FINITELY ADDITIVE MARKOV CHAINS

S. RAMAKRISHNAN

Thesis submitted to the Indian Statistical Institute
in partial fulfilment of the requirements
for the award of the degree of
Doctor of Philosophy

CALCUTTA

1980

ACKNOWLEDGEMENT

This thesis owes a lot to :

the Indian Statistical Institute for providing adequate research facilities and a congenial atmosphere;

Professor A. Maitra, my supervisor, who got me interested in this topic, gave useful suggestions, brought to my attention the right references, and with his incredible energy and never-say-no attitude made this a memorable experience for me;

Professor B.V. Rao, who throughout listened to me carefully and critically and gave me constant encouragement and valuable advice;

Mr. Dilip Kumar Bardhan for his enthusiastic and efficient typing;

Mr. Dilip Chatterjee for his careful duplicating;

my friends, colleagues and teachers, who offered comforting consolation during my moments of depression;

my relatives and wellwishers, who continually asked when I was going to finish "that thing" in voices that could not be interminably denied ([3]).

S. Ramakrishnan

C O N T E N T S

*****

Introduction

The inclusion of the axiom of countable additivity in the definition of a probability has been questioned by several leading probabilists. Some of the subjectivists are prepared to accept the condition of countable additivity as a useful regularity condition but not as a natural one in every situation. According to them, there are situations where the natural chance function to be considered is only finitely additive. There have been some attempts to relax the condition of countable additivity and see how much can be achieved. Special mention should be made of De Finetti [19], who has introduced the basic concepts in probability theory and statistics in a finitely additive setting.

The first serious attempt to study discrete-parameter stochastic processes in a finitely additive setting was made by Dubins and Savage [14], who studied the theory of gambling in a finitely additive setting. The basic fortune space for them is an arbitrary nonempty set F, equipped with the discrete topology. A strategy associates with each finite sequence of elements of F, a finitely additive probability defined on all subsets of F. Imposing a natural condition, they obtain, corresponding to every strategy, a finitely additive probability measure on all clopen sets (in the product topology) of the countably infinite product of F. The technique used in defining the measure is an induction on the structure of clopen sets. Having obtained this measure on the clopen sets, they are able to formulate and prove weak versions of various limit theorems.

Later, Purves and Sudderth ([29] and [30]) showed that this measure (corresponding to a strategy) obtained on the clopen sets, can be extended uniquely subject to simple regularity conditions, to a field including the Borel sets in the product space. Having obtained this extension, they proved a strong law of large numbers for uniformly bounded random variables in case of an independent strategy. They also proved some 0-1 laws and a martingale convergence theorem.

Robert Chen ([4], [5], [6] and [7]) followed this up with a systematic study of almost sure convergence for independent strategies and martingales. He proved, among other results, a version of Kolmogorov's strong law of large numbers and a law of iterated logarithm for i.i.d strategies.

It seems worthwhile, therefore, to carry out a study of finitely additive probability theory and see how much of conventional probability theory can be carried through. It is in this spirit that we take up here the study of discrete-parameter finitely additive Markov chains with stationary transitions.

Almost throughout, we assume that the state space is an arbitrary nonempty set (not necessarily countable), equipped with the discrete topology. The transition probabilities, as well as the initial distribution, are assumed only to be finitely additive. The result of this study is quite encouraging. Many of the results in the conventional theory of Markov chains, with stationary transitions and a countable state space, go through in this general setting. The proofs are not necessarily much more difficult. Generally, familiar arguments are carried through with the help of stop rule techniques. We have drawn a lot of inspiration from the books of Chung [8] and Freedman [20], which have greatly influenced our presentation. We give below a summary of the main results in various sections of the thesis.

Section 1 sets up the basic framework and states some of the fundamental results in the finitely additive theory, which will be of use to us in the sequel.

In section 2, we formulate and prove the Strong Markov property. Most of our proofs in later sections require the use of the Strong Markov property rather than the usual Markov property.

In section 3, we classify the state space into communicating classes. It is necessary to consider two notions of communication, weak communication and strong communication (the two notions are equivalent in the countably additive case). While studying the period of a state, cyclically moving subclasses etc., it seems appropriate to consider strong communication. For other notions like recurrence, transience (section 5), it seems more natural to consider weak communication. The main result used in the classification is a version of the Chapman-Kolmogorov equations.

Section 4 studies closed sets. We prove that starting in a closed set, the probability of remaining forever in it is one. We also mention some differences with the countably additive theory, concerning the relationships between essential classes and closed sets.

The notions of recurrence and transience are studied in section 5. We prove that starting from a recurrent state, the probability of infinitely many returns to it, is one, whereas the probability of infinitely many returns, starting from a transient state is zero. We also show that, in a weakly communicating class, either all states are recurrent or none is.

The blocks theorem is proved in section 6. It asserts that for any recurrent state i, the sequence of i-blocks is i.i.d, given that the process starts at i. This theorem enables us to use in our theory known results for i.i.d strategies.

Section 7 proves some results on finitely additive integration. Although the results of this section are proved in a more general setting than others, the main objective here is to prove a version of Wald's identity which is used in the next section.

In section 8, we study positive recurrence. We show that in a weakly communicating recurrent class, either all states are positive recurrent or none is. Further, a weakly communicating positive recurrent class is necessarily strongly communicating.

Section 9 obtains the limits of the probabilities of n-step transition to a state and the average number of visits to a state. The results are very much the same as in the countably additive case and they are proved using a well-known renewal theorem.

In section 10, we prove a strong law of large numbers for recurrent chains, using the blocks theorem and the strong law of large numbers for i.i.d strategies. We also prove a few other related results on almost sure convergence.

The problem of mean convergence is taken up in section 11 and we prove a mean ergodic theorem for recurrent chains, when the initial distribution is concentrated on a single state.

We prove a ratio limit theorem in section 12 under more general initial distributions. We use this in the next section on stationary initial distributions.

Section 13 starts by proving the existence of a (finitely additive) stationary initial distribution for any Markov strategy, then proves the uniqueness of the countably additive part of the stationary initial distribution for a large class of Markov chains with a countable state space. An example is given to show that we cannot prove uniqueness of the stationary initial distribution in general.

The main result in this section is the construction of a canonical stationary initial distribution when the Markov chain is a positive recurrent class and identifying the limit in the strong law of large numbers as the integral with respect to this canonical stationary initial distribution.

In section 14, we prove the ergodicity of recurrent chains under stationary initial distributions. We also prove that the Markov strategic measure, under which the state space is a recurrent class, is countably additive on the shift invariant σ-field.

Section 15 proves the results of Blackwell [1], on almost closed sets, in the finitely additive setting. We also characterize the Feller boundary. For many of the results in this section we impose a mild condition on the Markov strategy, namely, that under the strategy, every state has positive probability of being reached.

In section 16 we prove that all bounded real valued solutions of the Choquet - Deny equations (on the set of integers with respect to a finitely additive measure γ defined on all subsets of the integers) are constants if γ has a non-trivial countably additive part or a non-trivial translation invariant part. As in [1] we use this to study random walks induced by such measures on integers. We prove that the Markov strategic measure corresponding to these random walks is 0-1 valued on the shift invariant σ-field. This helps us prove a 0-1 law in case of an i.i.d stategy for an important class of exchangeable sets.

In section 17 we take up the question of a Hewitt-Savage 0-1 law in the finitely additive case. We have not been able to prove it in its generality, nor could we give a counterexample. However we prove that, in case of an i.i.d strategy, an exchangeable G_δ set which is a countable intersection of exchangeable open sets, has measure zero or one.

The main result in section 18 is the Riesz decomposition theorem which asserts that every superharmonic function corresponding to a Markov strategy can be expressed uniquely as the sum of a harmonic function and a potential. We then use this theorem to obtain characterizations of recurrence and transience. We have followed the presentation in Neveu [28] of the corresponding results in the countably additive case.

1. The Basic set up. We start with an arbitrary non-empty set I which we shall refer to as the state space; the elements of I will be called states. N will throughout stand for the set of positive integers. We equip $H = I^N$ with the product of discrete topologies. Let I^* be the set of all finite sequences of elements of I, including the empty one. Let Γ be the set of all finitely additive probabilities defined on all subsets of I. The term 'probability' for us will mean a nonnegative, finitely - additive, normalised set function. We shall explicitly say 'countably - additive probability' in case the function is also countably additive.

Definition. A strategy σ is a function on I^* into Γ.

For $p, q \in I^*$ and $h \in H$, pq will stand for the element of I^* whose terms consist of the terms of p followed by the terms of q and ph will stand for the element of H whose terms consist of the terms of p followed by the terms of h. If $A \subseteq H$, $Ap = \{h \in H : ph \in A\}$.

Definition. If σ is a strategy and $p \in I^*$, the conditional strategy σ given p, denoted by $\sigma[p]$ is the strategy defined by $\sigma[p](q) = \sigma(pq)$ for all $q \in I^*$.

In [30], Purves and Sudderth have shown that every strategy σ indhces a probability on a field $\mathcal{A}(\sigma)$ of subsets of H including $\underline{\underline{B}}$, the σ-field of Borel subsets of H. This induced probability will also be denoted by σ. Thus if $p \in I^*$ and σ is a strategy, $\sigma[p]$ will stand for the conditional strategy σ given p as well as for the measure on $\mathcal{A}(\sigma[p])$ induced by this conditional strategy. It has been shown in [30] that the induced probability is unique subject to certain regularity conditions.

We shall state some of the basic properties of σ which we shall be using very often, but before that we need a few more definitions.

<u>Definition</u>. A stop rule s is a function on H into N such that if $s(h) = n$ and h' agrees with h through the first n coordinates, then $s(h') = n$.

<u>Definition</u>. An incomplete stop rule t is a function on H into $N \cup \{\infty\}$ such that if $t(h) = n$ for some $n \in N$ and h' agrees with h through the first n coordinates, then $t(h')=n$. For any $h \in H$ and $n \in N$, $p_n(h) = (h_1, \ldots, h_n)$, the finite sequence of the first n coordinates of h. If s is a stop rule, $p_s(h) = p_n(h)$ where $s(h) = n$.

Let $\{A^n\}$ be a sequence of subsets of H belonging to $\underline{\underline{B}}$. Let s be a stop rule. Then A^s will stand for the set $\bigcup_{n=1}^{\infty} [A^n \cap \{s(h) = n\}] = \{h : h \in A^{s(h)}\}$.

We now state below five basic properties of the probability σ which we shall need in the sequel.

I) For every $A \in \underline{\underline{B}}$ and every stop rule s,

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

II) For every open set $U \subseteq H$,

$$\sigma(U) = \text{Sup} \{\sigma(K) : K \subseteq U \text{ and } K \text{ clopen}\}$$

(inner regularity)

For every $A \in \underline{\underline{B}}$,

$$\sigma(A) = \text{Inf} \{\sigma(U) : U \supseteq A, \; U \text{ open}\}$$

(outer regularity)

It is easy to see, because of inner and outer regularity, that for every $A \in \underline{\underline{B}}$ and every $\varepsilon > 0$, there exists a clopen $K \subseteq H$ such that $\sigma(A \triangle K) < \varepsilon$, where $\triangle$ denotes the symmetric difference between the two sets.

A measure on the Borel σ-field of a topological space satisfying inner and outer regularity, as defined above, will be called regular.

III) Let $\{A^n\}$ be a sequence of sets in $\underline{\underline{B}}$.

a) If $\{A^n\}_{n \geq 1}$ is a non-decreasing sequence, then

$\sigma(A) = \text{Sup} \{ \sigma(A^s) : s \text{ a stop rule} \}$, where $A = \bigcup_n A^n$.

b) If $\{A^n\}_{n \geq 1}$ is a non-increasing sequence, then

$\sigma(A) = \text{Inf} \{ \sigma(A^s) : s \text{ a stop rule} \}$ where $A = \bigcap_n A^n$.

The following remark will be useful later. Suppose s_1 and s_2 are stop rules such that $s_1(h) \leq s_2(h)$ for all $h \in H$. Under the hypothesis of a) we have $A^{s_1} \subseteq A^{s_2}$ and under the hypothesis of b) we have $A^{s_1} \supseteq A^{s_2}$. Hence in the assertions of a) and b) it is enough to take the supremum and infinimum respectively, over all stop rules s such that $s(h) \geq s_o(h)$ for all $h \in H$ where s_o is any fixed stop rule.

IV) Let σ be a strategy and $\{A^n\}_{n \geq 1}$ a sequence of sets in $\underline{\underline{B}}$. For some $\varepsilon > 0$, suppose

$\sigma[p_n(h)] \, (A^n p_n(h)) \leq \frac{\varepsilon}{2^n}$ for all $h \in H$ and each $n \in N$.

Then $\sigma(\bigcup_{n=1}^{\infty} A^n) \leq \varepsilon$.

<u>Definition</u>. A clopen subset K of H is said to be determined by stop rule s if $K p_s(h) = H$ for $h \in K$

$= \emptyset$ for $h \in K^c$.

For every clopen set K of H there exists a stop rule s such that K is determined by s. A proof of this fact can be found in [14].

V) Let σ be any strategy. Let $\{L_n\}_{n \geq 1}$ be a sequence of clopen subsets of H and $\{t_n\}$ a sequence of strictly increasing stop rules (i.e. $t_m(h) > t_n(h)$ for all $m > n$ and for all $h \in H$) and $\{\alpha_n\}$ a sequence of real numbers such that

i) L_n is determined by t_n for all $n \in N$,

ii) $\sigma(L_1) \geq \alpha_1$ and for all $h \in H$,

$$\sigma[p_{t_n}(h)]\,(L_{n+1}\,p_{t_n}(h)) \geq \alpha_{n+1},\ n \in N.$$

Then $\sigma(\bigcap_{n=1}^{\infty} L_n) \geq \prod_{n=1}^{\infty} \alpha_n$.

The proofs of I - V can be found in [30].

We now proceed to define Markov strategies.

<u>Definition.</u> A strategy σ is called a Markov strategy if $\sigma(i_1,\ldots,i_n)$ depends only on n and i_n for all $n \in N$ and $i_1,\ldots,i_n \in I$.

<u>Definition.</u> A Markov strategy σ is said to be a Markov strategy with stationary transition probabilities in case $\sigma(i_1,\ldots,i_n) = \sigma(i_n)$ for all $i_1,\ldots,i_n \in I$ and $n \in N$.

Since we shall never have occasion to deal with Markov strategies which do not have stationary transition probabilities, in future when we say that σ is a Markov strategy, we shall mean that σ is a Markov strategy with stationary transition probabilities. The measure induced on $(H, \underline{\underline{B}})$ by a Markov strategy σ will be called a Markov measure and the process of coordinate maps will be said to form a Markov chain.

The measure σ_o which the strategy associates with the empty sequence will be called the initial distribution of the strategy. Many of our results do not depend on σ_o. In such cases, by abuse of notation we shall say that $\{\sigma(i)\}_{i \varepsilon I}$ is the Markov strategy. In case the initial distribution σ_o is fixed, we shall sometimes say that $\{\sigma_o, \{\sigma(i)\}_{i \varepsilon I}\}$ is the Markov strategy.

2. The Strong Markov Property. We begin this section by strengthening (I). This will help us formulate and prove a very useful property for Markov chains called the strong Markov property.

Lemma 2.1. Let t be an incomplete stop rule and let s be a stop rule on H. Then $s_1 = \min\ (s,t)$ is a stop rule.

This lemma is wellknown (see [14]) and since the proof is simple and straightforward, we omit it.

<u>Proposition 2.2</u>. Let t be an incomplete stop rule and let s be a stop rule on H. Suppose that $A \varepsilon \underline{\underline{B}}$ and that $A \subseteq \{t \leq s\}$ where $\{t \leq s\} = \{h \varepsilon H : t(h) \leq s(h)\}$. Then

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

<u>Proof</u>. Define $s_1 = \min(s,t)$. By lemma 2.1, s_1 is a stop rule. Thus by (I) of the previous section

$$\sigma(A) = \int \sigma[p_{s_1}(h)]\ (A\ p_{s_1}(h))\ d\ \sigma(h) \qquad \ldots.(2.1)$$

We first observe that for any h such that $t(h) > s(h)$, $A\ p_{s_1}(h) = \emptyset$. For otherwise, let $h' \varepsilon A\ p_{s_1}(h) = A\ p_s(h)$ (because in this case $s_1(h) = s(h)$). Then $h'' = p_s(h)\ h' \varepsilon A$. Since h'' and h agree through the first $s(h)$ coordinates, $s(h'') = s(h)$. Since $h'' \varepsilon A$, $t(h'') \leq s(h'') = s(h)$. Now h'' and h agree through the first $t(h'')$ coordinates. Therefore $t(h) \leq s(h)$, a contradiction ! Therefore $A\ p_{s_1}(h) = \emptyset$.

Again, for h such that $t(h) \leq s(h)$, $s_1(h) = t(h)$. Therefore, $p_{s_1}(h) = p_t(h)$.

Thus (2.1) reduces to the required result.

<u>Theorem 2.3</u>. Let t be an incomplete stop rule. Let $A \varepsilon \underline{\underline{B}}$ be such that $A \subseteq \{t < \infty\}$ where $\{t < \infty\} = \{h \varepsilon H : t(h) \varepsilon N\}$.

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

Proof. Let $A^n = A \cap (t \leq n)$, $n \in N$.

Since $A \subseteq \{t < \infty\}$, we have $A^n \uparrow A$. Therefore by (III a) of the previous section, $\sigma(A) = \text{Sup}\, \{\sigma(A^s): s \text{ a stop rule}\}$. It is easy to see that $A^s = A \cap \{t \leq s\}$. Therefore $A^s \subseteq \{t \leq s\}$.

By applying proposition 2.2, we get

$$\sigma(A) = \text{Sup}_s\, \sigma(A^s) = \text{Sup}_s \int_{\{t \leq s\}} \sigma[p_t(h)] (A^s\, p_t(h))\, d\, \sigma(h) \quad \text{....(2.2)}$$

We now claim that for h such that $t(h) \leq s(h)$,

$A^s\, p_t(h) = A\, p_t(h)$.

Since $A^s \subseteq A$, we have $A^s\, p_t(h) \subseteq A\, p_t(h)$.

For the other part, if $h' \in A\, p_t(h)$, then $h'' = p_t(h)\, h' \in A$. Now h'' and h agree through the first $t(h)$ coordinates and so $t(h'') = t(h)$. Also $t(h'') \leq s(h'')$; for if not, then $s(h'') < t(h'')$ and since h'' and h agree through $t(h'')$ coordinates, they agree through $s(h'')$ coordinates.

Then we have $s(h) = s(h'') < t(h'') = t(h)$, a contradiction.

Therefore $t(h'') \leq s(h'')$ and hence $h'' \in A \cap \{t \leq s\} = A^s$, consequently $h' \in A^s p_t(h)$ which shows that $A\ p_t(h) \subseteq A^s\ p_t(h)$ and hence $A\ p_t(h) = A^s\ p_t(h)$.

We have thus shown that

$$\sigma(A) = \sup_s \int_{\{t \leq s\}} \sigma[p_t(h)]\ (A\ p_t(h))\, d\,\sigma(h). \qquad \ldots(2.3)$$

From (2.3) it clearly follows that

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

Again, given $\varepsilon > 0$, there exists a stop rule s_o such that $\sigma(\{t < \infty\}) < \sigma(\{t \leq s_o\}) + \varepsilon$. (This follows from IIIa of previous section applied to $A^n = \{t \leq n\}$, $n \in N$).

Then
$$\begin{aligned}\sigma(A) &= \sup_s \int_{\{t \leq s\}} \sigma[p_t(h)]\ (A\ p_t(h))\ d\,\sigma(h)\\
&\geq \int_{\{t \leq s_o\}} \sigma[p_t(h)]\ (A\ p_t(h))\ d\,\sigma(h)\\
&\geq \int_{\{t < \infty\}} \sigma[p_t(h)]\ (A\ p_t(h))\ d\,\sigma(h)\\
&\qquad - [\sigma(\{t < \infty\}) - \sigma(\{t \leq s_o\})]\\
&\geq \int_{\{t < \infty\}} \sigma[p_t(h)]\ (A\ p_t(h))\ d\,\sigma(h) - \varepsilon.\end{aligned}$$

Since ε is arbitrary, the proof of the theorem is now complete.

<u>Definition</u>. Let $A \in \underline{\underline{B}}$ and t be an incomplete stop rule. We shall say that A is conditionally determined given t if

i) $A \subseteq \{t < \infty\}$, ii) there exists a $B \subseteq H$ such that $A\, p_t(h) = B$ for all h such that $t(h) < \infty$.

Theorem 2.4. (Strong Markov Property) Suppose that $A \in \underline{\underline{B}}$ is conditionally determined given t, an incomplete stop rule, and further that there is $i \in I$ such that for all $h \in \{t < \infty\}$, the $t(h)^{th}$ coordinate of h is i. Then

$\sigma(A) = \sigma[i]\,(B)\,\sigma(\{t < \infty\})$ where $B \in \underline{\underline{B}}$ is such that $A\, p_t(h) = B$ for all $h \in \{t < \infty\}$.

Proof. We know by Theorem 2.3 that

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

$$= \int_{\{t < \infty\}} \sigma[i]\,(B)\, d\,\sigma(h)$$ by the hypothesis and since σ is Markov,

$$= \sigma[i]\,(B)\,\sigma(\{t < \infty\}).$$

As we shall soon see, most of our results will be proved using the Strong Markov property.

3. Classification of states. For $j \in I$ and $n \in N$ let A_j^n denote the set of all $h \in H$ such that the n^{th} coordinate of h is j. Let $B_j^n = \bigcup_{k=1}^{n} A_j^k$, $j \in I$, $n \in N$. Let $A_j = \bigcup_{n=1}^{\infty} A_j^n = \bigcup_{n=1}^{\infty} B_j^n$. For $i, j \in I$, $\sigma[i]\,(A_j)$ will be denoted by f_{ij}^*.

Definition. Let $i, j \varepsilon I$. We shall say that i weakly leads to j (denoted by $i \xrightarrow{w} j$) in case $f^*_{ij} > 0$.

We shall say that i strongly leads to j (denoted by $i \xrightarrow{s} j$), if $\sigma[i](A^n_j) > 0$ for some $n \varepsilon N$.

Clearly if $i \xrightarrow{s} j$ then $i \xrightarrow{w} j$. Also if $i \xrightarrow{w} j$ and if σ is countably additive, that is, if $\sigma(p)$ is a countably additive probability for all $p \varepsilon I^*$, then $i \xrightarrow{s} j$. However $i \xrightarrow{w} j$ does not in general imply that $i \xrightarrow{s} j$. We give below an example to establish this.

Example 3.1. Let $I = N$. Let σ be a Markov strategy such that $\sigma(1) = \gamma$, a purely finitely additive measure and $\sigma(n+1) = \delta_n$ for $n \varepsilon N$ where δ_n is the Dirac measure at n, i.e.

$$\delta_n(E) = 1 \text{ if } n \varepsilon E$$
$$= 0 \text{ if } n \not\varepsilon E$$
$$\text{for all } E \subseteq I = N.$$

Since the relations $\xrightarrow{w}$ and $\xrightarrow{s}$ depend only on conditional strategies, it is not necessary while studying these relations to specify the initial distribution of the strategy.

In the above example we claim that $1 \xrightarrow{w} 1$, however $1 \not\xrightarrow{s} 1$ (read as "$1 \xrightarrow{s} 1$ is not true").

Proof. $\sigma[1](A^1_1) = \gamma(\{1\}) = 0$ since γ is diffuse (zero on singletons).

Again for $n \in N$,

$$\sigma[1]\,(A_1^{n+1}) = \int \sigma[1,i]\,(A_1^{n+1}\,i)\,d\gamma(i)$$ by (I) of section 1 applied to the stop rule $s \equiv 1$ and the change of variable theorem [15],

$$= \int_{\{i \neq 1\}} \sigma[i]\,(A_1^n)\,d\gamma(i) + \sigma[1]\,(A_1^{n+1}\,1).\,\gamma(1)$$

The second term is zero since γ is diffuse.

Therefore, $\sigma[1]\,(A_1^{n+1}) = \int_{\{i \neq 1\}} \sigma[i]\,(A_1^n)\,d\gamma(i)$.

Now note that for $i \neq 1$, $\sigma[i]\,(A_1^n) = 1$ if $n = i - 1$
$= 0 \quad n \neq i - 1$

Therefore, $\sigma[1]\,(A_1^{n+1}) = 0$, once again because γ is diffuse. This proves that $1 \overset{s}{\not\to} 1$.

Let s be the stop rule on H defined by $s(h) = h_1$, the first coordinate of h. Recall the definitions of $\{A_j^n\}$ and A_j^s. Then clearly $f_{ij}^* = \sigma[i]\,(A_j) \geq \sigma[i]\,(A_j^s)$. We shall now show that $\sigma[1]\,(A_1^s) = 1$ and that will show that $1 \overset{w}{\to} 1$.

$$\sigma[1]\,(A_1^s) = \int_{\{i \neq 1\}} \sigma[1,i]\,(A_1^s\,i)\,d\gamma(i)$$

$$= \int_{\{i \neq 1\}} \sigma[i]\,(A_1^{i-1})\,d\gamma(i).$$

Note that the integrand is identically equal to one and hence the proof is complete.

We next prove a lemma giving equivalent conditions for 'weakly leading to' and then proceed to prove the Chapman - Kolmogorov equations.

<u>Lemma 3.1</u>. For $i, j \in I$ the following are equivalent

a) $i \xrightarrow{w} j$.

b) $\sigma[i]\,(B_j^s) > 0$ for some stop rule s.

c) $\sigma[i]\,(A_j^s) > 0$ for some stop rule s.

<u>Proof</u>. a) $\Longrightarrow$ b). By definition $\sigma[i]\,(A_j) > 0$. Now $B_j^n \uparrow A_j$. Therefore by (III a) of section 1, $\sigma[i]\,(A_j) = \operatorname{Sup}_s \sigma[i]\,(B_j^s)$. Thus b) follows.

b) $\Longrightarrow$ c). Suppose s is a stop rule such that $\sigma[i]\,(B_j^s) > 0$. Consider the incomplete stop rule t defined by $t(h) = n$ if $h_n = j$ and $h_m \neq j$ for $1 \leq m \leq n-1$

$= \infty$ if such an n does not exist.

Then t is called the time of first occurrence of j. Let $s_1 = \min(s,t)$. By Lemma 2.1, s_1 is a stop rule. It is easy to see that $B_j^s \subseteq A_j^{s_1}$. Therefore $\sigma[i](A_j^{s_1}) \geq \sigma[i](B_j^s) > 0$.

c) $\Longrightarrow$ a) is trivial since $\sigma[i](A_j) \geq \sigma[i](A_j^s)$ for every stop rule s. Thus the proof of the lemma is complete.

<u>Remark</u>. From the proof it is also evident that

$$F^*_{ij} = \operatorname*{Sup}_s \sigma[i]\,(B^s_j) = \operatorname*{Sup}_s \sigma[i]\,(A^s_j) .$$

<u>Definition</u>. Let s_1, s_2 be stop rules. The stop rule s_1 composed with s_2, denoted by $s_1 * s_2$, is defined by

$s_1 * s_2\,(h) = s_1(h) + s_2(g)$ where $g \in H$ is such that

$g_n = h_{(s_1(h) + n)}$ for all $n \in N$.

