](h) \\ &= \sigma(i)(E) + m_{ii} \sum_{n=1}^m \int_{\{j \neq i\}} \sigma(j)(E) d\alpha_{n,i}(j)\end{aligned}$$

by conditioning at the n th coordinate and then using the change of variable theorem,

$$= \sigma(i)(E) + m_{ii} \int_{\{j \neq i\}} \sigma(j)(E) d\lambda_{m,i}(j).$$

Therefore

$$\lambda_{m+1,i}(E) = \frac{\sigma(i)(E)(1 - m_{ii}\lambda_{m,i}(\{i\}))}{m_{ii}} + \int \sigma(j)(E) d\lambda_{m,i}(j) \quad \text{for all } m \in N \text{ and } E \subseteq I. \quad (7)$$

From (6) and the fact that $\lim_{m \rightarrow \infty} m_{ii}\lambda_{m,i}(\{i\}) = \mu_i(\partial_i) = 1$, it follows by taking limits as $m \rightarrow \infty$ on both sides of (7) that $\lambda_i(E) = \int \sigma(j)(E) d\lambda_i(j)$. Therefore λ_i is a stationary initial distribution.

EXAMPLE. Let $I = N$. Define $\sigma(1) = \mu$ where μ is a purely finitely additive probability on N and $\sigma(n+1) = \partial_1$, $n \in N$, ∂_1 being the Dirac measure at 1. Clearly 1 is a positive recurrent state and the method of Theorem 1 gives $\lambda = \frac{1}{2}\mu + \frac{1}{2}\partial_1$ as a stationary initial distribution.

THEOREM 2. Let $\{\sigma(i)\}_{i \in I}$ be a Markov strategy. Let i be positive recurrent and $i \xrightarrow{w} j$. Then $\lambda_i = \lambda_j$. (By Theorem 7 of §6, j is positive recurrent and λ_j is defined.)

PROOF. This follows immediately from Corollary 6 of §7.

COROLLARY 3. Let I be a positive recurrent class. Then $\lambda_i = \lambda_j$ for all $i, j \in I$.

PROOF. Follows trivially from Theorem 2.

We have therefore shown that if I is positive recurrent, there is a stationary initial distribution λ with

$$\lambda(\{i\}) = \frac{\mu_i(\partial_i)}{m_{ii}} = \frac{1}{m_{ii}} \quad \text{for all } i \in I.$$

Clearly $\lambda(\{i\}) > 0$ for all $i \in I$. Therefore λ has a nontrivial countably additive component. This λ will be called the canonical stationary initial distribution for the positive recurrent chain.

This canonical stationary initial distribution λ is important for the Markov chain and has desirable properties. As stated in the Strong Law of Large Numbers for a positive recurrent chain, for a function f such that for some $i \in I$, either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$,

$$\sigma\left(\left\{h: \frac{\sum_{k=1}^n f(h_k)}{n} \rightarrow \frac{\mu_i(f)}{m_{ii}}\right\}\right) = 1.$$

This almost sure limit is actually the integral of f with respect to the canonical stationary initial distribution.

THEOREM 4. Let I be a positive recurrent class and λ the canonical stationary initial distribution.

For any real valued function f on I such that for some $i \in I$ either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$,

$$\frac{\mu_i(f)}{m_{ii}} = \int f d\lambda.$$

PROOF. The proof is routine, first done for simple functions and then the general case by approximation with simple functions.

We conclude this section with an interesting characterization of countably additive positive recurrent chains.

THEOREM 5. Let I be a weakly communicating class under $\{\sigma(i)\}_{i \in I}$. Then I is positive recurrent and $\sigma(i)$ is countably additive for each $i \in I$ if and only if $\sum_{i \in I} m_{ii}^{-1} = 1$.

PROOF. Only if part. Since I is positive recurrent, there is a stationary initial distribution such that $\lambda(\{i\}) > 0$ for all $i \in I$. (The canonical stationary initial distribution has this property.) Therefore I has to be countable. Further given that

$\sigma(i)$ is countably additive for each i , it is known that $\pi(\{1\}) = m_{ii}^{-1}$, $i \in I$, is a countably additive stationary initial distribution. (See [5] or [10] for a proof of this.)

Therefore $\sum_{i \in I} m_{ii}^{-1} = 1$.

If *part*. $\sum_{i \in I} m_{ii}^{-1} = 1$ implies that $m_{ii}^{-1} > 0$ for some i , and hence $m_{ii} < \infty$, consequently i is positive recurrent. As I is a communicating class, I is a positive recurrent class. If λ is the canonical stationary initial distribution, we know that $\lambda(\{i\}) = \mu_i(\partial_i)/m_{ii} = m_{ii}^{-1}$ for all i . Therefore $\sum_{i \in I} m_{ii}^{-1} = 1$ implies that λ is countably additive. Therefore by Theorem 13 of §9, the result follows.

11. Ergodicity of the shift. The shift transformation T on H is defined by $T(h_1, h_2, \dots) = (h_2, h_3, \dots)$ for all $h \in H$. The invariant σ -field \mathcal{G} is the collection of all sets $A \in \mathfrak{B}$ such that $T^{-1}(A) = A$.

T is said to be ergodic under σ if (i) $\sigma(T^{-1}(A)) = \sigma(A)$ for all $A \in \mathfrak{B}$ and (ii) $\sigma(A) = 0$ or 1 for $A \in \mathcal{G}$.