It is easy to check that $s_1 * s_2$ is a stop rule.

<u>Theorem 3.2</u>. (Chapman - Kolmogorov Equations). Let $i, j \in I$ and s_1 , s_2 be stop rules.

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

where μ is the finitely additive probability defined on all subsets of I by $\mu(E) = \sigma[i]\,(\{h : h_{s_1(h)} \in E\}),\ E \subseteq I$.

<u>Proof</u>. The proof is carried out by applying (I) of section 1 to the stop rule s_1 and applying the change of variable theorem.

$$\sigma[i]\,(A_j^{s_1 * s_2}) = \int \sigma[i\,p_{s_1}(h)]\,(A_j^{s_1 * s_2}\,p_{s_1}(h))\,d\,\sigma[i](h)$$

$$= \int \sigma[h_{s_1}(h)]\,(A_j^{s_2})\,d\,\sigma[i](h)$$

because for all $h \in H$, $A_j^{s_1 * s_2}\,p_{s_1}(h) = A_j^{s_2}$. Now observe that under the transformation $h \longmapsto h_{s_1(h)}$ defined on $(H, \underline{\underline{B}})$ into

(I, P(I)) where P(I) is the set of all subsets of I, $\sigma[i]$ is carried over to the measure μ and the change of variable theorem completes the proof.

<u>Corollary 3.3</u>. Let i, j, k ε I. Then

$$i \xrightarrow{w} k \text{ and } k \xrightarrow{w} j \text{ implies that } i \xrightarrow{w} j \text{ ;}$$

$$\text{and } i \xrightarrow{s} k \text{ and } k \xrightarrow{s} j \text{ implies that } i \xrightarrow{s} j \text{ .}$$

That is to say, the relations " $\xrightarrow{w}$ " and " $\xrightarrow{s}$ " are transitive.

<u>Proof</u>. Suppose $i \xrightarrow{w} k$ and $k \xrightarrow{w} j$. By lemma 3.1, there exist stop rules s_1 and s_2 such that $\sigma[i]\,(A_k^{s_1}) > 0$ and $\sigma[k]\,(A_j^{s_2}) > 0$.

By theorem 3.2,

$$\sigma[i]\,(A_j^{s_1 * s_2}) = \int \sigma[l]\,(A_j^{s_2})\, d\,\mu(l) \quad \text{where } \mu \text{ is}$$

defined as in the theorem,

$$\geq \sigma[k]\,(A_j^{s_2})\,\mu(\{k\})$$

$$= \sigma[k]\,(A_j^{s_2}) \,.\, \sigma[i]\,(A_k^{s_1}) > 0 \,.$$

Thus by lemma 3.1 again, we have $i \xrightarrow{w} k$.

The proof of the other part is similar. We have only got to note in addition that if $s_1(h) = n_1$ and $s_2(h) = n_2$

for all $h \in H$ where $n_1, n_2 \in N$, then $s_1 * s_2 (h) = n_1 + n_2$ for all $h \in H$.

<u>Corollary 3.4</u>. For $i, j, k \in I$,

$$f^*_{ij} \geq f^*_{ik} f^*_{kj} .$$

<u>Proof</u>. Let s_1, s_2 be stop rules. Then as in the proof of the previous corollary it follows from theorem 3.2 that $\sigma[i] (A_j^{s_1 * s_2}) \geq \sigma[i] (A_k^{s_1}) \sigma[k] (A_j^{s_2})$. Taking supremum over all stop rules s_1, s_2 on both sides, we get

$$\sup_{s_1, s_2} \sigma[i] (A_j^{s_1 * s_2}) \geq \sup_{s_1} \sigma[i] (A_k^{s_1}) . \sup_{s_2} \sigma[k] (A_j^{s_2})$$

$$= f^*_{ik} \; f^*_{kj} .$$

Now $f^*_{ij} = \sup_s \sigma[i] (A_j^s) \geq \sup_{s_1, s_2} \sigma[i] (A_j^{s_1 * s_2})$

Thus the proof of the corollary is complete.

<u>Definition</u>. Let $i, j \in I$. We say that i weakly communicates with j, (denoted by $i \overset{w}{\longleftrightarrow} j$) in case $i \overset{w}{\longrightarrow} j$ and $j \overset{w}{\longrightarrow} i$. We say that i strongly communicates with j, (denoted by $i \overset{s}{\longleftrightarrow} j$) in case $i \overset{s}{\longrightarrow} j$ and $j \overset{s}{\longrightarrow} i$.

<u>Proposition 3.5</u>. On the set of all states such that $i \overset{w}{\longrightarrow} i$ ($i \overset{s}{\longrightarrow} i$), $\overset{w}{\longleftrightarrow}$ ($\overset{s}{\longleftrightarrow}$) is an equivalence relation.

Proof. Symmetry and reflexivity of the relations are immediate from definitions and transitivity follows from corollary 3.3.

Therefore the part of the state space where $\overset{W}{\longleftrightarrow}$ ($\overset{S}{\longleftrightarrow}$) is reflexive is partitioned into equivalence classes which will be called weakly communicating classes (strongly communicating classes). Since $i \overset{S}{\longleftrightarrow} j$ implies that $i \overset{W}{\longleftrightarrow} j$, each strongly communicating class is contained in a weakly communicating class.

Definition. We say that a state i is weakly essential if for $j \in I$, $i \overset{W}{\longrightarrow} j$ implies that $j \overset{W}{\longrightarrow} i$. We say that i is strongly essential if for $j \in I$, $i \overset{S}{\longrightarrow} j$ implies that $j \overset{S}{\longrightarrow} i$.

Proposition 3.6. If i is weakly (strongly) essential and $i \overset{W}{\longrightarrow} j$ ($i \overset{S}{\longrightarrow} j$) then j is weakly (strongly) essential. Thus in a weakly (strongly) communicating class either all states are weakly (strongly) essential or none is.

Proof. Same as in the countably additive case (See [8]).

Remark. We have already used the following consequence of the Chapman - Kolmogorov equations in some of the corollaries above. For all n_1, $n_2 \in N$, and $i,j,k \in I$, we have

$$\sigma[i](A_k^{n_1+n_2}) \geq \sigma[i](A_j^{n_1})\,\sigma[j](A_k^{n_2}).$$

As a result of this, for some of our future propositions, the same proof as in the countably additive case works. In such cases we omit the proof and give a reference for the proof in the countably additive case.

Definition. If $i \xrightarrow{s} i$, then the greatest common divisor d_i of $\{n \in N : \sigma[i](A_i^n) > 0\}$ is called the period of i.

Proposition 3.7 : All states within a strongly communicating class have the same period.

Proof. Same as in the countably additive case (See [8]).

We now give two examples, the first one to show that in a weakly communicating class, for some states period might be defined while for others it might not and the second example to show in a weakly communicating class, even among states for which period is defined, the period need not be the same.

Example 3.2. $I = N$, $\sigma(1) = \gamma$ where γ is diffuse, $\sigma(2) = \frac{1}{2}\delta_2 + \frac{1}{2}\delta_1$ and $\sigma(i) = \delta_{i-1}$ for all $i \geq 3$.

We first show in this example that I is a weakly communicating class, that is $i \xleftrightarrow{w} j$ for all $i, j \in I$. In view of transitivity of $\xrightarrow{w}$, it is enough to show that $1 \xrightarrow{w} i$ for all $i \geq 2$, $i \xrightarrow{w} 2$ for all $i \geq 2$ and finally that $2 \xrightarrow{w} 1$.

Now for $i \geq 2$,

$$f^*_{1i} = \int_{\{j > i\}} \sigma[1,j](A_i)\, d\gamma(j) \quad \text{since } \gamma \text{ is diffuse}$$

$$= \int_{\{j > i\}} \sigma[j](A_i)\, d\gamma(j)$$

$$\geq \int_{\{j > i\}} \sigma[j](A_i^{j-i})\, d\gamma(j)$$

$$= 1$$

Since $\sigma[j]\ (A_i^{j-1}) = 1$ for all $j > i$ where $i \geq 2$.

We have in fact shown that $f_{1i}^* = 1$ for $i \geq 2$ and hence $1 \xrightarrow{W} i$ for $i \geq 2$.

Now $2 \xrightarrow{W} 2$ since $f_{22}^* \geq \sigma[2]\ (A_2^1) = \frac{1}{2}$,

and $i \xrightarrow{W} 2$ for $i \geq 3$, since $f_{i2}^* \geq \sigma[i]\ (A_2^{i-2}) = 1$.

Finally $2 \xrightarrow{W} 1$ since $f_{21}^* \geq \sigma[2]\ (A_1^1) = \frac{1}{2}$.

It is easy to see as in example 3.1 that $i \overset{s}{\not\longrightarrow} i$ for $i \neq 2$. Therefore period is not defined for $i \neq 2$. However, period of 2 is 1 since $\sigma[2]\ (A_2^1) = \frac{1}{2}$.

Example 3.3. $I = N$, $\sigma(1) = \frac{1}{2}\delta_1 + \frac{1}{2}\gamma$ where γ is diffuse,

$$\sigma(2) = \frac{1}{2}\delta_3 + \frac{1}{2}\delta_1,\ \sigma(i) = \delta_{i-1}, \text{ for all } i \geq 3.$$

It is easy to see as in the previous example that I is a weakly communicating class. Now $\sigma[1]\ (A_1^1) > 0$. Therefore the period of 1 is 1. We now claim that the period of 2 is 2. This is because it can be shown by induction that $\sigma[2](A_2^n) > 0$ if and only if n is even.

Proposition 3.8. Let $C(i)$ be the strongly communicating class containing i. Let d be the period of i (hence of every state in $C(i)$). To each $j \in C(i)$ there corresponds an integer r_j, $0 \leq r_j \leq d-1$ such that $\sigma[i]\ (A_j^n) > 0$ if and only if $n \equiv r_j \pmod d$.

Proof. Same as in the countably additive case (See [8]).

Proposition 3.8 enables us to partition $C(i)$ into d subclasses $C_r(i)$, $0 \leq r \leq d-1$, such that $j \in C_r(i)$ if and only if the residue class corresponding to j is r. For every integer r, we define $C_r(i) = C_{r'}(i)$ where $r = r'$ (mod d). Suppose $j \in C_r(i)$. It can be shown that $C_{r'}(j) = C_{r+r'}(i)$.

It is for this reason that these subclasses of a strongly communicating class are referred to as the cyclically moving subclasses.

We end this section with a few examples indicating differences with the countably additive theory. In a countably additive Markov chain with a countable state space, if C is an essential class (in this theory the notions of weakly essential and strongly essential are equivalent) of period d and if C_r, $0 \leq r \leq d-1$, are the cyclic subclasses and if $i \in C_r$ then $\sigma[i]\left(\bigcup_{j \in C_{r+n}} A_j^n\right) = 1$ for every positive integer n.

We give below an example to show that this is far from true in the general case.

Example 3.4. Let $I = N$, $\sigma(1) = \gamma$, a diffuse probability and $\sigma(i) = \delta_1$ for all $i \geq 2$.

In this example $\{1\}$ is a strongly as well as weakly communicating class and 1 is strongly as well as weakly essential. However, it can easily be shown by induction that $\sigma[1](A_1^{2n}) = 1$ for all $n \in N$ and $\sigma[1](A_1^{2n-1}) = 0$ for all $n \in N$.

Although the properties of being strongly essential and weakly essential are equivalent in the countably additive case, they are not in general comparable.

Example 3.5. Let $I = N$, $\sigma(1) = \frac{1}{2}\delta_1 + \frac{1}{2}\gamma$ where γ is diffuse, $\sigma(2) = \delta_2$ and $\sigma(i) = \delta_{i-1}$ for all $i \geq 3$.

Here 1 is strongly essential since $1 \xrightarrow{s} 1$ and $1 \not\xrightarrow{s} i$ for $i \geq 2$. But 1 is not weakly essential since $1 \xrightarrow{w} 2$ but $2 \not\xrightarrow{w} 1$. In Example 3.1, N is a communicating class which is weakly essential since $i \xleftrightarrow{w} j$ for all $i, j \in N$. However 2 is not strongly essential since $2 \xrightarrow{s} 1$ but $1 \not\xrightarrow{s} 2$.

4. Closed Sets.

Definition. A nonempty subset E of I is called a closed set if $\sigma(j)(E) = 1$ for all $j \in E$. A closed set E is called minimal closed if no proper subset of E is closed.

For $E \subseteqq I$, let E^n stand for the subset of H defined by $E^n = \{h \varepsilon H : h_i \varepsilon E \text{ for } 1 \leq i \leq n\}$, $n \varepsilon N$. Let $E^N = \bigcap_{n=1}^{\infty} E^n$ (This notation agrees with the usual meaning of E^N).

<u>Theorem 4.1</u>. If E is a closed set, then for each $j \varepsilon E$, $\sigma[j] (E^N) = 1$.

<u>Proof</u>. Since $E^n \downarrow E^N$, by (III b) of section 1, for each $j \varepsilon I$, $\sigma[j] (E^N) = \operatorname{Inf}_s \sigma[j] (E^s)$.

We shall prove by induction on the structure of stop rules that $\sigma[j] (E^s) = 1$ for every stop rule s and every $j \varepsilon E$. Let us first show that the result is true for all stop rules of structure zero. It is true by definition of a closed set for $s \equiv 1$. Assume that we have proved the result for $s \equiv n-1$. Let now $s \equiv n$.

For $j \varepsilon E$,

$$\sigma[j] (E^n) = \int \sigma[j\, x] (E^n x)\, d\, \sigma(j)\, (x)$$

$$= \int_E \sigma[x] (E^n x)\, d\, \sigma(j)\, (x) \quad \text{since } \sigma(j)\,(E) = 1 \text{ and } \sigma \text{ is Markov}$$

$$= \int_E \sigma[x] (E^{n-1})\, d\, \sigma(j)\, (x) \quad \text{since for all } x \varepsilon E,\ E^n x = E^{n-1}$$

$$= 1 \quad \text{by the induction hypothesis.}$$

Thus the result is true for all stop rules of structure zero. Assume that the result is proved for all stop rules of structure less than α. Let s now be a stop rule of structure α. Let $j \varepsilon E$. Then

$$\sigma[j](E^{s+1}) = \int \sigma[jx](E^{s+1}x) \, d\,\sigma(j)(x)$$

$$= \int_E \sigma[x](E^{s_x}) \, d\,\sigma(j)(x)$$

where for $x \varepsilon E$, s_x is the stop rule defined by $s_x(h) = s(xh)$ for all h. Clearly for each x, s_x is a stop rule of structure less than α and thus by the induction hypothesis.

$$\sigma[j](E^{s+1}) = 1 .$$

Therefore $\sigma[j](E^s) \geq \sigma[j](E^{s+1}) = 1$.

Hence the theorem is proved.

Proposition 4.2. The union of finitely many weakly (strongly) communicating class none of which is weakly (strongly) essential, cannot be closed.

Proof. Same idea as in the countably additive case (See [8]).

Remark. While in the countably - additive theory with a countable state space it is true that an essential class is minimal closed, it is no longer true in general.

Example 4.1. $I = N$. $\sigma(1) = \frac{1}{2}\gamma + \frac{1}{2}\left(\sum_{n=0}^{\infty} p_{2n+1}\,\delta_{2n+1}\right)$

where γ is diffuse and such that it gives probability one to the set of even integers and p_{2n+1} are positive real numbers such that $\sum_{n=0}^{\infty} p_{2n+1} = 1$, $\sigma(2n) = \delta_{2n}$, $n \in N$ and $\sigma(2n+1) = \delta_{2n-1}$, $n \in N$.

It is easy to see in the above example that the set of odd integers is a weakly as well as strongly communicating class and further it is weakly as well as strongly essential. However $\sigma(1)$ (odd integers) $= \frac{1}{2}$. Hence the set of odd integers is not closed.

However if a set of weakly communicating states is closed, it is minimal closed. For if E is one such set and $i \in E$ then for every $j \in E - \{i\}$ $i \overset{w}{\longleftrightarrow} j$, so $\sigma[j]$ (h : all coordinates of h belong to $E - \{i\}$) < 1. Hence $E - \{i\}$ cannot be closed.

5. Recurrent and Transient states.

Definition. A state i is called recurrent if $f^*_{ii} = 1$. Otherwise it is called transient.

For $j \in I$, let G_j stand for the set of all $h \in H$ such that infinitely many coordinates of h are j.

Let $g_{ij} = \sigma[i]\,(G_j)$, $i, j \in I$.

Proposition 5.1. $g_{ij} = f^*_{ij}\, g_{jj}$ for $i, j \in I$

Proof. Let t be the time of first occurrence of j. Clearly G_j is conditionally determined given t and $G_j\, p_t(h) = G_j$ for all $h \in \{t < \infty\}$. Therefore by the Strong Markov property,

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

Theorem 5.2. For $i, j \in I$, $g_{ij} = \begin{cases} f^*_{ij} & \text{if } j \text{ is recurrent} \\ 0 & \text{if } j \text{ is transient}. \end{cases}$

Proof. Let $\{G^n_j\}$ be the sequence defined by $G^n_j = \{h \in H$: at least n coordinates of h are $j\}$ and let $g_{ij}(n)$ denote $\sigma[i]\,(G^n_j)$, $n \in N$. Let t be the time of the first occurrence of j. Observe that G^{n+1}_j is conditionally determined given t and $G^{n+1}_j\, p_t(h) = G^n_j$ for all $h \in \{t < \infty\}$, $n \in N$. Then by the Strong Markov property,

$$g_{ij}(n+1) = \sigma[j]\,(G^n_j)\,\sigma[i]\,(\{t < \infty\}) = g_{jj}(n).f^*_{ij} \qquad \ldots(5.1)$$

By repeated application of (5.1), we get

$$g_{ij}(n+1) = f^*_{ij}\,(f^*_{jj})^n,\ n \in N \qquad \ldots(5.2)$$

In fact (5.2) is true by definitions for $n = 0$ as well.

If j is transient, i.e. if $f^*_{jj} < 1$,

$$\begin{aligned} g_{ij} &\leq \operatorname*{Inf}_s \sigma[i](G^s_j) \quad \text{(by III b of section 1 since } G^n_j \downarrow G_j) \\ &\leq \operatorname*{Inf}_n \sigma[i](G^{n+1}_j) \\ &= \operatorname*{Inf}_n f^*_{ij}(f^*_{jj})^n \\ &= 0. \end{aligned}$$

Therefore we have shown that if j is transient, then $g_{ij} = 0$. Suppose now that j is recurrent i.e. $f^*_{jj} = 1$. By (5.2) $g_{ij}(n) = f^*_{ij}$ for all $n \in N$, which is equivalent to saying that $\sigma[i](G^s_j) = f^*_{ij}$ for all stop rules of structure 0 and for all $i, j \in I$.

Assume that we have proved that $\sigma[i](G^s_j) = f^*_{ij}$ for all stop rules of structure less than α and for all $i, j \in I$. Let s now be a stop rule of structure α. By proposition 2.3,

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

where for $h \in \{t < \infty\}$, s_h is the stop rule defined by $s_h(h') = s(p_t(h)\, h')$ for all $h' \in H$. Since for each $h \in \{t<\infty\}$ s_h is a stop rule of structure strictly less than α, by the induction hypothesis,

$$\sigma[i]\,(G_j^{s+1}) = \int_{\{t<\infty\}} f_{jj}^* \, d\,\sigma[i](h) = f_{ij}^* \quad (\text{since } f_{jj}^* = 1).$$

So $f_{ij}^* = \sigma[i]\,(G_j^1) \geq \sigma[i]\,(G_j^s) \geq \sigma[i]\,(G_j^{s+1}) = f_{ij}^*$.

Therefore $\sigma[i]\,(G_j^s) = f_{ij}^*$.

The proof of the theorem is thus complete.

As an immediate consequence, we get the following

<u>Corollary 5.3</u>. $g_{ii} = 1$ if i is recurrent

$= 0$ if i is transient.

<u>Proposition 5.4</u>. A recurrent state is weakly essential.

<u>Proof</u>. Let i be recurrent and let $i \xrightarrow{w} j$, i.e. $f_{ij}^* > 0$.

We know that, $1 = g_{ii} = \sigma[i]\,(G_i)$

$$= \sigma[i]\,(G_i \cap A_j) + \sigma[i]\,(G_i \cap A_j^c)$$

(where A_j is defined as is section 3).

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

Therefore $\sigma[i]\,(G_i \cap A_j) > 0$. But $G_i \cap A_j$ is conditionally determined given t, the time of first occurrence of j, and $(G_i \cap A_j)\, p_t(h) = G_i$ for all $h \in \{t < \infty\}$.

So by the Strong Markov property,

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

Therefore $g_{ji} > 0$. But by theorem 5.2, $f^*_{ji} = g_{ji}$. Therefore $j \xrightarrow{w} i$. Hence i is weakly essential.

Remark. A recurrent state need not be strongly essential. In Example 3.1, it is easy to see that $f^*_{21} = 1$ and $f^*_{12} = 1$. So by corollary 3.4, $f^*_{22} \geq f^*_{21} f^*_{12} = 1$. Hence 2 is recurrent. However 2 is not strongly essential because $2 \xrightarrow{s} 1$ but $1 \not\xrightarrow{s} 2$.

Theorem 5.5. 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

$$s'(h) < s(h) \text{ for all } h \in H. \text{ Then}$$

$$\sigma[i](\{h : h_n = i \text{ for some } n \geq s(h) \text{ and } h_n \neq j \text{ for all } n \geq s'(h)\})$$

$$\leq \sigma[i]'\{h : (\exists m \geq s(h))\ (h_n \neq j \text{ for } s'(h) \leq n \leq s(h), h_n \neq i \text{ for } s(h) \leq n < m, h_m = i, h_n \neq j \text{ for } n > m)\})$$

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

Define $t(h) = n$ if $n \geq s(h)$, $h_n = i$, $h_m \neq j$ for $s'(h) \leq m \leq s(h)$ and $h_m \neq i$ for $s(h) \leq m < n$,

$= \infty$ if no such n exists

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,

$$\sigma[i](C_{s,s'}) = \sigma[i](A_j^c)\sigma[i](\{t < \infty\})$$

$$= (1 - f_{ij}^*)\sigma[i](\{h: h_n = i \text{ for some } n \geq s(h) \text{ and } h_m \neq j,\ s'(h) \leq m \leq s(h)\}).$$

Taking infimum over all stop rules s such that $s(h) \geq s'(h)$ for all h on the right side of the above equation and using (III) of section 1, we get

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

$$\leq (1 - f_{ij}^*)\sigma[i](\{h: h_n = i \text{ for infinitely many } n,\ h_n \neq j \text{ for } n \geq s'(h)\})$$

Since $f_{ij}^* > 0$, it follows that $\sigma[i](\{h: h_n = i \text{ for infinitely many } n, h_n \neq j \text{ for } n \geq s'(h)\}) = 0$.

This being true for every stop rule s', by (III) of section 1 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 5.6. If i is recurrent and $i \xrightarrow{w} j$, then j is recurrent; consequently in a weakly communicating class either all states are recurrent or none is. Further $g_{ij} = g_{ji} = 1$.

Proof. Under the hypothesis, by theorem 5.5, we have $g_{ij} = 1$. Since by proposition 5.1, $g_{ij} = g_{jj} f^*_{ij}$, we must have $g_{jj} = 1$, so by corollary 5.3, j is recurrent. Since i is recurrent, by proposition 5.4, i is weakly essential. Therefore $j \xrightarrow{w} i$. Consequently by theorem 5.5 once again, $g_{ji} = 1$. The proof of the theorem is now complete.

6. The Blocks Theorem. Let i be a recurrent state. For each $n \in N$, let $t(n)$ be the incomplete stop rule corresponding to the n^{th} occurrence of i. Observe that $t(n)$ is finite for each $n \in N$ on G_i, which is a set of $\sigma[i]$-measure one, since i is recurrent. Let F stand for the set of those elements in I^*, whose last coordinate is i, the remaining coordinates being different from i.

Define $\{\beta_n\}$, $n \in N$ a sequence of functions on G_i into F as follows :

$$\beta_1(h) = p_{t(1)}(h), \quad \beta_{n+1}(h) = (h_{t(n)(h)+1}, h_{t(n)(h)+2}, \ldots h_{t(n+1)(h)}), \qquad n \in N, \; h \in G_i .$$

Call β_n the n^{th} i-block variable.

Let γ be the probability defined on all subsets of F by

$$\gamma(C) = \sigma[i] (\beta_1^{-1}(C)), \ C \subseteq F.$$

Let $\Omega = F^N$ be equipped with the product of discrete topologies and $\underline{\underline{F}}$ be the σ-field of Borel subsets of Ω. Let π be the i.i.d strategy on F defined by $\pi(q) = \gamma$ for all finite sequences q of elements of F. This strategy induces a probability, also denoted by π on $(\Omega, \underline{\underline{F}})$.

Let Φ be the mapping on G_i into Ω as follows :

$$\Phi(h) = (\beta_1(h), \beta_2(h), \ldots).$$

Lemma 6.1. Φ is a homeomorphism from G_i, endowed with the relative topology, onto Ω .

Proof. Omitted since it is straightforward.

Proposition 6.2. Let m and m' be regular probabilities (see section 1 for definition) defined on the Borel σ-fields $\underline{\underline{C}}$ and $\underline{\underline{C}}'$ of topological spaces Y and Y' respectively. Suppose (i) Ψ is a homeomorphism on Y onto Y', and

(ii) for every clopen set in Y', $m'(K) = m(\Psi^{-1}(K))$.

Then for every $C \in \underline{\underline{C}}'$, $m'(C) = m(\Psi^{-1}(C))$.

Proof. Let U be an open set in Y'. Then by inner regularity of m'

$$m'(U) = \text{Sup}\{m'(K) : K \text{ clopen and } K \subseteq U\}$$
$$= \text{Sup}\{m(\Psi^{-1}(K)) : K \text{ clopen and } K \subseteq U\} \text{ by (ii),}$$
$$\leq m(\Psi^{-1}(U)).$$

Also Ψ being a homeomorphism $\Psi^{-1}(U)$ is open and by inner regularity of m, given $\varepsilon > 0$, there exists a clopen set $D \subseteq \Psi^{-1}(U)$ such that $m(\Psi^{-1}(U)) \leq m(D) + \varepsilon$. Since Ψ is a homeomorphism $\Psi(D))$ is clopen in Y' and $\Psi^{-1}(\Psi(D)) = D$. Therefore by (ii) $m(\Psi^{-1}(U)) \leq m'(\Psi(D)) + \varepsilon \leq m'(U) + \varepsilon$. This being true for every $\varepsilon > 0$, $m(\Psi^{-1}(U)) \leq m'(U)$. We have therefore shown that $m'(U) = m(\Psi^{-1}(U))$ for open $U \in \underline{C}'$.

Now take a general $C \in \underline{C}'$. By outer regularity of m'

$$m'(C) = \text{Inf}\{m'(U) : U \text{ open and } U \supseteq C\}$$
$$= \text{Inf}\{m(\Psi^{-1}(U)) : U \text{ open and } U \supseteq C\}$$
$$\geq m(\Psi^{-1}(C)).$$

Further by outer regularity of m, given $\varepsilon > 0$, there exists an open set $V \supseteq \Psi^{-1}(C)$ such that $m(V) \leq m(\Psi^{-1}(C)) + \varepsilon$. As Ψ is a homeomorphism, $\Psi(V)$ is open and $\Psi^{-1}(\Psi(V)) = V$. Therefore $m'(\Psi(V)) \leq m(\Psi^{-1}(C)) + \varepsilon$

Hence $m'(C) \leq m'(\Psi(V)) \leq m(\Psi^{-1}(C)) + \varepsilon$.

As ε is arbitrary, it follows that $m'(C) \leq m(\Psi^{-1}(C))$.

Therefore $m'(C) = m(\Psi^{-1}(C))$ for all $C \in \underline{C}'$.

<u>Theorem 6.3 (Blocks Theorem)</u>. Let i be a recurrent state. Let Φ, π, Ω, $\underline{\underline{F}}$ be defined as before Lemma 6.1. Then for every $B \in \underline{\underline{F}}$,

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

<u>Proof</u>. Clearly π on $(\Omega, \underline{\underline{F}})$ is regular as π is strategic. Again $\sigma[i]$ restricted to G_i is regular, since $\sigma[i]$ is regular on H and $\sigma[i](G_i) = 1$. Thus in view of Lemma 6.1 and proposition 6.2, it is enough to prove that $\pi(K) = \sigma[i] (\Phi^{-1}(K))$ for all sets K, K clopen. We prove this by induction on the structure of the clopen set K. If $K = \emptyset$, $\pi(K) = 0 = \sigma[i] (\Phi^{-1}(K))$.
If $K = \Omega$, $\sigma[i] (\Phi^{-1}(\Omega)) = \sigma[i] (G_i) = 1 = \pi(\Omega)$.
Hence the result is true for clopen sets of structure zero.

Now assume that the result is true for all clopen sets of structure less than α. Let K be a clopen set of structure α.

Then $\pi(K) = \int \pi(Kx) \, d\gamma(x)$

$= \int \sigma[i] (\Phi^{-1}(Kx)) \, d\gamma(x)$ by the induction hypothesis

$= \int_{\{t(1) < \infty\}} \sigma[i] (\Phi^{-1}(K p_{t(1)}(h))) \, d\sigma[i](h)$

The last step is obtained by first observing that the map $h \longmapsto p_{t(1)}(h)$ on $\{t(1) < \infty\}$ into F carries the measure $\sigma[i]$ to λ and then using the change of variable theorem.

$$\text{So} \quad \pi(K) = \int_{\{t(1) < \infty\}} \sigma[i]\,(\Phi^{-1}(K\,p_{t(1)}(h))\,d\,\sigma[i](h),$$

$$= \int_{\{t(1) < \infty\}} \sigma[i]\,\{(\Phi^{-1}(K))\,p_{t(1)}(h)\}\,d\,\sigma[i](h)$$

$$= \sigma[i]\,(\Phi^{-1}(K)) \quad \text{by Theorem 2.3.}$$

Thus the proof of the theorem is complete.

7. Some results on Finitely Additive Integration.

So far we only needed to integrate bounded functions with respect to probabilities. In order to study the notion of positive recurrence we need to integrate unbounded functions, more precisely, we need to integrate incomplete stop rules with respect to Markov strategic measures.

Dunford and Schwartz [15] give us a definition where the integral of a non-negative function g w.r.t a measure P is defined as $\int g\,d\,P = \lim_n \int g \wedge n\,d\,P$ where $g \wedge n = \min(g,n)$, and the integral for a general real-valued g is defined by $\int g\,d\,P = \int g^+ d\,P - \int g^- d\,P$ where g^+ and g^- are the positive and negative parts of g respectively, whenever the right side makes sense. Otherwise $\int g\,d\,P$ is not defined.

For finitary functions on H and for a strategic measure σ, Dubins and Savage [14] have given an alternative definition of $\int g \, d\sigma$ by induction on the structure of g. These two definitions of integrals do not even coincide on proper stop rules integrated with respect to Markov strategic measures ; a minor modification of an example due to Robert Chen [7] shows this.

<u>Example 7.1</u>. $I = N$, $\sigma(\emptyset) = \gamma$ a diffuse measure,

$$\sigma(i) = \frac{1}{i} \sum_{j=1}^{i} \delta_j \ , \ i \in N .$$

Let s be the stop rule defined by

$$s(h) = h_1 + 2 \quad \text{if } h = (h_1, h_2, \dots) \text{ and } h_1 = h_2 .$$
$$= 2 \quad \text{otherwise.}$$

It can be seen easily that $\sigma(\{h : h_1 = h_2\}) = 0$. Therefore $\sigma(\{h : s(h) = 2\}) = 1$. Consequently the integral of s in the Dunford-Schwartz sense is equal to 2. Again for fixed $n \geq 2$, the function $s \circ n$ defined by $s \circ n\,(h) = s(n,h)$ takes values 2 and $2 + n$ with probabilities $\frac{n-1}{n}$ and $\frac{1}{n}$ respectively. Therefore $\int s \circ n \, d\sigma[n] = 3$ for every $n \geq 2$. Hence the Dubins - Savage integral of s with respect to σ is 3.

We shall use the Dunford-Schwartz definition of the integral since this yields desirable results. In this section we prove a few results on finitely additive integration theory needed later.

For the rest of this section let Y be an arbitrary non-empty set, $\underline{\underline{C}}$ a σ-field of subsets of Y and P a finitely additive probability on $\underline{\underline{C}}$.

Definition. A random variable on $(Y, \underline{\underline{C}})$ is an extended - real valued measurable function on $(Y, \underline{\underline{C}})$.

Proposition 7.1. Let ξ be a non-negative random variable on $(Y, \underline{\underline{C}})$ (can assume $+\infty$). If $\lim_n P(\{\xi \geq n\}) = 0$, then

$$\lim_n \int_{\{\xi < n\}} \xi \, dP < \infty \text{ if and only if } \sum_{n=1}^{\infty} P(\{\xi \geq n\}) < \infty .$$

Proof.
$$\lim_n \int_{\{\xi < n\}} \xi \, dP \leq \sum_{k=0}^{\infty} (k+1) P(\{k \leq \xi < k+1\})$$
$$= \sum_{n=0}^{\infty} \sum_{k=n}^{\infty} P(\{k \leq \xi < k+1\})$$
$$\leq \sum_{n=0}^{\infty} P(\{\xi \geq n\})$$
$$= \sum_{n=1}^{\infty} P(\{\xi \geq n\}) + 1 .$$

Again $\lim_{n} \int_{\{\xi < n\}} \xi \, dP \geq \sum_{k=0}^{\infty} k \, P(\{k \leq \xi < k+1\})$

$$= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P(\{k \leq \xi < k+1\})$$

$$= \sum_{n=1}^{\infty} P(\{\xi \geq n\}) .$$

The last step follows since $\lim_{n} P(\{\xi \geq n\}) = 0$ implies that $P(\{\xi \geq n\}) = \sum_{k=n}^{\infty} P(\{k \leq \xi < k+1\})$ for each n.

The result now follows immediately.

Proposition 7.2. Let ξ be a nonnegative random variable on $(Y, \underline{\underline{C}})$. If $\lim_{n} P(\{\xi \geq n\}) = 0$, then $\int \xi \, d P = \lim_{n} \int_{\{\xi < n\}} \xi \, d P$.

Proof. $\int \xi \, d P = \lim_{n} \int \xi \wedge n \, d P = \lim_{n} \left[\int_{\{\xi < n\}} \xi \, dP + n \, P(\{\xi \geq n\}) \right]$ ....(7.1)

If $\sum_{n=1}^{\infty} P(\{\xi \geq n\}) = \infty$, by proposition 7.1,

$$\lim_{n} \int_{\{\xi < n\}} \xi \, dP = \infty$$

and hence so is $\int \xi \, dP$ by (7.1).

If $\sum_{n=1}^{\infty} P(\{\xi \geq n\}) < \infty$, then $\operatorname{Lim}_n \; n\, P(\{\xi \geq n\}) = 0$ and hence (7.1) once again gives us $\int \xi \, dP = \operatorname{Lim}_n \int_{\{\xi < n\}} \xi \, dP$.

The proposition is hence proved.

Corollary 7.3. Let ξ be a random variable on $(Y, \underset{=}{C})$. Then

$$\int \xi \, d\,P < \infty \quad \text{iff} \quad \sum_{n=1}^{\infty} P(\{|\xi| \geq n\}) < \infty .$$

Proof. No loss of generality in assuming that ξ is non-negative since for the general case we can work with the positive and negative parts of ξ. If $P(\{\xi \geq n\}) \not\longrightarrow 0$ as $n \longrightarrow \infty$, then $\int \xi \, d\,P = \infty = \sum_{n=1}^{\infty} P(\{\xi \geq n\})$. If $P(\{\xi \geq n\}) \longrightarrow 0$, propositions 7.1 and 7.2 give us the result.

Theorem 7.4. Let ξ be a non-negative extended integer valued random variable on $(Y, \underset{=}{C})$. Then

$$\int \xi \, d\,P = \sum_{n=1}^{\infty} P(\{\xi \geq n\}).$$

Proof. If $P(\{\xi \geq n\})$ does not converge to zero, then $\sum_{n=1}^{\infty} P(\{\xi \geq n\}) = \infty$, and also $\int \xi \, d\,P = \infty$ because of (7.1).

If $\operatorname{Lim}_n P(\{\xi \geq n\}) = 0$, by proposition 7.2,

$$\int \xi \, d P = \operatorname{Lim}_n \int_{\{\xi < n\}} \xi \, d P = \operatorname{Lim}_n \sum_{k=1}^{n} k \, P(\{\xi = k\})$$

$$= \sum_{k=1}^{\infty} k \, P(\{\xi = k\})$$

$$= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P(\{\xi = k\})$$

$$= \sum_{n=1}^{\infty} P(\{\xi \geq n\}).$$

For the last step, observe that $\operatorname{Lim}_n P(\{\xi \geq n\}) = 0$ implies that $P(\{\xi \geq n\}) = \sum_{k=n}^{\infty} P(\{\xi = k\})$.

<u>Proposition 7.5</u>. Let $\{\xi_n\}$ be a sequence of non-negative random variables on $(Y, \underline{\underline{C}})$. Suppose

i) If for some $x \in Y$ and $n_o \in N$, $\xi_{n_o}(x) = 0$, then $\xi_n(x) = 0$ for all $n \geq n_o$.

ii) There exists a real number $c > 0$ such that $P(\ 0 < \xi_n < c\ ) = 0$ for all $n \in N$. (Note that ii) is automatically satisfied if ξ_n's are integer - valued).

iii) $\operatorname{Lim}_n P(\{\xi \geq n\}) = 0$ where $\xi = \sum_{n=1}^{\infty} \xi_n$.

Then $\int \xi \, d P = \sum_{n=1}^{\infty} \int \xi_n \, d P$.

Proof. Since ξ_n's are non-negative, we have

$$\int \xi \, dP \geq \sum_{n=1}^{\infty} \int \xi_n \, dP$$

because for every $k \in N$, $\int \xi \, dP \geq \int (\sum_{n=1}^{k} \xi_n) dP = \sum_{n=1}^{k} \int \xi_n dP$.

On the other hand since $\lim_n P(\{\xi \geq n\}) = 0$, by proposition 7.2

$$\int \xi \, dP = \lim_n \int_{\{\xi < n\}} \xi \, dP$$

$$= \lim_n \int_{\{\xi < n\}} (\xi_1 + \ldots + \xi_{[n/c]+1}) d\, P \quad \text{because of i) and ii)}$$

$$\leq \lim_n \sum_{k=1}^{[n/c]+1} \int \xi_k d\, P \quad ([n/c] \text{ denotes the integral part of } n/c)$$

$$\leq \sum_{n=1}^{\infty} \int \xi_n d\, P .$$

It thus follows that $\int \xi \, dP = \sum_{n=1}^{\infty} \int \xi_n d\, P$ and the theorem is proved.

We conclude this section by proving a finitely additive version of Wald's identity which we shall use in the next section.

Definition. Let ξ be a non-negative random variable on $(Y, \underset{=}{C})$. Let $A \in \underset{=}{C}$. We say that A is orthogonal to ξ in case

$$\int_A \xi \, dP = (\int_Y \xi \, dP) . P(A)$$

(Convention : $0 . \infty = 0$).

<u>Theorem 7.6</u> (Wald's Identity). Let $\{\xi_n\}$ be sequence of strictly positive random variables defined on $(Y, \underline{\underline{C}})$. Let τ be a positive extended integer valued random variable on $(Y, \underline{\underline{C}})$. Suppose

i) For every Borel subset A of the extended real line, $P(\xi_n^{-1}(A)) = P(\xi_1^{-1}(A))$ for all $n \varepsilon N$.

ii) There exists a real number $C > 0$ such that $P(\{\xi_n \geq C\})=1$ for all $n \varepsilon N$.

iii) $\{\tau < n\}$ is orthogonal to ξ_n for all $n \varepsilon N$.

Then $\int S_\tau d P = (\int \xi_1 d P)(\int \tau d P)$

where $S_\tau(x) = \xi_1(x) + \ldots + \xi_{\tau(x)}(x)$ for all $x \varepsilon Y$

$$= \sum_{n=1}^{\infty} \xi_n(x) \text{ if } \tau(x) = \infty .$$

<u>Proof</u>. Since τ is positive integer valued $\int \tau d P \geq 1$ and since $P(\xi_1 \geq C) = 1$, $\int \xi_1 d P \geq C > 0$. If $\int \xi_1 d P = \infty$, since $S_\tau(x) \geq \xi_1(x)$ for all $x \varepsilon Y$, $\int S_\tau d P = \infty$. If $\int \tau d P = \infty$, by Theorem 7.4, $\sum_{n=1}^{\infty} P(\{\tau \geq n\}) = \infty$. Now for each $n \varepsilon N$,

$$P(\{\tau \geq n\}) = P(\{\tau \geq n \text{ and } \xi_i \geq C, \text{ for } i=1,\ldots n\})$$
$$\leq P(\{S_\tau \geq nC\})$$
$$\leq P(\{S_\tau \geq [nC]\})$$

Therefore $\sum_{n=1}^{\infty} P(\{\tau \geq n\}) = \infty$ implies $\sum_{n=1}^{\infty} P(\{S_\tau \geq n\}) = \infty$.

Therefore by Corollary 7.3, $\int S_\tau \, d P = \infty$.

Thus Wald's identity holds if either $\int \xi_1 d P = \infty$ or $\int \tau \, d P = \infty$. Let us therefore assume that $\int \xi_1 d P < \infty$ and $\int \tau \, d P < \infty$. It follows easily from i) that $\int \xi_n d P = \int \xi_1 d P$ for all $n \varepsilon N$. In particular $\int \xi_n d P < \infty$ for all $n \varepsilon N$. Thus condition iii) implies that $\{\tau \geq n\}$ is orthogonal to ξ_n for all $n \varepsilon N$.

Let $\eta_n = \xi_n \cdot 1_{\{\tau \geq n\}}$, $n \varepsilon N$. Note that $S_\tau = \sum_{n=1}^{\infty} \eta_n$. We shall now check that $\{\eta_n\}$ satisfies the hypotheses of proposition 7.5.

If for some $x \varepsilon Y$ and $n_o \varepsilon N$, $\eta_{n_o}(x) = 0$ i.e. $\xi_{n_o} 1_{\{\tau \geq n_o\}}(x) = 0$, then $\tau(x) < n_o$ because ξ_n's are positive and hence $\tau(x) < n$ for all $n \geq n_o$. Consequently $\eta_n(x) = 0$ for all $n \geq n_o$.

Further since $\{0 < \eta_n < C\} \subseteq \{0 < \xi_n < C\}$ and $P(\{\xi_n \geq C\}) = 1$, $P(\{0 < \eta_n < C\}) = 0$ for all $n \varepsilon N$. We now have to check that $\lim_n P(\{S_\tau \geq n\}) = 0$.

Let $\varepsilon > 0$. Choose $n_o \varepsilon N$ such that $P(\{\tau \geq n_o\}) < \varepsilon/2$. Such a choice is available since we have assumed that $\int \tau \, dP < \infty$ and hence $\sum_{n=1}^{\infty} P(\{\tau \geq n\}) < \infty$. Then choose $n_1 \varepsilon N$ such that

$P(\{\xi_1 \geq n_1\}) < \frac{\varepsilon}{2\,n_o}$. Such a choice is possible since we have assumed that $\int \xi_1 d\, P < \infty$ and hence $\sum_{n=1}^{\infty} P(\{\xi_1 \geq n\}) < \infty$.

Therefore for any $n \geq n_o\, n_1$,

$$
\begin{aligned}
P(\{S_\tau \geq n\}) \leq P(\{S_\tau \geq n_o n_1\}) &\leq P(\{\tau \geq n_o\}) + P(\{S_\tau \geq n_o n_1 \text{ and } \tau < n_o\}) \\
&\leq P(\{\tau \geq n_o\}) + P\Big(\bigcup_{i=1}^{n_o-1} \{\xi_i \geq n_1\}\Big) \\
&\leq P(\{\tau \geq n_o\}) + \sum_{i=1}^{n_o-1} P(\{\xi_i \geq n_1\}) \\
&< P(\{\tau \geq n_o\}) + n_o .\, P(\{\xi_1 \geq n_1\}) \text{ by i)} \\
&< \varepsilon/2 + n_o .\, \frac{\varepsilon}{2\,n_o} = \varepsilon .
\end{aligned}
$$

Therefore $\lim_n P(\{S_\tau \geq n\}) = 0$.

Consequently proposition 7.5 applies to $\{\eta_n\}$ and we have

$$
\begin{aligned}
\int S_\tau d\, P &= \sum_{n=1}^{\infty} \int \eta_n d\, P \\
&= \sum_{n=1}^{\infty} (\int \xi_n d\, P) .\, P(\{\tau \geq n\}) \quad \text{since } \{\tau \geq n\} \text{ is orthogonal to } \xi_n . \\
&= (\int \xi_1 d\, P) \sum_{n=1}^{\infty} P(\{\tau \geq n\}) \\
&= (\int \xi_1 d\, P)(\int \tau\, d\, P) \quad \text{by Theorem 7.4.}
\end{aligned}
$$

The proof of the theorem is hence complete.

8. Positive Recurrence.

For $i, j \in I$, let m_{ij} stand for $\int t_j d\sigma[i]$ where t_j is the time of first occurrence of j.

Definition. A state i is called positive recurrent if $m_{ii} < \infty$.

Note that if $m_{ii} < \infty$ then by Theorem 7.4 $\sum_{n=1}^{\infty} \sigma[i](\{t_i \geq n\}) < \infty$. Consequently $\lim_n \sigma[i](\{t_i \geq n\}) = 0$ and hence $f^*_{ii} \geq \lim_n \sigma[i](\{t_i < n\}) = 1$. Therefore i is recurrent.

If a recurrent state is not positive recurrent, it is called null recurrent.

Lemma 8.1. Suppose i is a recurrent state, $i \neq j$ and $i \xrightarrow{W} j$. Let t_i, t_j be the times of first occurrence of i, j respectively. Then $\sigma[i](\{t_j < t_i\}) > 0$ and $\sigma[j](\{t_i < t_j\}) > 0$.

Proof. It is enough to show that $\sigma[i](\{t_j < t_i\}) > 0$ because i recurrent and $i \xrightarrow{W} j$ implies j recurrent and $j \xrightarrow{W} i$ and the roles of i and j can be interchanged.

If $\sigma[i](\{t_j < t_i\}) = 0$, then $\sigma[i](\{t_j > t_i\}) = 1$, i.e. $\sigma[i](\beta_1^{-1}(C)) = 1$ where C is the set of all those elements of F in which j does not occur, (β_1 and F are as in section 6). Therefore $\gamma(C) = 1$ (γ again as in section 6). The i.i.d strategy π defined by γ is a Markov strategy and C is a closed set. Therefore, by Theorem 4.1, $\pi(C^N) = 1$. Consequently by the Blocks Theorem $\sigma[i](\Phi^{-1}(C^N)) = 1$. Now $\Phi^{-1}(C^N) \subset A_j^c$ Therefore $\sigma[i](A_j^c) = 1$. This is a contradiction because actually $\sigma[i](A_j) = f_{ij}^* = 1$. Therefore the lemma is proved.

Let C be as in the proof of Lemma 8.1. Let θ be defined on G_i as follows :

$$\theta(h) = n \quad \text{if} \quad n \text{ is such that } \beta_1(h)\varepsilon C, \ldots, \beta_{n-1}(h)\varepsilon C, \beta_n(h)\varepsilon C^c$$
$$= \infty \quad \text{if no such } n \text{ exists.}$$

i.e. $\theta(h) = n$ in case j does not occur in the first $n-1$ i-blocks but occurs in the n^{th} i-block and $\theta(h) = \infty$ in case $h \varepsilon A_j^c$.

Define $\theta(h) = \infty$ on G_i^c.

Lemma 8.2. If i is recurrent, $i \neq j$ and $i \xrightarrow{W} j$, $\int \theta \, d\sigma[i] < \infty$.

Proof. Under the hypotheses, by lemma 8.1, $\sigma[i](t_j < t_i) > 0$, where t_i, t_j are the times of first occurrence of i, j respectively, that is $\sigma[i](\beta_1^{-1}(C^c)) > 0$ i.e. $\gamma(C^c) > 0$. Say $\gamma(C^c) = p$.

By theorem 7.3, $\int \theta \, d\, \sigma[i] = \sum_{n=1}^{\infty} \sigma[i] (\theta \geq n)$

$$= \sum_{n=1}^{\infty} \sigma[i] \left(\bigcap_{k=1}^{n-1} \beta_k^{-1}(C) \right)$$

$$= \sum_{n=1}^{\infty} (1-p)^{n-1} \quad \text{by the Blocks Theorem}$$

$$= \frac{1}{p} < \infty \quad \text{since } p > 0 .$$

Let $\{\lambda_n\}$ be the sequence of functions on H defined by $\lambda_1(h) = t(1)(h)$ and $\lambda_{n+1}(h) = t(n+1)(h) - t(n)(h)$, $n \in N$, $h \in G_i$ where $t(n)$, as in section 6, is the time of n^{th} occurrence of i and $\lambda_n(h) = \infty$, $h \in G_i^c$ and $n \in N$.

Let S_θ be defined by $S_\theta(h) = \sum_{n=1}^{\theta(h)} \lambda_n(h)$ for $h \in G_i$ and $S_\theta(h) = \infty$ if $h \notin G_i$.

<u>Lemma 8.3</u>. If i is recurrent, $i \neq j$ and $i \xrightarrow{W} j$, we have $\int S_\theta \, d\sigma[i] = m_{ij} + m_{ji}$.

<u>Proof</u>. Let t_j be the time of first occurrence of j. It is easy to see that $t_j(h) \leq S_\theta(h)$ for all $h \in H$.

Define on H the function r by $r(h) = S_\theta(h) - t_j(h)$ if $t_j(h) < \infty$

$= \infty$ otherwise

Clearly r is a non-negative integer valued function. Further $S_\theta(h) = t_j(h) + r(h)$ for all h.

Therefore $\int S_\theta d\,\sigma[i] = \int t_j d\,\sigma[i] + \int r\,d\,\sigma[i]$.

Since by definition, $\int t_j d\,\sigma[i] = m_{ij}$, it is enough to show that $\int r\,d\,\sigma[i] = m_{ji}$

Now $\int r\,d\,\sigma[i] = \sum_{n=1}^{\infty} \sigma[i](\{r \geq n\})$ by Theorem 7.4

$$= \sum_{n=1}^{\infty} \sigma[i](\{t_j < \infty \text{ and } r \geq n\}) \text{ since } f^*_{ij} = 1,$$

$$= \sum_{n=1}^{\infty} \sigma[i](\{t_j < \infty \text{ and } S_\theta \geq t_j + n\})$$

The event A_n (say) in the n^{th} term of the summation is conditionally determined given t_j and $A_n\, p_{t_j}(h) = \{t_i \geq n\}$, $h \in \{t_j < \infty\}$ and $n \in N$.

Therefore by the Strong Markov Property,

$$\int r\,d\,\sigma[i] = \sum_{n=1}^{\infty} \sigma[j](\{t_i \geq n\})\, \sigma[i](\{t_j < \infty\})$$

$$= \sum_{n=1}^{\infty} \sigma[j](\{t_i \geq n\}) = m_{ji} \quad \text{by Theorem 7.4}.$$

<u>Lemma 8.4</u>. Let i be recurrent and $i \xrightarrow{w} j$. Then

$$m_{ii} \leq m_{ij} + m_{ji} \quad \text{and} \quad m_{jj} \leq m_{ij} + m_{ji}.$$

<u>Proof</u>. Once again it is enough to prove the first one of the two inequalities because the other follows by interchanging the roles of i and j.

By Lemma 8.3, $\int S_\Theta d\, \sigma[i] = m_{ij} + m_{ji}$.

But $t_i(h) \leq S_\Theta(h)$ for $h \in H$, where t_i is the time of first occurrence of i.

Therefore $m_{ii} = \int t_i \, d\, \sigma[i] \leq \int S_\Theta d\, \sigma[i] = m_{ij} + m_{ji}$.

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 8.5. 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. If $i=j$, the result is trivial. So assume that $i \neq j$.

In view of the remark after lemma 8.4, it is enough to prove that $m_{ji} + m_{ji} < \infty$. By Lemma 8.3, this is equivalent to proving that $\int S_\Theta d\, \sigma[i] < \infty$. We shall prove this using Wald's Identity.

We will check that $\{\lambda_n\}$ satisfies the hypotheses of Wald's Identity (Theorem 7.5).

Clearly λ_n's are positive - integer valued measurable functions on H. Consider the map λ defined on F into N

defined by $\lambda(a)$ = the number of terms in a, $a \in F$. Observe that for each $n \in N$, $\lambda_n(h) = \lambda \circ \beta_n(h)$ for $h \in G_i$, where $\{\beta_n\}$ is as defined in section 6.

Now for a Borel subset A of the extended real line,

$$\sigma[i](\lambda_n^{-1}(A)) = \sigma[i](\beta_n^{-1}(\lambda^{-1}(A)))$$
$$= \sigma[i](\beta_1^{-1}(\lambda^{-1}(A)))$$
$$= \gamma(\lambda_1^{-1}(A)) \text{ by the Blocks theorem, for all } n \in N.$$

Since λ_n'S are integer-valued, condition ii) of Wald's Identity is automatically satisfied with say $C = \frac{1}{2}$.

We finally need to check that $\{\theta < n\}$ is orthogonal to λ_n, for each $n \in N$. Fix $n \in N$.

By theorem 7.4, $\int \lambda_n . 1_{\{\theta < n\}} d\,\sigma[i]$

$$= \sum_{k=1}^{\infty} \sigma[i](\lambda_n 1_{\{\theta < n\}} \geq k)$$

$$= \sum_{k=1}^{\infty} \sigma[i](\{\lambda_n \geq k\} \cap \{\theta < n\})$$

Since $(\theta < n)$ is an event depending only on the first n-1 blocks and since $(\lambda_n \geq k)$ is an event depending on the n^{th} block, by the Blocks Theorem, $\sigma[i](\{\lambda_n \geq k\} \cap \{\theta < n\})$

$$= \sigma[i](\lambda_n \geq k) . \sigma[i](\{\theta < n\}).$$

Therefore $\int \lambda_n 1_{\{\theta<n\}} d\sigma[i] = \sum_{k=1}^{\infty} \sigma[i](\lambda_n \geq k)\, \sigma[i](\{\theta<n\})$

$$= \sigma[i](\{\theta<n\}) \int \lambda_n d\sigma[i], \text{ by}$$

Theorem 7.4.