THEOREM 1. *Let I be a recurrent class under $\{\sigma(i)\}_{i \in I}$. Let λ be a stationary initial distribution. Then the shift T is ergodic under $\sigma = \{\lambda, \{\sigma(i)\}_{i \in I}\}$.*

PROOF. For $A \in \mathfrak{B}$

$$\begin{aligned} \sigma(T^{-1}(A)) &= \int \sigma[h_1, h_2] \{ (T^{-1}(A))h_1, h_2 \} d\sigma(h) \\ &= \int \sigma[h_2](Ah_2) d\sigma(h) \\ &= \int \sigma[j](Aj) d\lambda(j), \end{aligned}$$

by the change of variable theorem and the stationarity of λ ,

$$= \sigma(A).$$

Therefore condition (i) for ergodicity of σ is proved.

Let $B \in \mathcal{G}$. Let $i, j \in I$ and let t be the first time of occurrence of j . Since I is a recurrent class, $\sigma[i](\{t < \infty\}) = f_{ij}^* = 1$.

So $\sigma[i](B) = \sigma[i](B \cap \{t < \infty\}) = \sigma[i](t < \infty) \cdot \sigma[j](B) = \sigma[j](B)$, by the strong Markov property. Thus $\sigma[j](B) = \alpha$ for all $j \in I$ for some α , $0 \leq \alpha \leq 1$. Hence $\sigma(B) = \alpha$. We shall prove that α has to be 0 or 1.

Let $\varepsilon > 0$. Since α is regular, there exists a clopen set K such that $\sigma(B \Delta K) < \varepsilon/2$. Let s be a stop rule such that $Kp_s(h) = \emptyset$ for $h \in K^c$ and $Kp_s(h) = H$ for $h \in K$. (See [8] for the proof of the existence of such a stop rule.)

We then have

$$\begin{aligned} \sigma(B \cap K) &= \int \sigma[p_s(h)](B \cap K) p_s(h) d\sigma(h) \\ &= \int_K \sigma[p_s(h)](B) d\sigma(h) \\ &= \alpha \cdot \sigma(K). \end{aligned}$$

Thus

$$\begin{aligned}\alpha - \alpha^2 &= \sigma(B) - [\sigma(B)]^2 \\ &\leq |\sigma(B) - \sigma(B \cap K)| + |\alpha \cdot \sigma(K) - [\sigma(B)]^2| \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon.\end{aligned}$$

As ε is arbitrary, $\alpha - \alpha^2 = 0$ and hence $\alpha = 0$ or 1. The proof of the theorem is thus complete.

REMARK. It is clear from the proof above that the triviality of the invariant σ -field for a recurrent chain does not depend on the stationarity of the initial distribution. The proof given above has the same idea as the Purves-Sudderth [14] proof of the Kolmogorov 0-1 law for the independent case. In fact the triviality of the invariant σ -field for recurrent chains could be proved by going over to the blocks space using the blocks theorem and then applying the Kolmogorov 0-1 law.

ACKNOWLEDGEMENT. I am greatly indebted to Professor A. Maitra for many useful suggestions, to Professor B. V. Rao for several helpful discussions, and to the referee for his thoughtful comments.

REFERENCES

1. L. Breiman, *Probability*, Addison-Wesley, Reading, Mass., 1968.
2. R. Chen, *On almost sure convergence in a finitely additive setting*, Z. Wahrsch. Verw. Gebiete 37 (1977), 341–356.
3. ———, *Some finitely additive versions of the strong law of large numbers*, Israel J. Math. 24 (1976), 244–259.
4. K. L. Chung, *A course in probability theory*, 2nd ed., Academic Press, New York, 1974.
5. ———, *Markov chains with stationary transition probabilities*, 2nd ed., Springer-Verlag, Berlin, 1967.
6. C. Derman, *A solution to a set of fundamental equations in Markov chains*, Proc. Amer. Math. Soc. 5 (1954), 332–334.
7. L. E. Dubins, *On Lebesgue-like extensions of finitely additive measures*, Ann. Probab. 2 (1974), 456–463.
8. L. E. Dubins and L. J. Savage, *How to gamble if you must: Inequalities for stochastic processes*, McGraw-Hill, New York, 1965.
9. N. Dunford and J. T. Schwartz, *Linear operators*. I, Interscience, New York, 1958.
10. D. A. Freedman, *Markov chains*, Holden-Day, San Francisco, Calif., 1971.
11. T. E. Harris, *The existence of stationary measures for certain Markov processes*, Proc. 3rd Berkeley Sympos., vol. II, Univ. of California Press, Berkeley and Los Angeles, 1956, pp. 113–124.
12. ———, *Transient Markov chains with stationary measures*, Proc. Amer. Math. Soc. 8 (1957), 937–942.
13. R. A. Purves and W. D. Sudderth, *Some finitely additive probability*, Ann. Probab. 4 (1976), 259–276.
14. ———, *Some finitely additive probability*, Univ. of Minnesota School of Statistics Tech. Report No. 220, 1973.
15. S. Ramakrishnan, *Finitely additive Markov chains*, Ph.D. thesis, Indian Statistical Institute, 1980 (submitted).

STATISTICS AND MATHEMATICS DIVISION, INDIAN STATISTICAL INSTITUTE, CALCUTTA 700 035, INDIA