Therefore $\{\theta < n\}$ is orthogonal to λ_n. Therefore by Wald's Identity,

$$\int S_\theta d\sigma[i] = (\int \lambda_1 \, d\sigma[i])(\int \theta \, d\sigma[i])$$

Now $\int \theta \, d\sigma[i] < \infty$ by Lemma 8.2 and $\int \lambda_1 \, d\sigma[i] = m_{ii} < \infty$ because i is positive recurrent. Hence $m_{ij} + m_{ji} = \int S_\theta d\sigma[i] < \infty$. The proof of the theorem is therefore complete.

<u>Corollary 8.6</u>. Let i be positive recurrent and $i \xrightarrow{w} j$. Then $i \xrightarrow{s} j$. Consequently a positive recurrent weakly communicating class is strongly communicating.

<u>Proof</u>. If i is positive recurrent and $i \xrightarrow{w} j$, by Theorem 8.5, $m_{ij} < \infty$ and hence as before it can be shown that

$\lim_n \sigma[i](t_j \geq n) = 0$. Therefore $\lim_n \sigma[i](t_j < n) = 1$. Hence for some n, $\sigma[i](A_j^n) > 0$. Consequently $i \xrightarrow{s} j$. The rest of the assertion of the corollary now follows easily, once again using Theorem 8.5.

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 $\sigma(1) = \sum_{n=2}^{\infty} p_n \delta_n + (1 - \sum_{n=2}^{\infty} p_n)\gamma$, where $0 < p_n < 1$ for all $n \geq 2$ and $\sum_{n=2}^{\infty} p_n < 1$ and γ is a diffuse measure, $\sigma(n+1) = \delta_1$, $n \in N$.

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

9. Limits of n-step transition probabilities and average number of visits to a state.

For $i, j \in I$, $\sigma[i](A_j^n)$ will be called the probability of n-step transition to j from i and $\frac{1}{n}\sum_{k=1}^{n} \sigma[i](A_j^k)$ will be called the expected proportion of visits to j upto time n from i. In this section, we study the limits of these quantities as $n \longrightarrow \infty$.

Theorem 9.1. If j is transient, $\lim_{n \longrightarrow \infty} \sigma[i](A_j^n) = 0$ for all $i \in I$.

Proof. Let $Y(j)(h) = \sum_{n=1}^{\infty} \delta_j(h_n)$, $h \in H$, where $\delta_j(i) = 1$ if $i = j$, and 0 if $i \neq j$.

Then $\int Y(j)(h)\, d\,\sigma[i](h) = \sum_{m=1}^{\infty} \sigma[i](\{Y(j) \geq m\})$

$$= \sum_{m=1}^{\infty} f_{ij}^{*} (f_{jj}^{*})^{m-1} \quad , \quad \text{by } (5.2)$$

$$= \frac{f_{ij}^{*}}{1-f_{jj}^{*}} < \infty, \quad \text{since } j \text{ is transient.}$$

Further $\int Y(j)(h)\, d\,\sigma[i](h) = \int (\sum_{n=1}^{\infty} \delta_j(h_n))\, d\,\sigma[i](h)$

$$\geq \sum_{n=1}^{\infty} \int \delta_j(h_n)\, d\,\sigma[i](h)$$

$$= \sum_{n=1}^{\infty} \sigma[i](A_j^n)$$

Therefore $\sum_{n=1}^{\infty} \sigma[i](A_j^n) < \infty$. Consequently $\lim_{n \to \infty} \sigma[i](A_j^n) = 0$.

Corollary 9.2. If j is transient, $\frac{1}{n} \sum_{k=1}^{n} \sigma[i](A_j^k) \longrightarrow 0$ as $n \longrightarrow \infty$.

Proof. Immediate from theorem 9.1.

For the next theorem we shall be needing a version of the Renewal theorem which can be found in [18].

Renewal Theorem. Let f_n, $n \geq 1$, be a sequence of non-negative real numbers such that $\sum_{n=1}^{\infty} f_n \leq 1$.

Suppose $\{u_n\}$ is a sequence of real numbers satisfying,

$$u_n = f_1 u_{n-1} + f_2 u_{n-2} + \dots + f_n u_o \quad \text{for all } n \in N,$$

where $u_o = 1$. Then

a) if $\sum_{n=1}^{\infty} f_n = 1$, we have

$u_{nd} \longrightarrow \frac{d}{\mu^{-1}}$ as $n \longrightarrow \infty$, where d is the greatest common divisor (g.c.d) of those n for which $f_n > 0$ and

$\mu = \sum_{n=1}^{\infty} n f_n$, and

b) if $\sum_{n=1}^{\infty} f_n < 1$, then $u_n \longrightarrow 0$ as $n \longrightarrow \infty$.

(<u>Remark</u>. Although this version has not been explicitly stated in the above form in [18], it follows immediately from theorem 1 and theorem 2 of chapter XIII. 10).

<u>Theorem 9.3</u>. Let j be a recurrent state. If period of j is defined and is d, then $\sigma[j](A_j^{nd}) \longrightarrow \frac{d}{m_{jj}}$ as $n \longrightarrow \infty$, and if not defined, $\sigma[j](A_j^n) \longrightarrow 0$.

<u>Proof</u>. If period of j is not defined, $\sigma[j](A_j^n) = 0$ for all $n \in N$ and hence the result is trivial. Suppose the period of j is defined and is d. Define $f_n = \sigma[j](t_j = n)$ where t_j is the time of first occurrence of j and define

$u_n = \sigma[j](A_j^n)$, $n \in N$.

Clearly $u_n = \sigma[j](A_j^n) = \sum_{k=1}^{n} \sigma[j](\{t_j=k\} \cap A_j^n)$.

Conditioning the k^{th} term at the k^{th} coordinate, we get

$$u_n = \sum_{k=1}^{n} \sigma[j](\{t_j=k\}).\sigma[j](A_j^{n-k}), \text{ where } \sigma[j](A_j^o) = 1.$$

$$= \sum_{k=1}^{n} f_k\, u_{n-k}. \qquad \ldots.(9.1)$$

If $\sum_{n=1}^{\infty} f_n = 1$, then $\sum_{n=1}^{\infty} \sigma[j](t_j = n) = 1$,

therefore $\sigma[j](t_j \geq n) \longrightarrow 0$ as $n \longrightarrow \infty$.

Consequently, $m_{jj} = \sum_{n=1}^{\infty} n.\sigma[j](t_j = n) = \sum_{n=1}^{\infty} n\, f_n$.

If $\sum_{n=1}^{\infty} f_n < 1$, then $\sigma[j](t_j \geq n) \not\longrightarrow 0$ as $n \longrightarrow \infty$,

so $m_{jj} = \infty$.

By (9.1), the renewal theorem applies and we have

$$u_{nd} = \sigma[j](A_j^{nd}) \longrightarrow \frac{d}{m_{jj}} \quad \text{if} \quad \sum_{n=1}^{\infty} f_n = 1,$$

and $u_n = \sigma[j](A_j^n) \longrightarrow 0 = \dfrac{d}{m_{jj}}$ if $\sum_{n=1}^{\infty} f_n < 1$

[(we of course used the additional fact that

$$\text{g.c.d}\, \{ n : \sigma[j](A_j^n) > 0\} = \text{g.c.d}\, \{n: \sigma[j](t_j = n) > 0\}).$$

A proof of this can be found in [8] and [20] .]

The proof of the theorem is now complete.

Theorem 9.4. Suppose j is recurrent. a) If $m_{jj} = \infty$, then $\lim_{n \to \infty} \sigma[i](A_j^n) = 0$ for all $i \in I$. b) If $m_{jj} < \infty$ and period of j is equal to d, then

$$\lim_{n \to \infty} \sigma[i](A_j^{nd+r}) = \frac{d}{m_{jj}} \sum_{m=0}^{\infty} \sigma[i](\{t_j = md+r\})$$

where t_j is the time of first occurrence of j.

Proof. For each $n \in N$,

$$\sigma[i](A_j^n) = \sum_{k=1}^{n} \sigma[i](\{t_j = k\} \cap A_j^n)$$

$$= \sum_{k=1}^{n} \sigma[i](\{t_j = k\}).\sigma[j](A_j^{n-k}) \quad \ldots.(9.2)$$

by conditioning the k^{th} term at the k^{th} coordinate.

By applying the dominated convergence theorem for summation to (9.2) and applying theorem 9.3, we get both a) and b).

Corollary 9.5. If j is recurrent then

a) $\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \sigma[i](A_j^k) = 0$ if $m_{jj} = \infty$,

b) $\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \sigma[j](A_j^k) = \frac{1}{m_{jj}}$ if $m_{jj} < \infty$,

c) $\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \sigma[i](A_j^k) = \frac{\sum_{n=1}^{\infty} \sigma[i](t_j = n)}{m_{jj}}$ if $m_{jj} < \infty$ and $i \neq j$,

where t_j is the time of first occurrence of j.

Proof. a) follows immediately from part a) of Theorem 9.4.

b) follows from theorem 9.3 and c) from part b) of Theorem 9.4 if we use, in addition, the following elementary result from real analysis :

Let $\{a_n\}$ be a sequence of real numbers such that for some $d \in N$, $a_{nd+r} \longrightarrow d_r$ as $n \longrightarrow \infty$ for $r = 0,1,\ldots,d-1$. Then

$$\frac{1}{n} \sum_{k=1}^{n} a_k \longrightarrow \frac{\sum_{r=0}^{d-1} d_r}{d} \quad \text{as } n \longrightarrow \infty .$$

10. The Strong Law of Large numbers (SLLN).

We now prove a Strong Law of large numbers for Markov chains. For that we need the following version of the Strong law of large numbers for i.i.d strategies.

Let E be an arbitrary set. Recall that ρ is an i.i.d strategy on E if $\rho(p) = \rho(q)$ for all $p, q \in E^*$, the set of all finite sequences of elements in E. Let $Q = E^N$ be equipped with the product of discrete topologies and $\underline{\underline{E}}$ be the Borel σ-field on Q. A sequence of random variables $\{Z_n\}$ defined on Q is called a sequence of identical coordinate mappings if there exists a function f on E such that for each $w \in Q$ and $n \in N$, $Z_n(w) = f(w_n)$, where w_n is the n^{th} coordinate of w.

<u>SLLN for i.i.d strategies.</u> Let ρ be an i.i.d strategy and $\{Z_n\}$ a sequence of identical coordinate mappings on Q. If $\int Z_n d\rho = \mu$ for all $n \varepsilon N$ where μ is a real number, then $\rho(\{w : \frac{S_n(w)}{n} \longrightarrow \mu\}) = 1$

$$\text{where} \quad S_n(w) = \sum_{k=1}^{n} Z_k(w), \quad w \varepsilon Q.$$

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 $n \varepsilon N$, then $\rho(\{w : \frac{S_n(w)}{n} \longrightarrow \infty\})=1$. Consequently if $\int Z_n d\rho$ exists and is equal to $-\infty$ for all $n \varepsilon N$, then

$$\rho(\{w : \frac{S_n(w)}{n} \longrightarrow -\infty\}) = 1.$$

(For a proof of this see [6] and [7]).

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 $n \varepsilon N$, let $l(n)$ be the random variable defuned by $l(n)(h) = \sum_{k=1}^{n} \delta_i(h_k)$ where $\delta_i(j) = 1$ if $i=j$ and 0 if $i \neq j$.

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

<u>Remark</u>. 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 $\mu_i(f)$ may be defined even when $\mu_i(f^+) = \mu_i(f^-) = \infty$).

<u>In what follows i is assumed to be a recurrent state</u>.

<u>Lemma 10.1</u> If either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$, then

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

<u>Proof</u>. We shall prove the result for a general non-negative f. In that case we shall have the required result for f^+ and f^- and that will 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 $Z(i_1, \ldots, i_n) = \sum_{k=1}^{n} f(i_k)$.

Now define a sequence $\{Z_n\}$ of identical coordinate mappings on $(\Omega, \underline{\underline{F}})$ by $Z_n(w) = Z(w_n)$, $w \in \Omega$, $n \in N$, where w_n is the n^{th} coordinate of w. (Ω, $\underline{\underline{F}}$, Φ and π are as defined in section 6). Since π is an i.i.d strategy, by SLLN for i.i.d strategies, $\pi(B) = 1$ where $B = \left\{w : \frac{\sum_{k=1}^{n} Z_k(w)}{n} \longrightarrow \mu_i(f)\right\}$.

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

$$\frac{\sum_{k=1}^{l(n)(h)} Z_k(\Phi(h))}{l(n)(h)} \leq \frac{\sum_{k=1}^{n} f(h_k)}{l(n)(h)} \leq \frac{\sum_{k=1}^{l(n)(h)+1} Z_k(\Phi(h))}{l(n)(h)}$$

for all $h \in G_i$, and large enough n.

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

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

and the proof of the lemma is complete.

Corollary 10.2. 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 $0 < |\mu_i(g)| < \infty$,

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

Proof. In view of assumptions i) and ii), lemma 10.1 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 10.3. 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

$$\sigma[i](\{h: \frac{S_n(h)}{n} \longrightarrow \frac{\mu_i(f)}{m_{ii}}\}) = 1, \text{ where } S_n(h)=f(h_1)+\ldots+f(h_n).$$

Proof. Let g be the function defined by $g(j) = 1$ for all $j \in I$. Now f satisfies condition i) of corollary 10.2 by assumption and g satisfies ii) because g is non-negative. 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 10.2, the result follows.

Proposition 10.4. 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

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

Proof. Observe that if $A = \{h: \frac{S_n}{n} \longrightarrow \frac{\mu_i(f)}{m_{ii}}\}$, then $A \cap \{t < \infty\}$ where t is the time of first occurrence of 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,

$$\sigma[i](A) = \sigma[i](A \cap \{t < \infty\}) = \sigma[j](A).\sigma[i](\{t < \infty\})$$
$$= \sigma[j](A).$$

The result now follows from corollary 10.3.

<u>Corollary 10.5</u>. 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

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

<u>Proof</u>. Since i is positive recurrent and $i \xrightarrow{W} j$, j is also positive recurrent. Therefore by corollary 10.3 applied to j, $\frac{S_n}{n}$ converges to $\frac{\mu_j(f)}{m_{jj}}$ on a set of $\sigma[j]$-measure one. Also by proposition 10.4, $\frac{S_n}{n}$ converges to $\frac{\mu_i(f)}{m_{ii}}$ on a set of $\sigma[j]$ -measure one. Therefore the result follows.

<u>Corollary 10.6</u>. If i is positive recurrent and $i \xrightarrow{W} j$, for each subset E of the state space,

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

<u>Proof</u>. Immediate from corollary 10.5 by taking $f = 1_E$.

<u>Theorem 10.7</u> (SLLN). Let I be a recurrent class under a Markov strategy σ. Fix $i \in I$. Let f be a function on I such that either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$. Assume that either I is a positive recurrent class or $0 < |\mu_i(f)| < \infty$.

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

Proof. By proposition 10.4, $\sigma[j](\{h: \frac{S_n}{n} \longrightarrow \frac{\mu_i(f)}{m_{ii}}\}) = 1$ for all $j \in I$. Further if σ_o is the initial distribution of σ,

$$\sigma(\{h: \frac{S_n}{n} \longrightarrow \frac{\mu_i(f)}{m_{ii}}\}) = \int \sigma[j](\{h: \frac{S_n}{n} \longrightarrow \frac{\mu_i(f)}{m_{ii}}\}) d\,\sigma_o(j)$$

by I of section 1. Therefore the assertion of the theorem follows immediately.

We shall later on, in section 13, identify $\frac{\mu_i(f)}{m_{ii}}$ as the the integral of f with respect to a special stationary initial distribution, in case the chain is positive recurrent.

Many important limit theorems for Markov chains follow as special cases of theorem 10.7. We mention below only two of them.

Corollary 10.8. Let I be a recurrent class. Let $i \in I$ and suppose $l(n)$ is defined as in the beginning of this section.

Then

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

Proof. Notice that the function f defined by $f(j) = \delta_i(j)$ for all $j \in I$ satisfies the hypotheses of Theorem 10.7 and $\mu_i(f) = 1$. Further for this f, $S_n = l(n)$. Consequently the result follows.

<u>Proposition 10.9</u>. Let I be a positive recurrent class and let $i \varepsilon I$. Then $\sigma(\{h: \frac{t(l(n))}{n} \longrightarrow 1\}) = 1$, where for each $k \varepsilon N$, $t(k)$ is the time of k^{th} occurrence of i.

<u>Proof</u>. In view of corollary 10.8, it is enough to show that

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

This will follow if we show that $\sigma[j](\{h: \frac{t(l(n))}{l(n)} \longrightarrow m_{ii}\}) = 1$ for each $j \varepsilon I$.

We first show that $\sigma[i](\{\frac{t(l(n))}{l(n)} \longrightarrow m_{ii}\}) = 1$. This follows from the Blocks theorem and SLLN for i.i.d strategies applied to the block-length functions $\lambda_n = \lambda \circ \beta_n$, $n \varepsilon N$, where λ_n's , λ and β_n's are as defined in section 8.

Now an argument similar to the one used in the proof of proposition 10.4 shows that

$$\sigma[i](\{\frac{t(l(n))}{l(n)} \longrightarrow m_{ii}\} \cap G_i) = \sigma[j](\{\frac{t(l(n))}{l(n)} \longrightarrow m_{ii}\} \cap G_i)$$

for all $j \varepsilon I$.

Therefore $\sigma[j](\{h: \frac{t(l(n))}{l(n)} \longrightarrow m_{ii}\}) = 1$ for all $j \varepsilon I$.

The proposition follows immediately.

11. <u>A Mean Ergodic Theorem</u>. In the last section we proved the almost sure convergence of $\frac{S_n}{n}$. In this section we shall prove the convergence of the means of $\frac{S_n}{n}$ i.e. of $\int \frac{S_n}{n} \, d\,\sigma[i]$,

Let i be recurrent. Let $\mu_i(f)$ and $\{Z_n\}$ be defined as in the previous section.

<u>Lemma 11.1</u>. $\int Z_n(\Phi(h)) \, d\,\sigma[i]\,(h) = \mu_i(f)$ for all $n \in N$, provided $\mu_i(f)$ exists.

<u>Remark</u>. Φ has only been defined on G_i, hence we cannot talk about $Z_n(\Phi(h))$ for $h \in G_i^c$. However $\sigma[i](G_i) = 1$ and we are interested only in the integral of $Z_n \circ \Phi$ w.r.t $\sigma[i]$. So it does not matter how Φ is defined outside G_i. To fix matters, here and in future, whenever we talk about the integral of a partially defined function, we shall interpret the function to be zero outside its domain.

<u>Proof</u>. By the change of variable theorem and the blocks theorem, $\int Z_n(\Phi(h)) \, d\,\sigma[i]\,(h) = \int Z(w_n) \, d\,\pi(w)$ where Z, π etc. are as in section 10. Since π is i.i.d, $\int Z(w_n) \, d\,\pi\,(w) = \int Z(w_1) \, d\,\pi\,(w)$ (This can be verified the usual way by first considering indicator functions, then simple functions and so on). It is easy to see from definitions that $\int Z(w_1) \, d\,\pi\,(w) = \int Z_1(\Phi(h)) \, d\,\sigma[i]\,(h) = \mu_i(f)$.
This completes the proof of the lemma.

<u>Lemma 11.2</u>. For each $n \in N$, and $m \leq n+1$,

$$\int_{\{l(n)\geq m-1\}} Z_m(\Phi(h))\, d\,\sigma[i] = \left(\int Z_m(\Phi(h))d\,\sigma[i]\right)\cdot\sigma[i](\{l(n)\geq m-1\})$$

provided the integrals exist.

<u>Proof</u>. Observe that $\{l(n) \geq m-1\}$ is the event that the sum of the first $(m-1)$ i-block lengths is less than or equal to n. Therefore it depends only on the first $(m-1)$ i-blocks. However Z_m depends only on the m^{th} block. We now prove the result first for the case when the integrand is an indicator function depending only on the m^{th} block, using the Blocks theorem, then for the case when the integrand is a simple function depending on the m^{th} block, then for a bounded measurable function, then a non-negative function and finally we obtain the result in general.

<u>Lemma 11.3</u>. Assume that f is non-negative and $\mu_i(f) < \infty$. Then

$$\frac{1}{n}\int\left\{\max_{1\leq m\leq n} Z_m(\Phi(h))\right\} d\,\sigma[i](h) \longrightarrow 0 \quad \text{as} \quad n\longrightarrow\infty.$$

<u>Proof</u>. Let $S_n^*(h) = \max_{1\leq m\leq n} Z_m(\Phi(h))$, $h \in G_i$, $n \in N$.

By the Blocks theorem, for $k, n \in N$

$$\sigma[i](\{S_n^* \geq k\}) = 1 - [\theta(k)]^n \leq n[1-\theta(k)] \quad \ldots(11.1)$$

where $\theta(k) = \sigma[i](\{h: Z_1(\Phi(h)) < k\})$.

Since $\int Z_1(\Phi(h))\, d\,\sigma[i](h) = \mu_i(f) < \infty$, it follows from the remark made after proposition 7.2, that $\sum_{k=1}^{\infty}(1-\theta(k)) < \infty$ and hence $\theta(k) \longrightarrow 1$ as $k \longrightarrow \infty$.

Consequently $\sigma[i](\{S_n^* \geq k\}) \longrightarrow 0$ as $k \longrightarrow \infty$.

Therefore by proposition 7.2, $\int S_n^*\, d\sigma[i] = \operatorname{Sup}_k \int_{\{S_n^*<k\}} S_n^* d\sigma[i]$.

From the proof of proposition 7.1, it is clear that

$\operatorname{Sup}_k \int_{\{S_n^*<k\}} S_n^*\, d\sigma[i]$ lies between $\sum_{k=1}^{\infty} \sigma[i](\{S_n^* \geq k\})$ and $\sum_{k=1}^{\infty} \sigma[i](\{S_n^* \geq k\}) + 1$.

Therefore in order to prove our lemma, it suffices to show that

$$\frac{1}{n} \sum_{k=1}^{\infty} \sigma[i](\{S_n^* \geq k\}) \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

For each n, by (11.1), $\sum_{k=1}^{\infty} \frac{1}{n} \sigma[i](\{S_n^* \geq k\}) \leq \sum_{k=1}^{\infty} [1-\theta(k)] < \infty$.

Therefore by the dominated convergence theorem,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{\infty} \sigma[i](\{S_n^* \geq k\}) = \sum_{k=1}^{\infty} \lim_{n \to \infty} \frac{1}{n} \sigma[i](\{S_n^* \geq k\}).$$

Therefore it is enough to show for each k, that

$$\lim_{n \to \infty} \frac{1}{n} \sigma[i](S_n^* \geq k) = 0.$$

But $\frac{1}{n}\sigma[i](S_n^* \geq k) \leq \frac{1}{n}[1-[\theta(k)]^n] \longrightarrow 0$ as $n \longrightarrow \infty$.

Therefore the lemma follows.

Lemma 11.4. If $f \geq \theta$ and $\mu_i(f) < \infty$, then

$$\lim_{n \longrightarrow \infty} \int \left[\frac{1}{n} \sum_{k=n+1}^{t(l(n)(h)+1)} f(h_k)\right] d\,\sigma[i](h) \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Proof. Clearly,

$$0 \leq \frac{1}{n} \sum_{k=n+1}^{t(l(n)(h)+1)} f(h_k) \leq \frac{n+1}{n} \cdot \frac{1}{n+1} \max_{1 \leq m \leq n+1} Z_m(\Phi(h))$$

Hence the result follows from lemma 11.3.

Theorem 11.5. Let i be recurrent and f be such that $\mu_i(f^+) < \infty$ and $\mu_i(f^-) < \infty$. Then

$$\lim_{n \longrightarrow \infty} \int \left(\frac{S_n}{n}\right) d\,\sigma[i] = \frac{\mu_i(f)}{m_{ii}},$$

where $S_n(h) = \sum_{i=1}^{n} f(h_i)$.

Proof. Since under the hypothesis it is enough to prove the result for f^+ and f^-, we assume that f is non-negative.

For $h \in G_i$, and $n \in N$,

$$\frac{S_n}{n} = \frac{\sum_{m=1}^{l(n)+1} Z_m(\Phi(h))}{n} - \frac{1}{n} \sum_{k=n+1}^{t(l(n)(h)+1)} f(h_k)$$

$$= \frac{\sum_{m=1}^{n+1} Z_m(\Phi(h)) . 1_{\{l(n) \geq m-1\}}}{n} - \frac{1}{n} \sum_{k=n+1}^{t(l(n)(h)+1)} f(h_k).$$

Now applying lemmas 11.1, 11.2 and lemma 11.4, we get :

$$\lim_{n \to \infty} \int \frac{S_n}{n} \, d\sigma[i] = \mu_i(f). \lim_n \frac{1}{n} \sum_{m=1}^{n+1} \sigma[i](l(n) \geq m-1)$$

$$= \mu_i(f). \lim_n \frac{1}{n} \int l(n) \, d\sigma[i], \text{ by theorem 7.4.}$$

It now suffices to show that $\lim_n \frac{1}{n} \int l(n) \, d\sigma[i] = \frac{1}{m_{ii}}$.

Now $\lim_n \frac{1}{n} \int l(n) \, d\sigma[i] = \lim_n \frac{1}{n} \int \left[\sum_{k=1}^{n} \delta_i(h_k) \right] d\sigma[i]$

$$= \lim_n \frac{1}{n} \sum_{k=1}^{n} \sigma[i](A_i^k)$$

$$= \frac{1}{m_{ii}} \quad \text{by corollary 9.5.}$$

This completes the proof of the theorem.

12. <u>A Ratio limit theorem</u>. We now prove a ratio limit theorem for more general initial distributions. This will be used later

while studying stationary initial distributions. Let σ be a Markov strategy.

Assume that I is a recurrent class and let $i, j \in I$, $i \neq j$. Since I is recurrent, $\sigma[1](G_i) = 1$ for all $1 \in I$. Therefore $\sigma(G_i) = 1$.

On G_i, define $\{Z_n\}$ as in section 11, corresponding to the function $f = \delta_j$.

<u>Lemma 12.1</u>. $\int Z_1(\Phi(h))\, d\,\sigma[j](h) < \infty$.

<u>Proof</u>. $\int Z_1(\Phi(h))\, d\,\sigma[j](h) = \sum_{m=1}^{\infty} \sigma[j](\{h: Z_1(\Phi(h)) \geq m\})$.

The m^{th} term, by repeated application of the Strong Markov property for the incomplete stop rule t' defined by

$$t' = t \quad \text{if} \quad t < t(1),$$
$$= \infty \quad \text{otherwise},$$

where $t(1)$, t are respectively the incomplete rules corresponding to the first occurrences of i, j respectively, reduces to

$$\sigma[j](\{h: Z_1(\Phi(h)) \geq m\}) = [\sigma[j](\{t < t(1)\})]^m$$

But $\sigma[j](\{t < t(1)\}) = 1 - \sigma[j](\{t(1) < t\}) < 1$, by Lemma 8.1.

Therefore $\int Z_1 \circ \Phi \; d\,\sigma[j] = \dfrac{\sigma[j](\{t < t(1)\})}{\sigma[j](\{t(1) < t\})} < \infty.$

Lemma 12.2. $\int Z_1 \circ \Phi \; d\,\sigma \leq \int Z_1 \circ \Phi \; d\,\sigma[j] + 1$.

Consequently, $\int Z_1 \circ \Phi \; d\,\sigma < \infty$.

Proof. $\int Z_1 \circ \Phi \; d\,\sigma = \sum_{m=1}^{\infty} \sigma(\{Z_1 \circ \Phi \geq m\})$

$$= \sum_{m=1}^{\infty} \sigma(\{t < t(1)\}).(\sigma[j](t < t(1)))^{m-1}$$

by repeated application of the Strong Markov property for the incomplete stop rule t' defined in the previous lemma.

Therefore $\int Z_1 \circ \Phi \; d\,\sigma = \sigma(\{t < t(1)\}). \sum_{m=1}^{\infty} [\sigma[j](\{t < t(1)\})]^{m-1}$

$$\leq \sum_{m=0}^{\infty} [\,\sigma[j](\{t < t(1)\})\,]^m$$

$$= \int Z_1 \circ \Phi \; d\,\sigma[j] + 1$$

Thus the lemma is proved.

Lemma 12.3. For $n \geq 2$, $\int Z_n \circ \Phi \; d\,\sigma = \int Z_1 \circ \Phi \; d\,\sigma[i] = \mu_i(\delta_j)$

Proof. We have $\int Z_1 \circ \Phi \; d\,\sigma[i] = \mu_i(\delta_j)$ directly from definitions.

For $n \geq 2$, $\int Z_n \circ \Phi \; d\sigma = \sum_{m=1}^{\infty} \sigma(\{ Z_n \circ \Phi \geq m \})$

$$= \sum_{m=1}^{\infty} \sigma(\{ t(n-1) < \infty \}) . \sigma[i](\{ Z_1 \circ \Phi \geq m \})$$

by the strong Markov property applied to $t(n-1)$, the time of the $(n-1)^{st}$ occurrence of i. Since I is recurrent, $\sigma(\{ t(n-1) < \infty \}) = 1$.

Therefore $\int Z_n \circ \Phi \; d\sigma = \sum_{m=1}^{\infty} \sigma[i](\{ Z_1 \circ \Phi \geq m \}) = \int Z_1 \circ \Phi \; d\sigma[i]$.

Lemma 12.4. For each $n \in N$ and $m \geq 2$,

$$\int_{\{l(n) \geq m-1\}} Z_m \circ \Phi \; d\sigma = \sigma(\{ l(n) \geq m-1 \}) \, \mu_i(\delta_j) .$$

Proof. $\int_{\{l(n) \geq m-1\}} Z_m \circ \Phi \; d\sigma = \sum_{k=1}^{\infty} \sigma(\{ Z_m \circ \Phi \geq k \text{ and } l(n) \geq m-1 \})$

$$= \sum_{k=1}^{\infty} \sigma(\{ l(n) \geq m-1 \}) . \sigma[i](\{ Z_1 \circ \Phi \geq k \})$$

by the Strong Markov property applied to $t'' = t(m-1)$ if $t(m-1) \leq n$,
$= \infty$ otherwise

$$= \sigma(\{ l(n) \geq m-1 \}) . \int Z_1 \circ \Phi \; d\sigma[i]$$

$$= \sigma(\{ l(n) \geq m-1 \}) . \mu_i(\delta_j).$$

Lemma 12.5. $\int \left\{ \sum_{k=n+1}^{t(l(n)(h)+1)} \delta_j(h_k) \right\} d\,\sigma(h) \leq \int Z_1 \circ \Phi \, d\sigma[j] + 1.$

Proof. $$\int \left\{ \sum_{k=n+1}^{t(l(n)(h)+1)} \delta_j(h_k) \right\} d\sigma(h) = \sum_{m=1}^{\infty} \sigma\left(\left\{ \sum_{k=n+1}^{t(l(n)+1)} \delta_j(h_k) \geq m \right\} \right)$$

$$= \sum_{m=1}^{\infty} \sigma(\{t''' < \infty\}) . \sigma[j](Z_1 \circ \Phi \geq m-1)$$

where $t'''(h) = t(h_{n+1}, h_{n+2}, \ldots) + n$ if $t(h_{n+1}, h_{n+2} \ldots) + n < t(l(n)+1)$

$= \infty$ otherwise,

$$\leq \sum_{m=1}^{\infty} \sigma[j](Z_1 \circ \Phi \geq m-1)$$

$$= \int Z_1 \circ \Phi \, d\sigma[j] + 1 .$$

Theorem 12.6. Let I be a recurrent class under σ. Let $i, j \in I$. Assume that $\int \sum_{k=1}^{n} \delta_i(h_k) \, d\sigma(h) \longrightarrow \infty$ as $n \longrightarrow \infty$.

Then

$$\lim_{n \to \infty} \frac{\int \sum_{k=1}^{n} \delta_j(h_k) \, d\,\sigma(h)}{\int \sum_{k=1}^{n} \delta_i(h_k) \, d\,\sigma(h)} = \mu_i(\delta_j) .$$

Proof. Observe that on G_i,

$$\sum_{k=1}^{n} \delta_j(h_k) = Z_1 \circ \Phi + \sum_{m=2}^{l(n)+1} Z_m \circ \Phi - \sum_{k=n+1}^{t(l(n)+1)} \delta_j(h_k)$$

where the middle term is taken to be zero if $l(n) = 0$.

Hence, $\sum_{k=1}^{n} \delta_j(h_k) = Z_1 \circ \Phi + \sum_{m=2}^{n+1} (Z_m \circ \Phi) 1_{\{l(n) \geq m-1\}} - \sum_{k=n+1}^{t(l(n)+1)} \delta_j(h_k)$

Further $\sigma(G_i) = 1$.

So $\sum_{k=1}^{n} \int \delta_j(h_k)\, d\,\sigma(h) = \int \left[\sum_{k=1}^{n} \delta_j(h_k) \right] d\,\sigma(h)$

$$= \int Z_1 \circ \Phi \, d\,\sigma + \sum_{m=2}^{n+1} \int_{\{l(n) \geq m-1\}} Z_m \circ \Phi \, d\,\sigma(h) - \int \left[\sum_{k=n+1}^{t(l(n)+1)} \delta_j(h_k) \right] d\sigma$$

(for sufficiently large n).

Using lemmas 12.3 and 12.4 and dividing by $\sum_{k=1}^{n} \int \delta_i(h_k) d\sigma(h)$,

we get, for large n ,

$$\frac{\sum_{k=1}^{n} \int \delta_j(h_k) d\sigma(h)}{\sum_{k=1}^{n} \int \delta_i(h_k) d\sigma(h)} = \frac{\int Z_1 \circ \Phi \, d\sigma}{\sum_{k=1}^{n} \int \delta_i(h_k) d\sigma(h)} + \frac{\mu_i(\delta_j) \sum_{m=2}^{n+1} \sigma(l(n) \geq m-1)}{\sum_{k=1}^{n} \int \delta_i(h_k) d\sigma(h)} - \frac{\int \left[\sum_{k=n+1}^{t(l(n)+1)} \delta_j(h_k) \right] d\sigma}{\sum_{k=1}^{n} \int \delta_i(h_k) d\sigma(h)}$$

By lemmas 12.1, 12.2, 12.5 and the assumption that

$\sum_{k=1}^{n} \int \delta_i(h_k)\, d\sigma(h) \longrightarrow \infty$ as $n \longrightarrow \infty$, the first and third terms on the right hand side converge to zero as $n \longrightarrow \infty$.

Further for each $n \in N$, since $l(n) \leq n$,

$$\sum_{m=2}^{n+1} \sigma(l(n) \geq m-1) = \int l(n)\, d\sigma = \int \sum_{k=1}^{n} \delta_i(h_k)\, d\sigma(h)$$

$$= \sum_{k=1}^{n} \int \delta_i(h_k)\, d\sigma(h).$$

Therefore the theorem follows.

We conclude this section with an example to show that we can have Markov chains where I is a recurrent class, however for all $i \in I$,

$$\lim_{n \longrightarrow \infty} \int \sum_{k=1}^{n} \delta_i(h_k)\, d\sigma(h) < \infty.$$

Example 12.1. Let $I = N$, $\sigma_o = \delta_1$, $\sigma(1) = \gamma$, a diffuse measure and $\sigma(n+1) = \delta_n$ for all $n \in N$.

Clearly in the above example I is a recurrent class and for each $k \in N$, $\int \delta_1(h_k)\, d\sigma(h) = 0$.

Therefore $\lim_{n \longrightarrow \infty} \int \sum_{k=1}^{n} \delta_1(h_k)\, d\sigma(h) = 0$.

By theorem 12.6, this would imply that $\lim_{n \longrightarrow \infty} \int \sum_{k=1}^{n} \delta_i(h_k) d\sigma(h)$ is finite for all $i \in I$.

<u>Remark</u>. However, in case I is a recurrent class, we always have $\int \sum_{k=1}^{\infty} \delta_i(h_k) \, d\,\sigma(h) = \infty$ for all $i \in I$.

13. <u>Stationary Initial Distributions</u>. Let $\{\sigma(i)\}_{i \in I}$ be a Markov strategy.

<u>Definition</u>. A stationary initial distribution γ for $\{\sigma(i)\}_{i \in I}$ is a finitely additive probability measure defined on all subsets of I such that

$$\int \sigma(j)(E) \, d\,\gamma(j) = \gamma(E)\,, \text{ for all } E \subseteq I. \qquad \ldots(13.1)$$

<u>Lemma 13.1</u>. Let $\{\gamma_n\}$ be a sequence of finitely additive non-negative measures defined on all subsets of I such that there is a positive real number k with $\gamma_n(I) \leq k$ for all $n \in N$. Define the finitely additive measure γ by $\gamma(E) = l(\{\gamma_n(E)\})$ for all $E \subseteq I$, where l is a Banach limit. Then for every bounded real valued function f on I,

$$l(\{\int f \, d\,\gamma_n\}) = \int f \, d\,\gamma.$$

In particular, if $\lim_n \gamma_n(E) = \gamma(E)$ for all $E \subseteq I$, then for every bounded real valued function f on I,

$$\lim_n \int f \, d\,\gamma_n = \int f \, d\,\gamma.$$

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. For a general bounded function f, given $\varepsilon > 0$, get a simple function g such that $|g(i)-f(i)| < \varepsilon/2k$ for all $i \varepsilon I$. Then

$$|l(\{\int f d\gamma_n\}) - \int f d\gamma| \leq |l(\{\int f d\gamma_n\}) - l(\{\int g d\gamma_n\})| + |\int g d\gamma - \int f d\gamma|$$

$$< \frac{\varepsilon}{2k} \cdot k + \frac{\varepsilon}{2k} \cdot k = \varepsilon .$$

As ε is arbitrary, the proof of the first part of the lemma is complete.

The second part of the lemma follows immediately from the first because under the hypotheses of the second part, for any Banach limit l, $l(\{\gamma_n(E)\}) = \gamma(E)$ for all $E \subseteq I$.

For each $n \varepsilon N$ and $i \varepsilon I$ we shall now define inductively a measure $\sigma^n(i)(.)$ on all subsets of I.

Set $\sigma^1(i) = \sigma(i)$, $i \varepsilon I$

and $\sigma^{n+1}(i)(E) = \int \sigma(j)(E)\, d\, \sigma^n(i)(j)$, $E \subseteq I$, $n \varepsilon N$.

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

Theorem 13.2. Every Markov strategy $\{\sigma(i)\}_{i \varepsilon I}$ has a stationary initial distribution.

Proof. Fix an $i \in I$ and a Banach limit l. Let $\gamma_n = \sigma^n(i)$, $n \in N$. Then we claim that γ defined as in lemma 13.1 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 (13.1).

As is evident from the proof of theorem 13.2, 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 13.1. Let $I = N$, $\sigma(1) = \sum_{n=2}^{\infty} p_n \delta_n$ where $\{p_n\}$ is a sequence of nonnegative real numbers such that i) $\sum_{n=2}^{\infty} p_n = 1$ ii) $\sum_{n=2}^{\infty} n\, p_n < \infty$, δ_n being the Dirac measure at n, and $\sigma(n+1) = \delta_n$, $n \in N$.

In the above example it is easy to see that I is a positive recurrent class. It is known in such a case there is a unique countably additive stationary initial distribution (See [8] or [20]). However every finitely additive probability γ defined on all subsets of N, such that $\gamma(E) = \gamma(E+1)$ 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 fairly large.

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 upto multiplication by a constant.

Lemma 13.3. If $i \varepsilon I$ is transient, then $\int Y(i)(h)d\sigma(h) < \infty$ where $Y(i)(h) = \sum_{n=1}^{\infty} \delta_i(h_n)$, δ_i being the function on I defined by

$$\delta_i(j) = 1 \quad \text{if} \quad i = j$$
$$= 0 \quad \text{otherwise.}$$

Proof. Since $Y(i)$ is nonnegative and 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,

$$\int Y(i)(h)d\sigma(h) = \sum_{n=1}^{\infty} \sigma(\{t(1)<\infty\}) \, \sigma[i](\{Y(i) \geq n-1\})$$

$$= \sum_{n=1}^{\infty} \sigma(\{t(1) < \infty\}) \, (f^{*}_{ii})^{n-1}$$

by repeated application of the Strong Markov property

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

Theorem 13.4. Let γ be a stationary initial distribution for a Markov strategy $\{\sigma(i)\}_{i\varepsilon I}$. Suppose $i \varepsilon I$ is transient. Then $\gamma(\{i\}) = 0$.

Proof. Consider the Markov strategy $\sigma = \{\gamma, \{\sigma(i)\}_{i\varepsilon I}\}$. Let us first observe that under this σ,

$$\int 1_E(h_n)\, d\,\sigma(h) = \gamma(E) \quad \text{for all } E \subseteq I, \text{ for each } n \varepsilon N \qquad \ldots.\ (13.2)$$

We verify (13.2) by induction on n. It is true for $n = 1$ because γ is the initial distribution of the strategy σ. Assuming that we have proved (13.2) for $n \leq k$, for $n = k+1$, by conditioning at the k^{th} coordinate and using the change of variable theorem and (13.1), we get the result.

In particular, from (13.2) it follows that

$$\int \delta_i(h_n)\, d\,\sigma(h) = \gamma(\{i\}) \quad \text{for every } n \varepsilon N.$$

Consequently $\sum_{n=1}^{\infty} \int \delta_i(h_n) d\sigma(h) < \infty$ if and only if $\gamma(\{i\}) = 0$.

Since i is transient, and since

$$\int Y(i)d\sigma(h) = \int \Big[\sum_{n=1}^{\infty} \delta_i(h_n) \Big] \, d\sigma(h) \geq \sum_{n=1}^{\infty} \int \delta_i(h_n) d\sigma(h),$$

lemma 13.3 completes the proof of the theorem.

Remark. We have actually shown that (13.2) holds in general for any stationary initial distribution γ for a Markov strategy $\{\sigma(i)\}_{i\varepsilon I}$.

Corollary 13.5. Let I be a countable transient class under $\{\sigma(i)\}_{i\varepsilon I}$. Then every stationary initial distribution is purely finitely additive.

Proof. By Theorem 13.4, for any stationary initial distribution γ, $\gamma(\{i\}) = 0$ for all $i \varepsilon I$. Since I is countable, this implies that γ is purely finitely additive.

Theorem 13.6. Let I be a recurrent class under $\{\sigma(i)\}_{i\varepsilon I}$. Let γ be a stationary initial distribution. If $\gamma(\{i\}) > 0$ for some i, then $\gamma(\{j\}) = \mu_i(\delta_j) . \gamma(\{i\})$ where $\mu_i(\delta_j)$ is as defined before. Hence $\gamma(\{j\}) > 0$ for all $j \varepsilon I$.

Proof. If $\sigma = \{\gamma, \{\sigma(i)\}_{i\varepsilon I}\}$, by the remark after theorem 13.4, for each $n \varepsilon N$, $\int \delta_i(h_n) \, d\sigma(h) = \gamma(\{i\})$, and

$$\int \delta_j(h_n) \, d\sigma(h) = \gamma(\{j\}) \text{ for all } j \varepsilon I.$$

Therefore by Theorem 12.6, $\lim_{n} \frac{n\,\gamma(\{j\})}{n\,\gamma(\{i\})} = \mu_i(\delta_j)$, $j \in I$.

Therefore $\gamma(\{j\}) = \mu_i(\delta_j) \, . \, \gamma(\{i\})$ for all $j \in I$.

To complete the proof, we need to remark that $\mu_i(\delta_j) > 0$.

This is because $\mu_i(\delta_j) = \int Z_1 \, d\sigma[i]$

$$\left[\text{ where } Z_1(h) = \sum_{n=1}^{\infty} \delta_j(h_n) \, 1_{(t(1) \geq n)} \right]$$

$\geq \int_{\{t<t(1)\}} Z_1 d\sigma[i]$, $t(1)$, t being times of first occurrence of i, j respectively.

$\geq \sigma[i](\{t < t(1)\}) > 0$ by lemma 8.1.

Theorem 13.7. Let I be a recurrent class under $\{\sigma(i)\}_{i\in I}$. Let γ_1 be a stationary initial distribution for $\{\sigma(i)\}_{i\in I}$ such that $\gamma_1(\{i\}) > 0$ for some $i \in I$. Then for any other stationary initial distribution γ_2 of $\{\sigma(i)\}_{i\in I}$, there exists a non-negative real number c such that $\gamma_2(\{j\}) = c \, . \, \gamma_1(\{j\})$ for all $j \in I$.

Proof. By the previous theorem, $\gamma_1(\{i\}) > 0$ for some $i \in I$ implies that $\gamma_1(\{i\}) > 0$ for all $i \in I$. If $\gamma_2(\{i\}) = 0$ for all $i \in I$, we take c to be zero. If $\gamma_2(\{i\}) > 0$ for some $i \in I$, by the previous theorem once again, $\gamma_2(\{i\}) > 0$

for all $i \in I$. Further

$$\gamma_2(\{j\}) = \mu_i(\delta_j)\, \gamma_2(\{i\})$$

$$\gamma_1(\{j\}) = \mu_i(\delta_j)\, \gamma_1(\{i\})$$

Therefore $\dfrac{\gamma_2(\{j\})}{\gamma_1(\{j\})} = \dfrac{\gamma_2(\{i\})}{\gamma_1(\{i\})}$ for all $i, j \in I$.

We take c to be this common value and we are done.

Corollary 13.8. Let I be a countable recurrent class such that for some $i \in I$, $\sum_{j\in I} \mu_i(\delta_j) = \infty$. Then every stationary initial distribution is purely finitely additive.

Proof. Since I is countable, it is enough to show that for any stationary initial distribution γ, $\gamma(\{i\}) = 0$ (for then $\gamma(\{j\}) = 0$ for all $j \in I$).

Suppose on the contrary that there exists a stationary initial distribution γ such that $\gamma(\{i\}) > 0$. Then by Theorem 13.6,

$$\gamma(\{j\}) = \mu_i(\delta_j)\, \gamma(\{i\}).$$

Therefore $\sum_{j\in I} \gamma(\{j\}) = (\sum_{j\in I} \mu_i(\delta_j)) \cdot \gamma(\{i\}) = \infty$.

This is a contradiction, since, γ being a probability, $\sum_{j\in I} \gamma(\{j\}) \leq 1$. Hence the corollary is proved.

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

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

Theorem 13.9. Let I be a countable, weakly communicating class. If there exists a countably additive stationary initial distribution, then for each i, $\sigma(i)$ is a countably additive probability.

Proof. By Theorem 13.4, existence of a countably additive stationary initial distribution implies that I cannot be transient. So I is recurrent. If γ is such an initial distribution, then $\sum_{i \varepsilon I} \gamma(\{i\}) = 1$, hence $\gamma(\{i\}) > 0$ for some $i \varepsilon I$, consequently by theorem 13.6, $\gamma(\{i\}) > 0$ for every

$i \varepsilon I$. Now by stationarity of γ,

$$\gamma(\{j\}) = \int \sigma(i)(\{j\}) d\gamma(i) = \sum_{i \varepsilon I} \sigma(i)(\{j\})\gamma(\{i\}) \text{ for all } j \varepsilon I,$$

as γ is countably additive.

Summing over all $j \varepsilon I$, we get

$$\sum_{j\varepsilon I} \gamma(\{j\}) = \sum_{j\varepsilon I} \sum_{i\varepsilon I} \sigma(i)(\{j\}) \gamma(\{i\})$$

$$= \sum_{i\varepsilon I} (\sum_{j\varepsilon I} \sigma(i)(\{j\}))\gamma(\{i\}).$$

Since $\sum_{j\varepsilon I} \gamma(\{j\}) = 1$, and $\gamma(\{i\}) > 0$ for all $i \varepsilon I$ and since $\sum_{j\varepsilon I} \sigma(i)(\{j\}) \leq 1$ for each $i \varepsilon I$, it follows from above that $\sum_{j\varepsilon I} \sigma(i)(\{j\}) = 1$ for each $i \varepsilon I$, which means that $\sigma(i)$ is countably additive for each $i \varepsilon I$. The proof of theorem is thus complete.

The problem of existence of countably additive stationary initial distributions has been studied extensively when $\sigma(i)$ is countably additive for each $i \varepsilon I$. References for this include [8], [20], [21] and [22].

We earlier 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 non-trivial countably additive part. This stationary initial distribution has some desirable properties as we shall see later in this section.

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

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

<u>Lemma 13.10</u>. γ_i is a finitely additive probability on all subsets of I.

<u>Proof</u>. It is clear that γ_i is a finitely additive non-negative measure on all subsets of I. Further

$$\gamma_i(I) = \frac{1}{m_{ii}} \int \Big[\sum_{n=1}^{\infty} 1_{\{t(1)\geq n\}} \Big] d\,\sigma[i] = \frac{1}{m_{ii}} \int t(1) d\sigma[i] = 1.$$

Therefore γ_i is a probability.

Define for each $n \varepsilon N$, the finitely additive measure $\alpha_{n,i}$ by

$$\alpha_{n,i}(E) = \frac{1}{m_{ii}} \int 1_E(h_n)\, 1_{\{t(1)\geq n\}}\, d\sigma[i](h).$$

Let $\gamma_{m,i} = \sum_{n=1}^{m} \alpha_{n,i}$, $m \varepsilon N$.

Lemma 13.11. $\gamma_{m,i}(E) \longrightarrow \gamma_i(E)$ as $m \longrightarrow \infty$ for all $E \subseteq I$.

Proof.

For $E \subseteq I$, $$\gamma_i(E) = \frac{1}{m_{ii}} \int \left[\sum_{n=1}^{\infty} \left(1_E(h_n) 1_{\{t(1)\geq n\}} \right) \right] d\sigma[i]$$

$$\geq \frac{1}{m_{ii}} \int \left[\sum_{n=1}^{m} \left(1_E(h_n) 1_{\{t(1)\geq n\}} \right) \right] d\sigma[i],$$

for each $m \in N$

$$= \frac{1}{m_{ii}} \sum_{n=1}^{m} \int \left(1_E(h_n) 1_{\{t(1)\geq n\}} \right) d\sigma[i]$$

$$= \gamma_{m,i}(E)$$

Therefore $\gamma_i(E) \geq \lim_m \gamma_{m,i}(E)$ for all $E \subseteq I$.

Also, $$\lim_m \gamma_{m,i}(I) = \frac{1}{m_{ii}} \sum_{n=1}^{\infty} \sigma[i]\left(\{t(1)\geq n\}\right) = 1 = \gamma_i(I).$$

Therefore the lemma follows.

Lemma 13.12. For each $m \in N$ and $E \subseteq I$

$$\gamma_{m+1,i}(E) = \frac{\sigma(i)(E)\ (1-m_{ii}\ \gamma_{m,i}(\{i\}))}{m_{ii}} + \int \sigma(j)(E) d\gamma_{m,i}(j)$$

Proof. $$m_{ii}\ \gamma_{m+1,i}(E) = \sum_{n\neq 1}^{m+1} \int 1_E(h_n).1_{(t(1)\geq n)}(h)\ d\sigma[i](h)$$

$$= \sigma(i)(E) + \sum_{n=1}^{m} m_{ii} \int_{\{j\neq i\}} \sigma(j)(E) d\alpha_{n,i}(j)$$

by conditioning the n^{th} term at the $(n-1)^{th}$ coordinate and using the change of variable theorem,

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

$$= \sigma(i)(E)(1-m_{ii}\, \gamma_{m,i}(\{i\})) + m_{ii} \int \sigma(j)(E) \, d\, \gamma_{m,i}(j)$$

Therefore $\gamma_{m+1,i}(E) = \dfrac{\sigma(i)(E)(1-m_{ii}\, \gamma_{m,i}(\{i\}))}{m_{ii}} + \int \sigma(j)(E) d\gamma_{m,i}(j).$

Theorem 13.13. If $i \varepsilon I$ is positive recurrent, then γ_i is a stationary initial distribution.

Proof. Taking limits as $m \longrightarrow \infty$ on both sides of the assertion of lemma 13.12 and using lemma 13.1 and the fact that

$$\lim_{n} \gamma_{m,i}(\{i\}) = \gamma_i(\{i\}) = \frac{1}{m_{ii}} ,$$

we get, $\gamma_i(E) = \int \sigma(j)(E) \, d\, \gamma_i(j).$

i.e. γ_i is a stationary initial distribution.

Example 13.2. Let $I = N$, $\sigma(1) = \gamma$, a diffuse measure and

$$\sigma(n+1) = \delta_1 \text{ , } n \varepsilon N.$$

Clearly 1 is positive recurrent. In fact $m_{11} = 2$. [Th]e above construction gives us the stationary initial distri[but]ion,

$$\gamma_1 = \frac{1}{2}\gamma + \frac{1}{2}\delta_1 .$$

Suppose i is positive recurrent and $i \xrightarrow{W} j$. It follows from corolllary 10.6 that $\gamma_i = \gamma_j$. Therefore if I is a positive recurrent class, we get a canonical stationary initial distribution which we shall denote by γ.

If I is a positive recurrent class and f is a real-valued function on I such that, for some $i \in I$, either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$, then the SLLN for a positive recurrent class asserts that

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

(here, σ_o is arbitrary)

We next show that the σ - a.e limit above is the integral of f with respect to the canonical stationary initial distribution γ.

<u>Theorem 13.14</u>. Let I be a positive recurrent class. If f is a real-valued function on I such that either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$, then $\dfrac{\mu_i(f)}{m_{ii}} = \int f \, d\gamma$.

<u>Proof</u>. If f is an indicator function, the result is true directly from definitions. Hence by linearity of $\mu_i(.)$ and the integral, the result follows for simple functions. For a

bounded function f, given $\varepsilon > 0$, get a simple function g such that $| g(j) - f(j) | < \frac{\varepsilon}{2}$ for all $j \varepsilon I$.

Then $\left| \frac{\mu_i(f)}{m_{ii}} - \int f d\gamma \right| \leq \left| \frac{\mu_i(f)}{m_{ii}} - \frac{\mu_i(g)}{m_{ii}} \right| + \left| \int g\, d\gamma - \int f\, d\gamma \right|$

$$< \frac{\varepsilon}{2m_{ii}} \cdot m_{ii} + \frac{\varepsilon}{2} = \varepsilon .$$

As ε is arbitrary, the result follows for bounded f.

If f is an unbounded nonnegative function,

$$\int f\, d\gamma = \sup_n \int f \wedge n\, d\gamma$$

Therefore, if $\int f\, d\gamma = \infty$, then for every $M > 0$, there is $m \varepsilon N$ such that

$$\int f \wedge m\, d\gamma > M. \text{ Hence } \frac{\mu_i(f)}{m_{ii}} \geq \frac{\mu_i(f \wedge m)}{m_{ii}} = \int f \wedge m\, d\gamma > M.$$

This being true for every M, $\frac{\mu_i(f)}{m_{ii}} = \infty = \int f\, d\gamma$.

If $\int f\, d\gamma < \infty$, given $\varepsilon > 0$, get $m \varepsilon N$ such that $\int f\, d\gamma < \int f \wedge m\, d\gamma + \varepsilon$.

Therefore $\int f\, d\gamma < \int f \wedge m\, d\gamma + \varepsilon = \frac{\mu_i(f \wedge m)}{m_{ii}} + \varepsilon \leq \frac{\mu_i(f)}{m_{ii}} + \varepsilon$.

As ε is arbitrary, $\int f\, d\gamma \leq \frac{\mu_i(f)}{m_{ii}}$.

Also $\frac{\mu_i(f)}{m_{ii}} = \frac{1}{m_{ii}} \int \sum_{k=1}^{\infty} (f(h_k) 1_{\{t(1) \geq k\}}(h)) \, d\, \sigma[i](h)$

$$= \frac{1}{m_{ii}} \operatorname{Sup}_n \int \Big(\sum_{k=1}^{\infty} (f(h_k) 1_{\{t(1) \geq k\}}(h))\Big) \wedge n \, d\, \sigma[i](h)$$

$$\leq \frac{1}{m_{ii}} \operatorname{Sup}_n \int \Big(\sum_{k=1}^{\infty} (f \wedge n)(h_k) 1_{\{t(1) \geq k\}}\Big) \, d\, \sigma[i](h)$$

$$= \operatorname{Sup}_n \int f \wedge n \, d\, \gamma$$

$$= \int f \, d\, \gamma$$

Hence $\frac{\mu_i(f)}{m_{ii}} = \int f \, d\, \gamma$.

For a general f such that either $\mu_i(f^+) < \infty$ or $\mu_i(f^-) < \infty$, the result follows because it is true for f^+ as well as f^-.

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

<u>Theorem 13.15</u>. Let I be a weakly communicating class under $\{\sigma(i)\}_{i \in I}$. The following statements are equivalent :

i) I is positive recurrent and $\sigma(i)$ is countably additive for each $i \in I$.

ii) $\sum_{i \in I} \frac{1}{m_{ii}} = 1$.

Proof. i) ==> ii). Under i) it is known that $\pi(\{i\}) = \frac{1}{m_{ii}}$, $i \in I$ is a countably additive stationary initial distribution (See [8] or [20]). Therefore $\sum_{i\in I} \frac{1}{m_{ii}} = \pi(I) = 1$.

ii) ==> i). If $\sum_{i\in I} \frac{1}{m_{ii}} = 1$, $\frac{1}{m_{ii}} > 0$ for some i, equivalently $m_{ii} < \infty$ for some i. So I, being a weakly communicating class, is positive recurrent. If γ is the canonical stationary initial distribution, then $\gamma(\{i\}) = \frac{1}{m_{ii}}$ for all $i \in I$. ii) implies that γ is countably additive. Then theorem 13.9 completes the proof.

14. Ergodicity of the Shift. The shift transformation T on H is defined by $T(h_1, h_2, \ldots,) = (h_2, h_3, \ldots,)$ for all $h \in H$. The invariant σ-field $\underline{\underline{I}}$ is the collection of all Borel sets A such that $T^{-1}(A) = A$.

Say that T is ergodic under σ if i) $\sigma(T^{-1}(A)) = \sigma(A)$ for all $A \in \underline{\underline{B}}$ and ii) $\sigma(A) = 0$ or 1 for all $A \in \underline{\underline{I}}$.

Proposition 14.1. Let I be a recurrent class under $\{\sigma(i)\}_{i\in I}$. Then under any initial distribution σ_0,

$$\sigma(A) = 0 \text{ or } 1 \text{ for all } A \in \underline{\underline{I}}.$$

Proof. Since I is a recurrent class, for every $i \in I$, $\sigma(t_i < \infty) = 1$ where t_i is the time of first occurrence of i. In fact $\sigma(G_i) = 1$, for all $i \in I$.

Now for $i \in I$, $A \in \underline{\underline{I}}$, $\sigma(A) = \sigma(A \cap (t_i < \infty)) = \sigma(t_i < \infty) . \sigma[i](A)$

by the Strong Markov property,

$$= \sigma[i](A).$$

Therefore, $\sigma(A) = \sigma[i](A)$ for $i \in I$ and $A \in \underline{\underline{I}}$.

Let $\varepsilon > 0$. Get a clopen set K such that $\sigma(A \triangle K) < \frac{\varepsilon}{2}$. Suppose s is stop rule such that K is determined by s.

$$\text{Now } \sigma(A \cap K) = \int \sigma[p_s(h)] \, [(A \cap K) \, p_s(h)] \, d\,\sigma(h)$$

$$= \int_K \sigma[p_s(h)](A) \, d\,\sigma(h)$$

$$= \sigma(A) . \sigma(K) .$$

Therefore $\sigma(A) - [\sigma(A)]^2 \leq | \sigma(A) - \sigma(A \cap K)|$

$$+ |\sigma(A).\sigma(K) - [\sigma(A)]^2 |$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon .$$

As ε is arbitrary, $\sigma(A) = [\sigma(A)]^2$. Therefore $\sigma(A) = 0$ or 1.

Remark. In course of the proof of proposition 14.1, we have actually proved that if I is a recurrent class under $\{\sigma(i)\}_{i \in I}$, then under any initial distribution σ_o , $\sigma(A) = \sigma[i](A)$ for all $i \varepsilon I$ and for all $A \varepsilon \underline{\underline{I}}$.

<u>Theorem 14.2</u>. Let I be a recurrent class under $\{\sigma(i)\}_{i\varepsilon I}$. Let γ be a stationary initial distribution for $\{\sigma(i)\}_{i\varepsilon I}$. Then the shift T is ergodic under $\sigma = \{\gamma, \{\sigma(i)\}_{i\varepsilon I}\}$.

<u>Proof</u>. In view of proposition 14.1, it is enough to verify that $\sigma(A) = \sigma(T^{-1}(A))$ for all $A \varepsilon \underline{\underline{B}}$.

For $A \varepsilon \underline{\underline{B}}$,
$$\sigma(T^{-1}(A)) = \int \sigma[h_1,h_2](T^{-1}(A)h_1,h_2)d\,\sigma(h)$$
$$= \int \sigma[h_2](Ah_2)\; d\;\sigma(h)$$
$$= \int \sigma[j](Aj)\; d\;\gamma(j)$$
, by change of variable theorem and (13.2),
$$= \sigma(A).$$

Hence the theorem is proved.

<u>Theorem 14.3</u>. Let I be a recurrent class under $\{\sigma(i)\}_{i\varepsilon I}$. Then under any initial distribution σ_o, the measure σ is countably additive on $\underline{\underline{I}}$, the invariant σ-field.

<u>Proof</u>. In view of proposition 14.1, it is sufficient to show that if $\{A_n\}_{n \geq 1}$ is a sequence of sets in $\underline{\underline{I}}$ and $\sigma(A_n) = 0$ for all $n \varepsilon N$, then $\sigma(\bigcup_n A_n) = 0$. Because of the remark

after proposition 14.1, $\sigma[i](A_n) = 0$ for all $i \in I$ and all $n \in N$. We can now apply proposition (IV) of section 1 and show that $\sigma(\bigcup_n A_n) < \varepsilon$ for every $\varepsilon > 0$.

This completes the proof of the theorem.

Remark. Theorem 14.3 is in the same spirit as the following result of Purves and Sudderth [29] :

An independent strategy σ is countably additive on the tail σ-field.

15. Almost closed sets and the Feller Boundary.

In this section we shall need the following generalization of proposition I of section 1.

Let g be a bounded Borel measurable function on H. Let s be any stop rule. Then for any strategy σ,

$$\int g \, d\sigma = \int \left\{ \int g \circ p_s(h) \, d\sigma[p_s(h)] \right\} d\sigma(h),$$

where for $h \in H$, $g \circ p_s(h)$ is the function on H defined by

$$g \circ p_s(h)(h') = g(p_s(h)\, h'), \; h' \in H.$$

Proposition I of section 1 is a special case of the above assertion with $g = 1_A$, $A \in \mathcal{B}$. The above assertion is proved from proposition I of section 1 the routine way by first proving it for simple functions and then for a general bounded measurable function by writing it as a uniform limit of simple functions.

We also need the following four additional theorems of Purves and Sudderth [29].

<u>Theorem 15.α</u> (Levy 0-1 law) : Let σ be a strategy and $A \varepsilon \underline{\underline{B}}$. Then $\sigma(\{h : \sigma[p_n(h)](A\, p_n(h)) \longrightarrow 1_A(h)\}) = 1$.

Let σ be a strategy. Let $\{Y^n\}_{n \geq 0}$ be a sequence of real valued functions on H such that for all n, Y^n is of structure at most n. So for $n \geq 1$, Y^n depends only on the first n coordinates and we can write $Y^n(h) = Y^n(h_1, \ldots, h_n)$. If $\int Y^1 d\sigma_o \leq Y^o$ and for all $n \geq 1$ and $i_1, \ldots, i_n \varepsilon I$, $\int Y^{n+1}(i_1, \ldots, i_n, j)\, d\sigma(i_1, \ldots, i_n)(j) \leq Y^n(i_1, \ldots, i_n)$, then $\{Y^n\}_{n \geq 0}$ is called a Supermartingale. If equality holds in all the inequalities above, $\{Y^n\}_{n \geq 0}$ is called a Martingale.

<u>Theorem 15.β</u> (Supermartingale Convergence Theorem). If $\{Y^n\}_{n \geq 0}$ is a uniformly bounded supermartingale with respect to a strategy σ, then $\sigma(\{h : Y^n(h) \text{ converges}\}) = 1$.

<u>Theorem 15.γ</u>. Let σ be a strategy and g a bounded, real-valued Borel measurable function. Define $Y^o(h) = \int g\, d\sigma$, $Y^n(h) = \int g \circ p_n(h)\, d\sigma[p_n(h)]$, for all $h \varepsilon H$, $n \varepsilon N$. Then

$\{Y^n\}_{n\geq 0}$ is a bounded martingale with respect to σ, and, hence, converges σ-almost surely. Moreover if $\emptyset = \limsup_n Y^n$, then for every $\varepsilon > 0$, $\sigma(\{h : |\emptyset(h) - g(h)| > \varepsilon\}) = 0$.

<u>Theorem 15.6</u>. Let $\{g_n\}_{n\geq 1}$ be a uniformly bounded sequence of real-valued functions on H such that for every n, g_n is of structure at most n. Then for every strategy σ,

$$\limsup_s \int g_s \, d\sigma = \int (\limsup_n g_n) \, d\sigma,$$

where the first lim sup is taken over stop rules and the second over positive integers, and $g_s(h) = g_{s(h)}(h)$.

Consequently, if $\lim_n g_n$ exists, then

$$\int \lim_n g_n \, d\sigma = \lim_s \int g_s \, d\sigma.$$

With these preliminaries we come back to a fixed Markov strategy σ. As in the previous section, $\underline{\underline{I}}$ will stand for the shift invariant σ-field.

<u>Notation</u>. F[illegible] subset E of I,

$$\{E \text{ i.o}\} = \{h \in H : h_n \in E \text{ for infinitely many } n\},$$

and $\{E \text{ eventually}\} = \{h \in H :$ there exists $n_o \in N$ such that $h_n \in E$ for all $n \geq n_o\}$.

Let $\underline{\underline{C}} = \{E : E \subseteq I \text{ and } \sigma(\{E \text{ i.o}\}) = \sigma(\{E \text{ eventually}\})\}$,

and $\underline{\underline{T}} = \{E : E \in \underline{\underline{C}} \text{ and } \sigma(\{E \text{ i.o}\}) = 0\}$.

<u>Lemma 15.1</u>. The collection $\underline{\underline{C}}$ is a field of subsets of I and the collection $\underline{\underline{T}}$ is an ideal of subsets of I.

<u>Proof</u>. Clearly $\underline{\underline{C}}$ contains I. It is closed under complementation and finite intersections because

$$\text{a) } \{E \text{ i.o}\} - \{E \text{ eventually}\} = \{E^c \text{ i.o}\} - \{E^c \text{ eventually}\}$$

$$\text{and b) } \{E_1 \cap E_2 \text{ i.o}\} - \{E_1 \cap E_2 \text{ eventually}\}$$
$$\subseteq \{E_1 \text{ i.o} - E_1 \text{ eventually}\} \cup \{E_2 \text{ i.o} - E_2 \text{ eventually}\}$$

for all $E, E_1, E_2 \subseteq I$.

The collection $\underline{\underline{T}}$ clearly contains the empty set, and if $E_1 \varepsilon \underline{\underline{T}}$ and $E_2 \subseteq E_1$, then $E_2 \varepsilon \underline{\underline{T}}$.

It is also closed under finite unions because

$$\text{c) } \{E_1 \cup E_2 \text{ i.o}\} - \{E_1 \cup E_2 \text{ eventually}\}$$
$$\subseteq \{E_1 \text{ i.o} - E_1 \text{ eventually}\} \cup \{E_2 \text{ i.o} - E_2 \text{ eventually}\}$$

The lemma is hence proved.

<u>Definition</u>. A set $A \varepsilon \underline{\underline{I}}$ will be called a σ-atom of $\underline{\underline{I}}$ in case $\sigma(A) > 0$ and for every $B \varepsilon \underline{\underline{I}}$, $B \subseteq A$ implies that either $\sigma(B) = 0$ or $\sigma(B) = \sigma(A)$.

Let '$\sim$' be the equivalence relation on $\underline{\underline{I}}$ defined by, $A \sim B$ if and only if $\sigma(A \triangle B) = 0$. Let $\underline{\underline{N}}$ be the σ-null sets of $\underline{\underline{I}}$. Clearly $\underline{\underline{N}}$ is an ideal. We have $A \sim B$ if and only if $A \triangle B \in \underline{\underline{N}}$. It is easy to see that if A is a σ-atom of $\underline{\underline{I}}$ and $A \sim B$, then B is also a σ-atom of $\underline{\underline{I}}$.

Let ξ be the mapping on $\underline{\underline{I}}$ into the power set of I defined by $\xi(A) = \{ i \in I : \sigma[i](A) > \frac{1}{2} \}$, $A \in \underline{\underline{I}}$.

<u>Lemma 15.2</u>. For $A \in \underline{\underline{I}}$, $A \sim \{\xi(A) \text{ i.o}\} \sim \{\xi(A) \text{ eventually}\}$. Consequently ξ is a map on $\underline{\underline{I}}$ into $\underline{\underline{C}}$.

<u>Proof</u>. By the Levy 0-1 law (Theorem 15.α), for any $A \in \underline{\underline{I}}$, $\sigma(\{h : \sigma[h_n](A) \longrightarrow 1_A(h)\}) = 1$. Let B denote this set of convergence. Then

$$A \cap B = \{\xi(A) \text{ i.o}\} \cap B = \{\xi(A) \text{ eventually}\} \cap B.$$

Hence the lemma follows.

<u>Lemma 15.3</u>. For $A, B \in \underline{\underline{I}}$, if $A \sim B$, then $\xi(A) \triangle \xi(B) \in \underline{\underline{T}}$. Consequently we can use ξ to define a map (also to be denoted by ξ) on $\underline{\underline{I}}/\underline{\underline{N}}$ onto $\underline{\underline{C}}/\underline{\underline{T}}$. This map is a Boolean algebra isomorphism on $\underline{\underline{I}}/\underline{\underline{N}}$ onto $\underline{\underline{C}}/\underline{\underline{T}}$.

<u>Proof</u>. Follows easily from lemmas 15.1 and 15.2.

<u>Definition</u>. A subset E of I is called almost closed if $\{E \text{ i.o}\} \sim \{E \text{ eventually}\}$ and $\sigma(\{E \text{ i.o}\}) > 0$. Equivalently, E is almost closed if and only if $E \in \underline{\underline{C}}$ and $E \notin \underline{\underline{T}}$.

Definition. An almost closed set is called an atomic almost closed set just in case it does not contain two disjoint almost closed sets.

Definition. An almost closed set is called completely non-atomic just in case it does not contain any atomic closed subset.

Lemma 15.4. If $A \in \underline{\underline{I}}$ is a σ-atom of $\underline{\underline{I}}$, then $\xi(A)$ is an atomic almost closed set. Also if E is an atomic almost closed set then $\{E \text{ i.o}\}$ is a σ-atom of $\underline{\underline{I}}$.

Proof. If $A \in \underline{\underline{I}}$ is a σ-atom of $\underline{\underline{I}}$, then by lemma 15.2, $\xi(A)$ is almost closed and further $\{\xi(A) \text{ i.o}\}$ is a σ-atom of $\underline{\underline{I}}$. Therefore if E_1 and E_2 are disjoint subsets of $\xi(A)$, then $\sigma(\{E_1 \text{ i.o}\})$ and $\sigma(\{E_2 \text{ i.o}\})$ cannot both be positive and hence E_1, E_2 cannot both be almost closed. Therefore $\xi(A)$ is an atomic almost closed set.

For the other part, suppose $\{E \text{ i.o}\}$ is not a σ-atom of $\underline{\underline{I}}$. Then, E being almost closed, $\{E \text{ eventually}\}$ is not a σ-atom of $\underline{\underline{I}}$. So there exist $A, B \in \underline{\underline{I}}$ such that $A \cap B = \emptyset$, $A, B \subseteq \{E \text{ eventually}\}$ and $\sigma(A) > 0$ and $\sigma(B) > 0$. It is now easy to see that $\xi(A) \cap E$ and $\xi(B) \cap E$ are disjoint almost closed subsets of E. Therefore E cannot be atomic, a contradiction. Hence the result follows.

<u>Theorem 15.5</u>. Let σ be a Markov strategy on I. Then there exists a sequence $\{E_n\}_{n\geq 0}$, of disjoint subsets of I such that

i) E_o, if non-empty, is a completely non-atomic almost closed set,

ii) E_n, if non-empty, is an atomic almost closed set, for $n \geq 1$,

iii) $\sigma(\bigcup_{n=0}^{\infty} \{E_n \text{ eventually}\}) = 1$.

<u>Proof</u>. Let $\underline{\underline{D}} = \{A \in \underline{\underline{I}} : A$ is a σ-atom of $\underline{\underline{I}}\}$. It is easy to see that there can only be countably many equivalence classes of $\underline{\underline{D}}$ under '$\sim$' defined before. Let $\{A_n\}_{n\geq 1}$, be a disjoint sequence in $\underline{\underline{I}}$ with the property that the sequence contains one and only one representative from each equivalence class of $\underline{\underline{D}}$ and if a set in the sequence is not a representative for any equivalence class of $\underline{\underline{C}}$, it is empty.

Now consider the disjoint sequence $\{\xi(A_n)\}_{n\geq 1}$, of subsets of I. If $\sigma(\bigcup_{n=1}^{\infty} \{\xi(A_n) \text{ eventually}\}) = 1$, then the sequence defined by $E_o = \emptyset$, $E_n = \xi(A_n)$, $n \geq 1$, by lemma 15.4, satisfies i), ii) and iii).

Suppose $B = \bigcup_{n=1}^{\infty} \{\xi(A_n) \text{ eventually}\}$ has σ-measure less than one. Let $E_o = \xi(B^c)$, $E_n = \xi(A_n) \cap [\xi(B^c)]^c$, $n \geq 1$.

We claim that $\{E_n\}_{n\geq 0}$ satisfies i), ii) and iii).

Clearly $\{E_n\}_{n\geq 0}$ is a disjoint sequence of subsets of I. Also for $n \geq 1$,

$$\{E_n \text{ eventually}\} = \{\xi(A_n) \text{ eventually}\} \cap \{[\xi(B^c)]^c \text{ eventually}\}$$
$$\sim \{\xi(A_n) \text{ eventually}\} \cap B$$
$$= \{\xi(A_n) \text{ eventually}\}$$
$$\sim A_n .$$

Therefore, for each $n \geq 1$, E_n is an atomic almost closed set provided A_n was one. Further by lemma 15.4, E_o cannot contain an atomic almost closed set, because for any σ-atom of Σ, there is a member of the sequence $\{A_n\}$ equivalent to it. Hence E_o has to be completely non-atomic.

Now observe that

$$\sigma(B \cap [\bigcup_{n=0}^{\infty} \{E_n \text{ eventually}\}])$$

$$= \sigma(B \cap \{E_o^c \text{ eventually}\} \cap [\bigcup_{n=0}^{\infty} \{E_n \text{ eventually}\}])$$

because $B \sim \{E_o^c \text{ eventually}\}$,

$$= \sigma(B \cap \{E_o^c \text{ eventually}\}),$$

(the last equality holds because the two sets under consideration are the same),

$$= \sigma(B) .$$

Further,

$$\sigma(B^c \cap [\bigcup_{n=0}^{\infty} \{E_n \text{ eventually}\}]) = \sigma(\{E_0 \text{ eventually}\})$$
$$= \sigma(B^c) .$$

Consequently iii) follows and the theorem is proved.

In order to prove uniqueness in some sense of the above decomposition, we need an additional assumption on σ .

Condition (*). A Markov strategy σ is said to satisfy condition (*) , if for each $i \varepsilon I$, $\sigma(\{t_i < \infty\}) > 0$, where t_i is the time of first occurrence of i .

Lemma 15.6. If the Markov strategy σ satisfies condition (*), then for every $A \varepsilon \underline{\underline{I}}$, $\sigma(A) = 0$ if and only if $\sigma[i](A) = 0$ for every $i \varepsilon I$.

Proof. The if part is clearly true for any Markov strategy σ . Further for each $i \varepsilon I$,

$$\sigma(A) \geq \sigma(A \cap \{t_i < \infty\})$$
$$= \sigma(\{t_i < \infty\}) \, \sigma[i](A), \text{ by the Strong Markov property.}$$

If $\sigma(A) = 0$ and condition (*) is satisfied, the above inequality implies that $\sigma[i](A) = 0$ for all $i \varepsilon I$. Hence the lemma is proved.

Lemma 15.7. If the Markov strategy σ satisfies condition (*), then for any sequence $\{A_n\}$ of sets in $\underline{\underline{I}}$, $\sigma(A_n) = 0$ for all $n \geq 1$ implies that $\sigma(\bigcup_{n=1}^{\infty} A_n) = 0$. Consequently $\underline{\underline{N}}$ is a σ-ideal.

Proof. We can apply proposition IV of section 1 to the sequence $\{A_n\}$ for every $\varepsilon > 0$, because of lemma 15.6. We shall then get $\sigma(\bigcup_{n=1}^{\infty} A_n) < \varepsilon$ for every $\varepsilon > 0$, i.e. $\sigma(\bigcup_{n=1}^{\infty} A_n) = 0$.

Remark. Although under condition (*), $\underline{\underline{N}}$ is a σ-ideal, the measure σ on $\underline{\underline{I}}$ need not be countably additive.

Example 15.1. Let $I = N \times N$, $\sigma_0 = \sum_{m,n\varepsilon N} p_{m,n}\, \delta_{(m,n)} + (1 - \sum_{m,n} p_{m,n})\gamma$ where $\{p_{m,n}\}$ are positive real numbers such that $0 < \sum_{m,n} p_{m,n} < 1$, γ is a measure such that $\gamma(E_m) = 0$ for all $m \varepsilon N$, where $E_m = \{(m,n) : n \varepsilon N\}$. Let $\sigma(m,n) = \delta_{(n,n+1)}$ for $m,n \varepsilon N$.

Then $\sigma(E_m \text{ eventually}) = \sum_{n=1}^{\infty} p_{m,n}$, for each $m \geq 1$.

However $\sigma(\bigcup_{m=1}^{\infty} E_m \text{ eventually}) = 1 > \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} p_{m,n}$.

By choice of σ_0, it is also easy to see that σ satisfies condition (*).

Definition. A subset E of $\underline{\underline{I}}$ is transient if $\sigma[i](\{E \text{ i.o}\})=0$ for each $i \varepsilon I$.

If σ satisfies (*), clearly E is transient if and only if $E \varepsilon \underline{\underline{T}}$.

Theorem 15.8. Suppose σ is a Markov strategy satisfying condition (*). If $\{E_n\} n \geq 0$ and $\{E'_n\}$ $n \geq 0$ are two sequences satisfying i), ii) and iii) of theorem 15.5, then

a) $(E_o \Delta E'_o)$ is transient

b) For each $n \varepsilon N$, if E_n is non-empty,

then $(E_n \Delta E'_{n'})$ is transient for some $n' \varepsilon N$.

Proof. For each $n \varepsilon N$, if E_n is non-empty, by lemma 15.4, $\{E_n \text{ i.o}\}$ is a σ-atom and so by i), ii), iii) and Lemma 15.4, there exists $n' \varepsilon N$ such that $\{E_n \text{ i.o}\} \sim \{E'_{n'} \text{ i.o}\}$. Therefore $\sigma\{(E_n \Delta E'_{n'}) \text{ i.o}\} = 0$. In view of lemma 15.6, this implies that $(E_n \Delta E'_{n'})$ is transient and thus b) is proved.

Now by Lemma 15.7, $[\bigcup_{n=1}^{\infty}\{E_n \text{ i.o}\}] \sim [\bigcup_{n=1}^{\infty}\{E'_n \text{ i.o}\}]$.

Hence by iii) and Lemma 15.1,

$$\{E_o \text{ i.o}\} \sim [\bigcup_{n=1}^{\infty}\{E_n \text{ i.o}\}]^c \sim [\bigcup_{n=1}^{\infty}\{E'_n \text{ i.o}\}]^c \sim \{E'_o \text{ i.o}\}$$

Hence a) is also satisfied. Thus the theorem is proved.

Remark. In the countably additive case, the almost closed sets could be so chosen that states in such a set are either all recurrent or all transient [1]. Such a decomposition is not in general possible as shown by the example below.

Example 15.2. $I = N$, $\sigma_0 = \delta_1$, $\sigma(1) = \gamma$ where γ is a diffuse measure and $\sigma(n) = \delta_1$, $n \geq 2$.
Here 1 is the only recurrent state but $\sigma(\{1 \text{ eventually}\}) = 0$.

Definition. A Markov chain will be called simply atomic in case in the above decomposition, we get only one non-empty set and that is an atomic almost closed set.

Therefore a Markov chain is simply atomic if and only if $\sigma(A) = 0$ or 1 for all $A \in \underline{\underline{I}}$.

Theorem 15.9. If I is a recurrent class under $\{\sigma(i)\}_{i\varepsilon I}$, then under any initial distribution σ_0, $\sigma = \{\sigma_0, \{\sigma(i)\}_{i\varepsilon I}\}$ is simply atomic.

Proof. The above theorem is just a restatement of proposition 14.1.

Definition. A real-valued $\underline{\underline{I}}$-measurable function on H will be called an invariant function.

Suppose σ satisfies condition (*). Then $\underline{\underline{N}}$ is a σ-ideal, and $L_\infty(H, \underline{\underline{I}}, \sigma)$, the space of invariant functions equipped with the essential supremum norm, is a Banach space.

For a Markov strategy σ, a bounded function g is said to be harmonic if it satisfies

$$g(i) = \int g(j)\, d\, \sigma(i)(j) \quad \text{for all} \quad i \in I \qquad \ldots(15.1)$$

It is easy to see that for any Markov strategy σ, the space of harmonic functions with the supremum norm is a Banach space.

<u>Theorem 15.10</u>. Let σ be a Markov strategy. Let f be a bounded invariant function. Then g, defined by $g(i) = \int f\, d\, \sigma[i]$, $i \in I$, is a harmonic function.

Conversely, if g is a harmonic function, there exists a bounded invariant function f such that

$$g(i) = \int f\, d\, \sigma[i]\,,\ i \in I.$$

If σ further satisfies condition (*), then $f \longmapsto g$, where $g(i) = \int f\, d\, \sigma[i]$, $i \in I$, is a linear isometric isomorphism of $L_\infty(H, \underline{\underline{I}}, \sigma)$ and the Banach space $\underline{\underline{H}}$ of harmonic functions with the supremum norm.

<u>Proof</u>. If f is bounded invariant, by the assertion stated at the beginning of this section,

$$g(i) = \int f\, d\, \sigma[i] = \int \left\{ \int f\, d\, \sigma[j] \right\} d\, \sigma(i)(j)$$
$$= \int g(j)\, d\, \sigma(i)(j).$$

Hence g is harmonic.

For the converse, set $f(h) = \limsup_n \, g(h_n)$. Clearly f is a bounded invariant function.

By theorem 15.6,

$$\int f \, d\,\sigma[i] = \limsup_s \int g(h_s) \, d\,\sigma[i](h) .$$

We shall verify by induction on the structure of stop rules, that $\int g(h_s) \, d\,\sigma[i](h) = g(i)$ for each stop rule s. For $s \equiv 1$, this follows from (15.1). Also for $n \in N$,

$$\int g(h_{n+1}) \, d\,\sigma[i](h) = \int \left\{ \int g \circ j(h_{n+1}) \, d\,\sigma[j](h) \right\} d\sigma(i)(j)$$

where $g \circ j$ is the function on H defined by $g \circ j(h) = g(jh)$, $h \in H$

$$= \int \left\{ \int g(h_n) \, d\,\sigma[j](h) \right\} d\,\sigma(i)(j).$$

Therefore it follows by ordinary induction on n that the assertion is true for all stop rules of structure zero. Assume that we have proved the result for stop rules of structure $< \alpha$. Let s be a stop rule of structure α.

Then, $\int g(h_s) \, d\,\sigma[i](h) = \int \left\{ \int g \circ j(h_s) \, d\,\sigma[j](h) \right\} d\sigma(i)(j)$

For $j \in I$ such that $s(jh) = 1$ for some $h \in H$ (and hence all $h \in H$), $g \circ j$ is the constant function $g(j)$. For $j \in I$ such that $s(jh) > 1$ for all $h \in H$, $g \circ j(h_s) = g(h_{s_j})$, where s_j is the stop rule defined by $s_j(h) = s(jh) - 1$, $h \in H$. Clearly s_j is a stop rule of structure $< \alpha$. Therefore by the induction hypothesis, $\int g \circ j(h_s)\, d\,\sigma[j](h) = g(j)$ for such j as well. Therefore $\int g(h_s)\, d\,\sigma[i](h) = \int g(j) d\sigma(i)(j) = g(i)$. Therefore the second part of the theorem is proved.

Suppose now σ satisfies condition (*). The map $f \longmapsto g$ on $L_\infty(H, \underline{\underline{I}}, \sigma)$ into $\underline{\underline{H}}$, defined in the statement of the theorem, is clearly linear and is onto by the second part of the theorem. We only need to check that it preserves norms. By lemma 15.7, $\sigma(\{|f| > \|f\|_\infty\}) = 0$. Therefore by lemma 15.6, $\sigma[i](\{|f| > \|f\|_\infty\}) = 0$ for each $i \in I$. Consequently $|\int f\, d\,\sigma[i]| \leq \int |f|\, d\,\sigma[i] \leq \|f\|_\infty$, which implies that $\|g\| = \sup_i |g(i)| \leq \|f\|_\infty$.

Again, by theorem 15.7 applied to f and by lemma 15.7,

$$\sigma\left(\left\{h : \int f\, d\,\sigma[h_n] \longrightarrow f(h)\right\}\right) = 1,$$

i.e. $\sigma\left(\left\{h : g(h_n) \longrightarrow f(h)\right\}\right) = 1$. Therefore

$$\sigma\left(\left\{h : |g(h_n)| \longrightarrow |f|(h)\right\}\right) = 1.$$

Clearly on this set of convergence $|f|(h) \leq \|g\|$.

Therefore $\|f\|_\infty \leq \|g\|$. Hence theorem 15.10 is completely proved.

We are now in a position to make a few remarks on the Feller boundary of a Markov chain. Let σ be a Markov strategy satisfying condition (*). The Feller boundary of the Markov chain is the Stone space of the extreme points of the convex set of nonnegative elements of the unit ball of $\underline{\underline{H}}$, equivalently, in view of theorem 15.10, the Stone space of the extreme points of the convex set of nonnegative elements of the unit ball of $L_\infty(H, \underline{\underline{I}}, \sigma)$. The nonnegative extreme points of the unit ball of $L_\infty(H, \underline{\underline{I}}, \sigma)$ are precisely the equivalence classes $[1_A]$, $A \in \underline{\underline{I}}$.

Therefore the extreme points of the nonnegative elements of the unit ball of $\underline{\underline{H}}$ can be identified with $\underline{\underline{I}}/\underline{\underline{N}}$. Since $\underline{\underline{N}}$ is a σ-ideal, $\underline{\underline{I}}/\underline{\underline{N}}$ is a Boolean σ-algebra. Further it satisfies the countable chain condition (because for any finitely additive probability, at most countably many disjoint sets can have positive probability). Therefore $\underline{\underline{I}}/\underline{\underline{N}}$ is a complete Boolean algebra. Hence its Stone space is a compact Hausdorff extremally disconnected space. Under condition (*), the map ξ defined in lemma 15.3 is an isomorphism of $\underline{\underline{I}}/\underline{\underline{N}}$ and $\underline{\underline{C}}/\underline{\underline{T}}$. Therefore the Feller boundary can also be looked upon as the Stone space of $\underline{\underline{C}}/\underline{\underline{T}}$. So we have proved the following

<u>Theorem 15.11</u>. Let σ be a Markov strategy satisfying condition (*). The Feller boundary M of I under σ is the Stone space of $\underline{\underline{C}}/\underline{\underline{T}}$ ($\underline{\underline{C}}$, $\underline{\underline{T}}$ as defined before) and is a compact Hausdorff extremally disconnected space.

As in the countably additive case, we can topologize IUM giving it a suitable topology, under which it is Hausdorff. Now on this space the harmonic functions admit a continuous extension. The details are the same as in the countably additive case, hence omitted. (See [15a] and [16]).

We conclude this section with a characterization of simply atomic chains which will be used in the next section.

<u>Theorem 15.12</u>. Let σ be a Markov strategy. If the only harmonic functions are constants, then the Markov chain is simply atomic. If σ is also assumed to satisfy condition(*), then the Markov chain is simply atomic implies that all harmonic functions are constants.

<u>Proof</u>. Let $A \in \underline{\underline{I}}$. Then by theorem 15.10, $g(i) = \sigma[i](A)$, $i \in I$ is a harmonic function. Also by the Levy 0-1 law, $\sigma(\{h : \sigma[h_n](A) \longrightarrow 1_A(h)\}) = 1$. Therefore if all harmonic functions are constants, then either $\sigma[i](A) = 1$ for all $i \in I$ or $\sigma[i](A) = 0$ for all $i \in I$. In the first case $\sigma(A) = 1$ and in the second $\sigma(A) = 0$. Therefore the chain is simply atomic.

Conversely, if the Markov chain is simply atomic and satisfies condition (*), then by definition of a simply atomic chain and lemma 15.6, for $A \in \underline{\underline{I}}$, either $\sigma[i](A) = 0$ for all $i \in I$ or $\sigma[i](A) = 1$ for all $i \in I$. So if $f = 1_A$, then $g(i) = \int f \, d\sigma[i]$ is a constant. Consequently, by linearity, if f is a simple function $\int f \, d\sigma[i]$ is a constant. Any bounded invariant function f is a uniform limit of a sequence $\{f_n\}$ of simple invariant functions and therefore $\int f \, d\sigma[i] = \operatorname{Lim}_n \int f_n \, d\sigma[i]$ for all $i \in I$. Consequently, if f is any bounded invariant function, $\int f \, d\sigma[i]$ is constant. By theorem 15.10 every harmonic function is of the above form and hence we are done.

16. Choquet - Deny Equations and Random Walks.

Throughout this section we shall assume that I is the group of all integers with '+' the usual addition.

Let γ be a finitely additive probability defined on all subsets of I. Consider the equations

$$f(i) = \int f(i+j) \, d\gamma(j), \quad \text{for all } i \in I \qquad \ldots(16.1)$$

(16.1) will be called the Choquet - Deny equations and we shall be studying bounded real-valued solutions of (16.1).

Proposition 16.1. If γ is a probability on all subsets of I such that $\gamma(\{n_o\}) > 0$ for some $n_o \in I$, then every bounded real-valued solution f of (16.1) is of period n_o, i.e. $f(i+n_o) = f(i)$ for all $i \in I$.

(<u>Remark</u>. f being of period n_o is equivalent to f being of period $|n_o|$).

<u>Proof</u>. If f is a bounded solution of (16.1), it is easy to check that g defined by $g(i) = f(i+n_o) - f(i)$, $i \in I$, is also a bounded solution of (16.1). Suppose $\operatorname{Sup}_{i\in I} g(i) = a$.

Then $g(i) \leq \alpha\, g(i+n_o) + (1-\alpha)\, a$ ....(16.2)

where $\alpha = \gamma(\{n_o\})$.

Therefore, $g(i) \leq \alpha^n g(i+nn_o) + (1-\alpha^n)a$, $n = 0,1,2,$ ....

by repeated application of (16.2).

Hence $g(i) \leq a + \alpha^n[g(i+nn_o) - a]$, $n = 0,1,2,\ldots.$

Let $\varepsilon > 0$ and M be any positive integer. Since $a = \operatorname{Sup}_i g(i)$, we can choose $i_o \in I$ such that $a - \varepsilon\alpha^M \leq g(i_o)$. Then

$$a - \varepsilon\alpha^M \leq g(i_o) \leq a + \alpha^n[g(i_o+nn_o) - a],\ n = 0,1,2,\ldots.$$

Therefore, $a-\varepsilon \leq g(i_o+nn_o)$, for $n = 0,1,2,\ldots, M-1$.

Adding, we get,

$$M(a-\varepsilon) \leq \sum_{n=0}^{M-1} g(i_o+nn_o) = f(i_o+Mn_o) - f(i_o) \leq 2\,C$$

where $C > 0$ is such that $|f(i)| \leq C$ for all $i \in I$.

Since the above inequality is true for all positive integers M and all $\varepsilon > 0$, it follows that $a \leq 0$ i.e. $\operatorname{Sup}_{i\in I} g(i) \leq 0$.

Working now with $-f$, which is also a bounded solution of (16.1), we can show that $\inf_i g(i) \geq 0$.

Consequently $g(i) = 0$ for all $i \in I$.

i.e. $f(i) = f(i+n_o)$ for all $i \in I$.

<u>Proposition 16.2.</u> Let γ be any finitely additive probability defined on all subsets of I. If f is a bounded solution of (16.1) of period $n_o (\in N)$, then either

$$\gamma(i \in I : i \text{ is a multiple of } n_o) = 1$$

or f is of period n_1 for some $n_1 \in N$ such that $n_1 < n_o$ and n_1 divides n_o.

<u>Proof.</u> Let $I_r = \{i \in I : i = nn_o + r \text{ for some integer } n\}$,

$$r = 0, \ldots, n_o - 1 .$$

Since f is of period n_o, f is constant on each of the I_r's. Suppose $f(i) = a_r$ for all $i \in I_r$, $r = 0, \ldots, n_o - 1$.

From (16.1), we now have

$$\left.\begin{aligned}
a_o &= a_o\,\gamma(I_o) + a_1\,\gamma(I_1) + \cdots + a_{n_o-1}\,\gamma(I_{n_o-1}) \\
a_1 &= a_1\,\gamma(I_o) + a_2\,\gamma(I_1) + \cdots + a_o\,\gamma(I_{n_o-1}) \\
a_2 &= a_2\,\gamma(I_o) + a_3\,\gamma(I_1) + \cdots + a_1\,\gamma(I_{n_o-1}) \\
&\vdots \\
a_{n_o} &= a_{n_o-1}\,\gamma(I_o) + a_o\,\gamma(I_1) + \cdots + a_{n_o-2}\,\gamma(I_{n_o-1})
\end{aligned}\right\} \quad \ldots (16.3)$$

Either $\gamma(I_o) = 1$, in which case we are done, or

$$\left.\begin{aligned} a_o &= a_1 \lambda_1 + \ldots + a_{n_o-1} \lambda_{n_o-1} \\ a_1 &= a_2 \lambda_1 + \ldots + a_o \lambda_{n_o-1} \\ &\vdots \\ a_{n_o-1} &= a_o \lambda_1 + \ldots + a_{n_o-2} \lambda_{n_o-1} \end{aligned}\right\} \qquad \ldots (16.4)$$

where $\lambda_r = \dfrac{\gamma(I_r)}{1-\gamma(I_o)}$, $r = 1, \ldots n_o-1$.

Clearly $\sum_{r=1}^{n_o-1} \lambda_r = 1$.

Let $r_1, \ldots r_\ell$ be all the indices r such that $\lambda_r > 0$.
Let S be the semigroup generated by $r_1, \ldots r_\ell$.
If k is such that $a_k = \max(a_o, a_1, \ldots a_{n_o-1})$, then it easily follows from (16.4) and the fact that $\sum_{r=1}^{n_o-1} \lambda_r = 1$ that,

$$a_k = a_{(k+s)(\text{mod } n_o)} \quad \text{for all } s \in S.$$

However S (mod n_o) is a subgroup of $\{0, 1, \ldots, n_o-1\}$ under addition modulo n_o. By definition of S, S contains positive integers strictly less than n_o. Let n_1 be the least positive integer in S (mod n_o). Then it follows that
$a_k = a_{(k+Mn_1)(\text{mod } n_o)}$ for all positive integers M and also that n_1 divides n_o. A similar argument applied to the

equations among (16.4) which do not involve

a_k , a_{k+n_1}, a_{k+2n_1},..., shows that f is of period n_1.

Hence the proposition is proved.

Corollary 16.3. If γ is a finitely additive probability on all subsets of I such that for every proper subgroup J of I , $\gamma(J) < 1$, then every bounded periodic solution of (16.1) is a constant.

Proof. Follows by repeated application of proposition 16.2.

Corollary 16.4. If γ is a probability on all subsets of I such that for every proper subgroup J of I, $\gamma(J) < 1$ and if $\gamma(\{n_o\}) > 0$ for some $n_o \varepsilon I$, then every bounded solution of (16.1) is a constant.

Proof. If follows immediately from proposition 16.1 that every bounded solution of (16.1) has to be of period $|n_o|$ and now corollary 16.3 gives us the required result.

Definition. A probability γ defined on all subsets of I is called translation invariant if $\gamma(E) = \gamma(E+1)$ for all subsets E of I where $E + 1 = \{i \varepsilon I : i = j + 1 \text{ for some } j \varepsilon E\}$.

Proposition 16.5. If $\gamma = \lambda \gamma_1 + (1-\lambda) \gamma_2$ where $0 \leq \lambda < 1$, γ_2 is a translation invariant probability and γ_1 any finitely additive probability, then any bounded solution of (16.1) is a constant.

<u>Proof</u>. Let f be a bounded solution of (16.1). Since γ_2 is translation invariant, $\int f(i+j)\, d\gamma_2(j) = K$ independent of i. Let $g(i) = f(i) - K$, $i \in I$.

Clearly g is also a solution of (16.1).

Further $g(i) = \lambda \int g(i+j)\, d\gamma_1(j) + (1-\lambda) \int g(i+j)\, d\gamma_2(j)$

$$= \lambda \int g(i+j)\, d\gamma_1(j) \quad \text{for all } i \in I.$$

Since $\lambda < 1$ and g is bounded, this implies that $g(i) = 0$ for all $i \in I$. Therefore $f(i) = K$ for all $i \in I$ which proves the proposition.

We now apply the above results to random walks induced by measures γ on all subsets of I.

<u>Definition</u>. The random walk induced by a probability γ on all subsets of I is the Markov strategy with stationary transition probabilities defined by $\sigma(i)(E) = \gamma(\{j: i+j \in E\})$, $E \subseteq I$, $i \in I$, and initial distribution $\sigma_0 = \gamma$.

<u>Theorem 16.6</u>. Let γ be a probability on all subsets of I such that

either i) γ is 0-1 valued,

or ii) γ has a non-trivial translation invariant part

or iii) γ has a non-trivial countably additive part.

Then the random walk σ induced by γ is simply atomic.

Proof. If i) holds, it is evident from the way the measure σ is defined on $\underline{\underline{B}}$ that σ is 0-1 valued on $\underline{\underline{B}}$, hence on $\underline{\underline{I}}$, i.e. the random walk is simply atomic.

Observe that the equations (15.1) for these random walks actually reduce to (16.1). Hence if ii) holds, by proposition 16.5, all bounded solutions of (15.1) are constants, consequently by theorem 15.12, the random walk is simply atomic.

If iii) holds, $\gamma(\{n_0\}) > 0$ for some $n_0 \in I$. Therefore there exists a smallest subgroup J of I such that $\gamma(J) = 1$. (This J has to include n_0). If $J = I$, by corollary 16.4, all bounded solutions of (15.1) are constants, and so, by theorem 15.12, we are done. If $J \subsetneqq I$, then by working with J and the measure γ restricted to J, one can show as shown in corollary 16.4, using theorem 15.12, that σ is 0-1 valued on any invariant subset of J^N. However by Theorem 4.1, $\sigma(J^N) = 1$. Therefore σ is 0-1 valued on $\underline{\underline{I}}$. Consequently the random walk is simply atomic.

We are now in a position to draw conclusions about the measure of certain important sets in $\underline{\underline{B}}$ under the i.i.d strategy ρ induced by a measure γ on all subsets of I.

Consider the mapping Ψ on $H(=I^N) \longrightarrow H$ defined by $(h_1, h_2, \ldots, h_n, \ldots) \longmapsto (h_1, h_1+h_2, \ldots, h_1+h_2+\ldots+h_n, \ldots)$.

Clearly Ψ is 1-1, onto and a homeomorphism.

Lemma 16.7. For each $i \in I$ and every clopen K in H, $\rho([\Psi^{-1}(K)]i) = \sigma[i](Ki)$, where ρ is the i.i.d strategy induced by a probability γ on all subsets of I and σ the random walk induced by γ.

Proof. If K is H or empty, the result is immediate. Assume the result is true for all i whenever K is a clopen set of structure less than α. Let K now be a clopen set of structure α.

$$\begin{aligned}
\rho([\Psi^{-1}(K)]i) &= \int \rho([\Psi^{-1}(K)]i,j)\, d\gamma(j) \\
&= \int \rho([\Psi^{-1}(Ki)]i+j)\, d\gamma(j) \\
&= \int \sigma[i+j](Ki,i+j)\, d\gamma(j), \text{ by the induction hypothesis,} \\
&= \int \sigma[j](Ki,j)\, d\sigma(i)(j), \text{ by change of variable} \\
&= \sigma[i](Ki).
\end{aligned}$$

Hence the lemma is proved by induction on the structure of the clopen set.

Corollary 16.8. For every clopen K in H,

$$\sigma(K) = \rho(\Psi^{-1}(K)).$$

Proof. $\sigma(K) = \int \sigma[i](Ki)\,d\gamma(i) = \int \rho([\Psi^{-1}(K)]i)\,d\gamma(i)$
$= \rho(\Psi^{-1}(K))$.

Proposition 16.9. For every $B \in \underline{\underline{B}}$, $\sigma(B) = \rho(\Psi^{-1}(B))$.

Proof. Since Ψ is a homeomorphism on H and σ, ρ are strategic and since the result is true for clopen sets (corollary 16.8), proposition 6.2 applies and we get the required result.

Theorem 16.10. Let γ be a probability on all subsets of I such that either i) γ is 0-1 valued

or ii) γ has a non-trivial translation invariant part

or iii) γ has a non-trivial countably additive part.

Let ρ be the i.i.d strategy corresponding to γ.

Then $\rho(\{h : h_1 + \ldots + h_n \in E$ for infinitely many $n\}) = 0$ or 1 for all $E \subseteq I$.

Proof. By theorem 16.6, the random walk σ induced by γ is atomic. Therefore $\sigma(\{E \text{ i.o}\}) = 0$ or 1 for all $E \subseteq I$.

Now proposition 16.9 completes the proof because

$$\Psi^{-1}\{E \text{ i.o}\} = \{h : h_1 + \ldots + h_n \in E \text{ for infinitely many } n\}.$$

Remark 1. If γ is a measure such that $\gamma(J) = 1$ for some proper subgroup J of I, f defined by $f(i) = 1$ for $i \in J$ and $f(i) = 0$, $i \notin J$ is a solution of (16.1) which is not constant.

If γ is such that $\gamma(J) < 1$ for all proper subgroups J of I, then all bounded solutions of (16.1) are constant if γ has a non-trivial countably additive part or a non-trivial translation invariant part (proposition 16.5 and corollary 16.4). However the answer for other γ is not known. Corollary 16.3 tells us that for γ such that $\gamma(J) < 1$ for all proper subgroups J of I, no bounded non-constant solution of (16.1) can be periodic.

Remark 2. Theorem 16.6 shows that for γ either 0-1 valued or having a non-trivial countably additive part or translation invariant part, the corresponding random walk is simply atomic. The answer is not known for general γ.

17. Zero-One laws for Exchangeable sets.

Let γ be a probability on all subsets of I. Let ρ be the i.i.d strategy induced by γ. On I^N consider the group $\underline{\underline{G}}$ of permutations of finitely many coordinates.

Definition. A set $A \in \underline{\underline{B}}$ is said to be exchangeable if $A = \pi(A)$ for all $\pi \in \underline{\underline{G}}$.

Let $\underline{\underline{E}}$ be the σ-field of exchangeable sets in $\underline{\underline{B}}$.

Definition. ρ is exchangeable if $\rho(A) = \rho(\pi(A))$ for all $\pi \in \underline{\underline{G}}$ and for all $A \in \underline{\underline{B}}$.

In case γ is countably additive, the Hewitt-Savage zero - one law states that $\rho(A) = 0$ or 1 for all $A \varepsilon \underline{\underline{E}}$. A fact which is used in the usual proof of this result is that, in case γ is countably additive, ρ is exchangeable. The following theorem shows that this result is far from true for general γ.

Theorem 17.1. Let I be countable and γ a probability defined on all subsets of I. If ρ is the corresponding i.i.d strategy, ρ is exchangeable if and only if γ is countably additive.

Proof. We shall prove only the 'only if' part since the 'if' part is wellknown [3].

Assume without loss of generality that $I=N$. Let $\gamma = \lambda \mu_1 + (1-\lambda)\mu_2$ where μ_1 is a countably additive probability and μ_2 a purely finitely additive probability and $0 \leq \lambda \leq 1$.

Let $A = \{h \varepsilon H : h_1 < h_2\}$.

Then $$\rho(A) = \lambda^2 \int \mu_1\{n : n>i\}\, d\mu_1(i) + \lambda(1-\lambda) \int \mu_2\{n : n>i\}\, d\mu_1(i)$$
$$+\lambda(1-\lambda) \int \mu_1\{n:n>i\}\, d\mu_2(i) + (1-\lambda)^2 \int \mu_2\{n:n>i\}\, d\mu_2(i)$$
$$= \lambda^2 \int \mu_1\{n:n>i\} d\mu_1(i) + \lambda(1-\lambda) + (1-\lambda)^2 \qquad \ldots.(17.1)$$

(The third term is zero since $\mu_1\{n: n > i\} \longrightarrow 0$ as $i \longrightarrow \infty$ and because μ_2 is purely finitely additive).

On the other hand, if π is the permutation which permutes the first two coordinates, then

$$\rho(\pi(A)) = \lambda^2 \int \mu_1\{n: n<i\}\, d\mu_1(i) + \lambda(1-\lambda) \int \mu_2\{n: n<i\} d\mu_1(i)$$
$$+ \lambda(1-\lambda) \int \mu_1\{n: n<i\}\, d\mu_2(i) + (1-\lambda)^2 \int \mu_2\{n: n<i\} d\mu_2(i)$$
$$= \lambda^2 \int \mu_1 \{n: n<i\}\, d\mu_1(i) + \lambda(1-\lambda) \qquad \ldots.(17.2)$$

(Now the second and fourth terms vanish since μ_2 is purely finitely additive).

The first terms on the right hand side of (17.1) and (17.2) are equal since μ_1 is countably additive. Therefore $\rho(A) = \rho(\pi(A))$ if and only if $(1-\lambda)^2 = 0$ i.e. if and only if $\lambda = 1$, which means that γ is countably additive. The theorem is hence proved.

Because of theorem 17.1, it appears unlikely that the Hewitt-Savage 0-1 law is true for general γ. However 0-1 laws hold for a large class of interesting exchangeable sets.

For all tail sets A, $\rho(A) = 0$ or 1 because of the Kolmogorov 0-1 law proved in [29]. Theorem 16.10 shows that a 0-1 law holds for a large class of exchangeable sets, which are not tail sets. A topological approach to the problem gives us more 0-1 laws.

For the rest of the section ρ on $(H, \underline{\underline{B}})$ is the i.i.d strategic measure corresponding to a probability γ on I.

Definition. For a stop rule s, the s-shift T_s is the mapping on H defined by $T_s(h) = (h_{s(h)+1}, h_{s(h)+2}, \ldots)$.
Clearly, for all $A \in \underline{\underline{B}}$, $T_s^{-1}(A) = \bigcup_{h \in H} p_s(h)A$,
where $p_s(h)A = \{h' \in H : h' = p_s(h) h'' \text{ for some } h'' \in A\}$.
The following lemmas study properties of the s-shift.

Lemma 17.2. For every $h \in H$ and $A \in \underline{\underline{B}}$, $[T_s^{-1}(A)]p_s(h) = A$.

Proof. Let $h' \in A$. Then $T_s(p_s(h) h') = h'$. Therefore $p_s(h) h' \in T_s^{-1}(A)$. Consequently, $h' \in [T_s^{-1}(A)]p_s(h)$.
On the other hand, if $h' \in [T_s^{-1}(A)]p_s(h)$, then $p_s(h) h' \in T_s^{-1}(A)$, therefore $T_s(p_s(h)h') \in A$ i.e. $h' \in A$.
Therefore $A = [T_s^{-1}(A)]p_s(h)$.

Lemma 17.3. For every $A \in \underline{\underline{B}}$, $\rho(T_s^{-1}(A)) = \rho(A)$.

Proof. $\rho(T_s^{-1}(A)) = \int \rho[(T_s^{-1}(A))p_s(h)]d\,\rho(h)$

$= \int \rho(A)\, d\,\rho(h)$ by lemma 17.2

$= \rho(A)$.

Lemma 17.4. Let $A \in \underline{\underline{E}}$. Suppose K is a clopen set such that $K \subseteq A$. Then for every stop rule s, $T_s^{-1}(K) \subseteq A$.

Proof. Let K be determined by the stop rule r.
Let $h \in T_s^{-1}(K)$. Then $h = p_s(h) h'$ where $h' \in K$. Since K is determined by r, $p_r(h')h^* \in K$ for all $h^* \in H$. In particular $p_r(h') p_s(h) h'' \in K$ where h'' is such that $h' = p_r(h') h''$.

Consequently $p_r(h') p_s(h)h'' \varepsilon A$. Since A is exchangeable, it follows that $p_s(h) p_r(h') h'' \varepsilon A$, i.e. $h \varepsilon A$. Therefore $T_s^{-1}(K) \subseteq A$.

<u>Lemma 17.5</u>. Let K be a clopen set determined by stop rule r. Then for every stop rule s, $T_s^{-1}(K)$ is a clopen set determined by stop rule $s * r$ defined in section 3.

<u>Proof</u>. Let $h \varepsilon T_s^{-1}(K)$ and suppose $h = p_s(h) p_r(h')h''$ Therefore $p_r(h')h'' \varepsilon K$. If $\hat{h}$ agrees with h in the first $s * r$ coordinates then $\hat{h} = p_s(h) p_r(h')h^*$ for some $h^* \varepsilon H$. Therefore $T_s(\hat{h}) = p_r(h')h^* \varepsilon K$ since $p_r(h')h'' \varepsilon K$ and K is determined by r. Therefore $\hat{h} \varepsilon T_s^{-1}(K)$. Consequentely $T_s^{-1}(K)$ is determined by $s * r$.

<u>Lemma 17.6</u>. Let K be a clopen set determined by stop rule s.

Then $\rho(K \cap T_s^{-1}(K)) = [\rho(K)]^2$

Consequently, $\rho(K \cup T_s^{-1}(K)) = 1 - [1 - \rho(K)]^2$

<u>Proof</u>. $$\begin{aligned}\rho(K \cap T_s^{-1}(K)) &= \int \rho([K \cap T_s^{-1}(K)]p_s(h))\, d\rho(h)\\ &= \int_K \rho([T_s^{-1}(K)]p_s(h)\, d\rho(h)\\ &= \int_K \rho(K)\, d\rho(h) \quad \text{by lemma 17.2}\\ &= [\rho(K)]^2.\end{aligned}$$

Now $\rho(K \cup T_s^{-1}(K)) = \rho(K) + \rho(T_s^{-1}(K)) - \rho(K \cap T_s^{-1}(K))$

$$= 2\rho(K) - [\rho(K)]^2$$

$(\rho(T_s^{-1}(K)) = \rho(K)$ by lemma 17.3)

$$= 1 - [1 - \rho(K)]^2.$$

Theorem 17.7. Let ρ be an i.i.d strategic measure on $(H, \underline{\underline{B}})$. Let $G \in \underline{\underline{E}}$ be open, where $\underline{\underline{E}}$ is the exchangeable σ-field. Then $\rho(G) = 0$ or 1.

Proof. Since G is open, by inner regularity of ρ

$\rho(G) = \text{Sup}\{\rho(K) : K \subseteq G\}$. If $\rho(K) = 0$ for all $K \subseteq G$, $\rho(G) = 0$.

Suppose, on the other hand, there exists K_1 clopen, $K_1 \subseteq G$ such that $\rho(K_1) = \alpha > 0$ and suppose K_1 is determined by stop rule s. Then $T_s^{-1}(K_1) \subseteq G$ by lemma 17.4 and therefore $K_2 = K_1 \cup T_s^{-1}(K_1) \subseteq G$ is a clopen set determined by $s * s$ by lemma 17.5 and $\rho(K_2) = 1 - (1 - \alpha)^2$ by lemma 17.6. Let $K_3 = K_2 \cup T_{s*s}^{-1}(K_2)$. Then $K_3 \subseteq G$ and $\rho(K_3) = 1-(1-\alpha)^4$. Proceeding thus, we get for each n, a clopen set K_n, $K_n \subseteq G$ such that $\rho(K_n) = 1-(1-\alpha)^{2^{n-1}}$. Since $\alpha > 0$, $\rho(K_n) \uparrow 1$ as $n \longrightarrow \infty$. Therefore $\rho(G) = 1$ and the proof of the theorem is complete.

<u>Theorem 17.8</u>. Let ρ be an i.i.d strategic measure on $(H,\underline{\underline{B}})$. Let $G \in \underline{\underline{E}}$ be such that $G = \bigcap_{n=1}^{\infty} G_n$ where for each $n \in N$, G_n is open and $G_n \in \underline{\underline{E}}$. (Consequently G is a G_δ). Then $\rho(G) = 0$ or 1.

<u>Proof</u>. By theorem 17.7, $\rho(G_n) = 0$ or 1 for all $n \in N$. If $\rho(G_n) = 0$ for some n, $\rho(G) = 0$. Assume therefore that $\rho(G_n) = 1$ for all n. Let $\varepsilon > 0$. Let $\{\varepsilon_n\}$ be a sequence of positive real numbers such that $\prod_{n=1}^{\infty} (1-\varepsilon_n) \geq 1-\varepsilon$. For $n \in N$, get a clopen set K_n such that $K_n \subseteq G_n$ and $\rho(K_n) \geq 1 - \varepsilon_n$. This can be done because of inner regularity. Suppose K_n is determined by stop rule s_n.

Set $L_1 = K_1$ and $t_1 = s_1$. Clearly $\rho(L_1) \geq 1 - \varepsilon_1$ and L_1 is determined by t_1. Also $L_1 \subseteq G_1$.

For $n \geq 2$, set $L_n = T^{-1}_{t_{n-1}}(K_n)$ and $t_n = t_{n-1} * s_n$.

Then $\rho[(L_n)p_{t_{n-1}}(h)] = \rho(K_n) \geq 1-\varepsilon_n$ for all $h \in H$,

and L_n is determined by t_n. Also $L_n \subseteq G_n$.

Therefore by (V) of section 1, we have

$$\rho(\bigcap_{n=1}^{\infty} L_n) \geq \prod_{n=1}^{\infty} (1 - \varepsilon_n).$$

Therefore, $\rho(G) = \rho(\bigcap_{n=1}^{\infty} G_n) \geq \rho(\bigcap_{n=1}^{\infty} L_n) \geq \prod_{n=1}^{\infty} (1-\varepsilon_n) \geq 1-\varepsilon$.

As ε is arbitrary, $\rho(G) = 1$ and hence the theorem is proved.

We conclude this section with the following three remarks.

Remark 1. If $G \varepsilon \underline{\underline{E}}$ and $\varepsilon > 0$, then by outer regularity there exists an open set $U_{\varepsilon} \supseteq G$ such that $\rho(U_{\varepsilon} - G) < \varepsilon$. Set $G_{\varepsilon} = \bigcap_{\pi \varepsilon \underline{\underline{G}}} \pi(U_{\varepsilon})$. Then $G_{\varepsilon} \varepsilon \underline{\underline{E}}$ and $\rho(G_{\varepsilon} - G) < \varepsilon$. Further $\underline{\underline{G}}$ being countable, G_{ε} is a G_{δ}. Let $G^{*} = \bigcap_{n=1}^{\infty} G_{1/n}$. Then i) $G^{*} \varepsilon \underline{\underline{E}}$ ii) $G^{*} \supseteq G$ iii) G^{*} is a G_{δ} and $\rho(G^{*} - G) = 0$. Consequently in order to prove a Hewitt-Savage 0-1 law for the finitely additive case, it is enough to prove it for G_{δ}'s in $\underline{\underline{E}}$.

Remark 2. Theorem 17.8 gives us a 0-1 law for some special G_{δ}'s in $\underline{\underline{E}}$, namely G_{δ}'s which can be expressed as a countable intersection of exchangeable open sets. Not every exchangeable G_{δ} is of this form. For example take $G = [\bigcup_{\pi \varepsilon \underline{\underline{G}}} \{\pi(h)\}]^{c}$ for a fixed $h \varepsilon H$ such that G is infinite. Clearly $G \varepsilon \underline{\underline{E}}$ and is a G_{δ} (G^{c} is a countable set and is the $\underline{\underline{E}}$-atom containing h. G cannot be written as $\bigcap_{n=1}^{\infty} G_{n}$ where G_{n}'s are exchangeable and open, for if it could, then G^{c} would be an F_{σ} which could be expressed as countable union of closed exchangeable sets. This is not possible since the only non-empty exchangeable subset of G^{c} is itself, and G^{c} is an F_{σ} which is not closed.

Remark 3. The exchangeable sets of the form $\{h : h_1+\dots+h_n \in E \text{ i.o}\}$, $E \subseteq I$, are G_δ exchangeable sets but not necessarily a countable intersection of exchangeable open sets.

18. Potentials. Let $\{\sigma(i)\}_{i\in I}$ be a Markov strategy. Let f be a real valued function on I. For $i \in I$, by the coordinate process $\{Y_n^i\}_{n\geq 0}$ induced by f, we shall mean the sequence of random variables defined by $Y_0^i \equiv f(i)$ and $Y_n^i(h) = f(h_n)$, $h \in H$, $n \geq 1$.

Definition. A bounded, nonnegative function f on I is called a superharmonic function if

$$f(i) \geq \int f(j)\, d\, \sigma(i)(j) \quad \text{for all } i \in I \qquad \dots(18.1)$$

If equality holds in (18.1) for all $i \in I$, then f is a harmonic function.

Definition. A bounded nonnegative function f on I is called a potential if f is a superharmonic function and, $0 \leq g(i) \leq f(i)$ for all $i \in I$ and g harmonic imply that $g(i) = 0$ for all $i \in I$.

Lemma 18.1. A bounded, nonnegative function f on I is superharmonic if and only if for each $i \in I$, the coordinate process $\{Y_n^i\}_{n\geq 0}$ induced by f, forms a supermartingale with respect to $(H, \underline{B}, \sigma[i])$.

(See section 15 for definition of supermartingale).

Proof. If for each i, $\{Y_n^i\}_{n\geq 0}$ forms a supermartingale, then $\int f(j)d\sigma(i)(j) = \int Y_1^i \, d\,\sigma[i] \leq Y_0^i = f(i)$ for each $i \varepsilon I$. Consequently f is superharmonic.

On the other hand if f is superharmonic, then for each i,

$$\int Y_1^i \, d\sigma(i) = \int f(j)d\sigma(i)(j) \leq f(i) = Y_0^i .$$

Also for each $n \varepsilon N$ and each finite sequence $(i_1,\ldots,i_n)$,

$$\int (Y_{n+1}^i \; i_1,\ldots,i_n) d\sigma[i,i_1,\ldots,i_n] = \int f(j)d\sigma(i_n)(j)$$

$$\leq f(i_n) = Y_n^i \; i_1,\ldots,i_n .$$

Therefore $\{Y_n^i\}_{n\geq 0}$ is a supermartingale.

Lemma 18.2. Let f be a superharmonic function. If $f_1(i) = \int \emptyset \, d\,\sigma[i]$, $i \varepsilon I$, where $\emptyset(h) = \limsup_n f(h_n)$, $h \varepsilon H$, then $0 \leq f_1(i) \leq f(i)$ for all $i \varepsilon I$ and f_1 is a harmonic function. Further if f is harmonic, then $f_1(i) = f(i)$ for all $i \varepsilon I$.

Proof. We first note that since $\emptyset$ is a bounded, nonnegative invariant function, by theorem 15.10, f_1 is harmonic. Also, by theorem 15.6, for each $i \varepsilon I$,

$$f_1(i) = \int \emptyset \, d\,\sigma[i] = \operatorname{Limsup}_s \int f(h_s) \, d\,\sigma[i](h) ,$$

where the limsup is taken over stop rules. We can show just as in the proof of the second part of theorem 15.10, by induction on the structure of stop rules, that for each $i \in I$ and each stop rule s, $\int f(h_s) d\,\sigma[i](h) \leq f(i)$ (equality holds if f is harmonic). Hence the result follows.

Lemma 18.3. A superharmonic function f on I is a potential if and only if $f_1(i) = 0$ for all $i \in I$, where f_1 is the function defined in Lemma 18.2.

Proof. Suppose that f is a potential. By lemma 18.2, f_1 is harmonic and $0 \leq f_1 \leq f$. Consequently, by definition of a potential, $f_1(i) = 0$ for all $i \in I$.

Conversely if $f_1(i) = 0$ for all $i \in I$ and g is a harmonic function such that $0 \leq g(i) \leq f(i)$ for all $i \in I$, then for each i,

$$0 \leq \int \operatorname{Lim\,sup}_n \; g(h_n) d\sigma[i] \leq \int \operatorname{Lim\,sup}_n \; f(h_n) d\sigma[i] = f_1(i).$$

But since g is harmonic, by lemma 18.1,

$$\int \operatorname{Lim\,sup}_n \; g(h_n)\, d\,\sigma[i] = g(i).$$

Therefore $0 \leq g(i) \leq f_1(i) = 0$. Consequently f is a potential.

Theorem 18.4 (Riesz decomposition theorem). Let $\{\sigma(i)\}_{i\in I}$ be Markov strategy. Let f be superharmonic. Then f can be niquely expressed as the sum of a non-negative harmonic function

f_1 and a potential f_2. Further if $e(i) = f(i) - \int f(j) d\sigma(i)(j)$, $i \in I$ (e is called the excess), then the potential f_2 of f is given by

$$f_2(i) = e(i) + \lim_s \widehat{\sigma[i]}\,(e(h_1) + \ldots + e(h_s))\ ,\ i \in I\ ,$$

where $\widehat{\sigma[i]}(\cdot)$ denotes the Dubins - Savage integral with respect to the strategic measure $\sigma[i]$, and the limit on the right side is taken over the net of stop rules.

Proof. The function f_1 defined as in lemma 18.2 is a harmonic function and $0 \leq f_1 \leq f$, by lemma 18.2.

Let $f_2 = f - f_1$. Then for each $i \in I$,

$$\int \limsup_n f_2(h_n) d\sigma[i] = \int \limsup_n f(h_n) d\sigma[i]$$
$$- \int \limsup_n f_1(h_n) d\sigma[i]$$

because, by the supermartingale convergence theorem, the three lim sup's are $\sigma[i]$ a.e. equal to the limits.

The right side equals $f_1(i) - f_1(i) = 0$ by lemma 18.2. Therefore by lemma 18.3, f_2 is a potential.

We shall now prove uniqueness of the decomposition. Let $f = f_1' + f_2'$ where f_1' is harmonic and f_2' a potential.

Then $f_1(i) = \int \lim_n f(h_n) d\sigma[i] = \int \lim_n f_1'(h_n) d\sigma[i]$

$$+ \int \lim_n f_2'(h_n) d\sigma[i]$$

$$= f_1'(i) + 0$$

by lemmas 18.3 and 18.2.

Therefore $f_1 = f_1'$ and hence $f_2 = f_2'$.

For the last part, it is enough to prove that for each stop rule s and each $i \in I$,

$$e(i) + \widehat{\sigma[i]}(e(h_1)+\ldots+e(h_s)) = f(i) - \int f(h_{s+1}) d\sigma[i](h) \quad \ldots.(18.3)$$

By taking limits on either side of the above expression over the net of stop rules and using theorems 15.β and 15.δ, the result would follow.

We prove (18.3) by induction on the structure of stop rules. For $s \equiv 1$, (18.3) follows because

$$e(i) + \widehat{\sigma[i]}(e(h_1)) = e(i) + \int \left\{ f(j) - \int f(k) d\sigma(j)(k) \right\} d\sigma(i)(j)$$

$$= f(i) - \int \left\{ \int f(k) d\sigma(j)(k) \right\} d\sigma(i)(j)$$

$$= f(i) - \int f(h_2)\, d\, \sigma[i](h).$$

Also for $n \in N$,

$$e(i) + \widehat{\sigma[i]}(e(h_1) + \ldots + e(h_{n+1}))$$

$$= e(i) + \int [e(j) + \widehat{\sigma[j]}(e(h_1) + \ldots + e(h_n))] \, d\,\sigma(i)(j)$$

(by the property of the Dubins- Savage integral)

$$= e(i) + \int [f(j) - \int f(h_{n+1}) d\sigma[j](h)] \, d\,\sigma(i)(j)$$

if (18.3) holds for $s \equiv n$,

$$= f(i) - \int f(h_{n+2}) \, d\,\sigma[i](h) .$$

Thus by ordinary induction on n, (18.3) follows for all stop rules of structure zero. Assume that (18.3) holds for all stop rules of structure $< \alpha$. Let s now be a stop rule of structure α. By the property of the Dubins - Savage integral

$$e(i) + \widehat{\sigma[i]}(e(h_1) + \ldots + e(h_s))$$

$$= e(i) + \int [e(j) + \widehat{\sigma[j]}(g_j)] \, d\,\sigma(i)(j) \quad \ldots .(18.4)$$

where g_j is the function on H defined by $g_j(h') = 0$ for all $h' \in H$ if $s(jh') = 1$ for all $h' \in H$, and if $s(jh') > 1$ for all $h' \in H$, $g_j(h') = e(h'_1) + \ldots + e(h'_{s_j})$, where s_j is the stop rule defined by $s_j(h') = s(jh') - 1$ for all $h' \in H$. Clearly s_j has structure less than α. By the induction hypothesis, (18.4) reduces to

$$e(i) + \widehat{\sigma[i]}(e(h_1) + \ldots + e(h_s)) = f(i) - \int f(h_{s+1}) d\sigma[i](h) .$$

Hence (18.3) is proved for all stop rules s and the proof of the theorem is complete.

Theorem 18.5. For a subset E of I, the function f_E defined by

$$f_E(i) = \sigma[i](\{h: h_n \in E \text{ for some } n \in N\}) \text{ if } i \notin E$$
$$= 1 \text{ otherwise,}$$

is a superharmonic function with values between 0 and 1. The harmonic component of its Riesz decomposition, f_{E1}, is equal to the probability of visiting E infinitely often, i.e

$$f_{E1}(i) = \sigma[i](\{E \text{ i.o}\}) \text{ for all } i \in I.$$

Finally, the excess e_E of f_E is given by

$$e_E(i) = \begin{cases} \sigma[i](\{h: h_n \notin E, \text{ for each } n \in N\}) & \text{if } i \in E \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Observe that by conditioning at the first coordinate, we get for any $i \in I$,

$$\sigma[i](\{h: h_n \in E \text{ for at least one } n \in N\}) = \int f_E(j)d\sigma(i)(j) \quad \text{....(18.5)}$$

Therefore if $i \notin E$, $f_E(i) = \int f_E(j)\, d\,\sigma(i)(j)$ ....(18.6)

and if $i \in E$, $1 = f_E(i) \geq \int f_E(j)\, d\,\sigma(i)(j)$.

Therefore f_E is superharmonic.

Further from (18.5) and (18.6), it also follows that

$$e_E(i) = 0 \quad \text{if} \quad i \notin E$$

$$= \sigma[i](\{ h_n \notin E \text{ for } n \geq 1 \}) \quad \text{if} \quad i \in E .$$

By theorem 18.4, $f_{E1}(i) = \int \lim_n f_E(h_n) \, d\, \sigma[i]$ for each i.

Therefore by theorem 15.δ,

$$f_{E1}(i) = \lim_s \int f_E(h_s) \, d\, \sigma[i]$$

$$= \lim_s \sigma[i] (\{h: h_n \in E \text{ for some } n \geq s(h)\})$$

$$= \sigma[i] (\{ E \text{ i.o} \}).$$

This completes the proof of the theorem.

We can now characterize transient sets in terms of potentials. Recall that E is transient if $\sigma[i](\{ E \text{ i.o} \}) = 0$ for all $i \in I$.

<u>Remark</u>. If $E = \{ i \}$, then $\{ i \}$ is transient if and only if i is transient.

<u>Corollary 18.6</u>. A subset E of I is transient if and only if there exists a potential g such that $g(i) \geq 1$ for all $i \in E$.

<u>Proof</u>. 'If part'. By lemma 18.3, we have

$$\int \limsup_n g(h_n) \, d\, \sigma[i](h) = 0 \quad \text{for each} \quad i \in E.$$

On $\{E \text{ i.o}\}$, $\limsup_n g(h_n) \geq 1$, therefore

$$\sigma[i](\{E \text{ i.o}\}) = 0 \text{ for all } i \in I, \text{ i.e. } E \text{ is}$$

transient.

'Only if'. If E is transient, the harmonic component of f_E , by theorem 18.5 is zero. Consequently f_E is a potential. Further f_E is equal to one on E by definition.

<u>Notation</u>. If $E = \{i\}$, we shall denote $f_{\{i\}}$ by f_i .

<u>Corollary 18.7</u>. For each $i \in I$, the superharmonic function f_i is either a potential or a harmonic function. In the first case i is transient and in the second case i is recurrent.

<u>Proof</u>. If f_i is not a potential, the harmonic part of its Riesz decomposition is non-trivial, therefore by theorem 18.5, $\sigma[j](G_i) = g_{ji} > 0$ for some $j \in I$. Since $G_i \in \underline{I}$, the invariant σ-field, by the Levy 0-1 law ,

$$\sigma[h_n](G_i) \longrightarrow 1_{G_i}, \sigma[j] \text{ a.e.. Since } \sigma[j](G_i) > 0,$$

there exists $h \in G_i$ for which $\lim_n \sigma[h_n](G_i) = 1$, but this limit has to be equal to $\sigma[i](G_i) = g_{ii}$. Therefore $g_{ii} = 1$, hence by corollary 5.3, i is recurrent. Therefore $e_i(i) = 1-1 = 0$. Consequently by theorem 18.5, $e_i(j) = 0$ for all $j \in I$, i.e. f is harmonic.

To complete the proof, we only need to observe that if f_i is a potential, then $\sigma[i](G_i) = 0$ by theorem 18.5, hence by corollary 5.3, i is transient.

We conclude this section with a characterization of recurrent chains in terms of superharmonic functions.

Theorem 18.8. Let $\{\sigma(i)\}_{i \varepsilon I}$ be a Markov Strategy. Then I is a weakly communicating recurrent class if and only if the only superharmonic functions are constants.

Proof. 'if part'. In this case for each $i \varepsilon I$, f_i is a constant, consequently harmonic, and therefore, by corollary 18.7, i is recurrent. Also since $f_i(i) = 1$, f_i is identically equal to 1 for each i, therefore $i \xrightarrow{W} j$ for all $i, j \varepsilon I$. Consequently I is a recurrent class.

'Only if part'. Let f be a superharmonic function. Let $i, j \varepsilon I$. Since I is a weakly communicating recurrent class, $\sigma[i](G_i \cap G_j) = 1$. Also by the Supermartingale convergence theorem, on a set of $\sigma[i]$-measure one, $f(h_n)$ converges. Consequently there exists $h \varepsilon G_i \cap G_j$ for which $f(h_n)$ converges. Since $h \varepsilon G_i \cap G_j$, this limit has to be equal to $f(i)$ as well as $f(j)$. Therefore $f(i) = f(j)$ for all $i, j \varepsilon I$. Consequently f is a constant.

REFERENCES

[1] D. BLACKWELL , On transient Markov Processes with a countable number of states and stationary transition probabilities, Ann. Math. Statist. 26 (1955), 654-658.

[2] L. BREIMAN , Transient atomic Markov chains with a denumerable number of states, Ann. Math. Statist. 29 (1958) , 212-218.

[3] L. BREIMAN , Probability, Addison-Wesley, Reading, Mass., 1968.

[4] R. CHEN , On finitely additive almost sure convergence, dissertation, University of Minnesota, 1974.

[5] R. CHEN . A finitely additive version of Kolmogorov's law of iterated logarithm, Israel J. Math. 23 (1976), 209-220.

[6] R. CHEN , Some finitely additive versions of the strong law of large numbers, Israel J. Math. 24 (1976), 244-259.

[7] R. CHEN , On almost sure convergence in a finitely additive setting, Z. Wahrscheinlichkeitstheorie und Verw. Gebeite, 37 (1977), 341-356.

[8] K.L. CHUNG , Markov chains with Stationary Transition Probabilities, 2nd edition, Springer, Berlin, 1967.

[9] K.L. CHUNG and W.H.J. FUCHS , On the distribution of values of sums of random variables, Mem.Amer.Math. Soc., No.6 (1951).

[10] C. DERMAN , A solution to a set of fundamental equations in Markov chains, Proc. Amer. Math. Soc., 5 (1954), 332-334.

[11] J.L. DOOB , Stochastic Processes, Wiley, New York, 1953.

[12] L.E. DUBINS , On Lebesque-like extensions of finitely additive measures. Ann. Probability 2 (1974) , 456-463.

[13] L.E. DUBINS , Finitely additive conditional probabilities, conglomerability and disintegrations, Ann. Probability 3 (1975), 89-99.

[14] L.E. DUBINS and L.J.SAVAGE, How to Gamble if you Must : Inequalities for Stochastic Processes, McGraw-Hill, New York, 1965.

[15] N. DUNFORD and J.T. SCHWARTZ , Linear Operators, Part 1, Interscience, New York, 1958.

[15a] J. FELDMAN , Feller and Martin boundaries for countable sets, Illinois J. Math. 6 (1962), 357-366.

[16] W. FELLER , Boundaries induced by nonnegative matrices, Trans. Amer. Math. Soc. 83 (1956), 19-54.

[17] W. FELLER , An Introduction to Probability theory and its Applications, Vol.2, Wiley, New York, 1966.

[18] W. FELLER , An Introduction to Probability theory and its Applications, Vol.I, 3rd edition, Wiley, New York, 1968.

[19] B. DE FINETTI , Theory of Probability, Vol.I and II, Wiley, New York, 1974.

[20] D. FREEDMAN , Markov Chains, Holden-Day, San Francisco, 1971.

[21] T.E. HARRIS , The existence of stationary measures for certain Markov processes, Proceedings of the third Berkeley symposium, Vol.II, 1956.

[22] T.E. HARRIS , Transient Markov chains with stationary measures, Proc. Amer. Math. Soc.8 (1957), 937-942.

[23] E. HEWITT and L.J. SAVAGE , Symmetric measures on cartesian products, Trans. Amer. Math. Soc. 80 (1955), 470-501.

[24] S. KARLIN , A First course in Stochastic Processes, Academic Press, New York, 1969.

[25] J.G. KEMENEY, J. SNELL and A.W. KNAPP , Denumerable Markov chains, Van Nostrand, Princeton, 1966.

[26] C. KURATOWSKI , Topologie I, 4th edition, Warsaw, 1958.

[27] J. NEVEN , Mathematical Foundations of the Calculus of Probability, Holden-Day, San Francisco, 1965.

[28] J. NEVEN , Discrete-parameter Martingales, North-Holland, Amsterdam, 1975.

[29] R.A. PURVES and W.D. SUDDERTH , Univ. of Minnesota School of statistics, Tech. Report No.220, 1973.

[30] R.A. PURVES and W.D. SUDDERTH , Some finitely additive probability, Ann. Probability, 4 (1976), 259-276.

[31] R. SIKORSKI , Boolean Algebras, 2nd edition, Springer, Berlin, 1964.