Internat. J. Math. & Math. Sci.
Vol. 2 No. 4 (1979) 537-587

ON THE ASYMPTOTIC EVENTS OF A MARKOV CHAIN

HARRY COHN

Department of Statistics
University of Melbourne
Parkville, Victoria 3052
Australia

(Received December 4, 1978)

ABSTRACT. In this paper we investigate some structure properties of the tail σ-field and the invariant σ-field of both homogeneous and nonhomogeneous Markov chains as representations for asymptotic events, descriptions of completely nonatomic and atomic sets and global characterizations of asymptotic σ-fields. It is shown that the Martin boundary theory can provide a unified approach to the asymptotic σ-fields theory.

KEY WORDS AND PHRASES. *Tail σ-field, Invariant σ-field, Atomic Set, Completely Nonatomic Set, Harmonic Function, Space-Time Harmonic Function, Martingale, Martin Boundary, Chacon-Ornstein Ergodic Theorem, 0-2 Laws.*

1980 MATHEMATICS SUBJECT CLASSIFICATION CODES. *60-02, 60J10, 60F99.*

1. INTRODUCTION.

The first result on the asymptotic events of a sequence of random variables was the 0 - 1 law given by Kolmogorov in 1933 [28]. In the years that followed the publication of Kolmogorov's book, the 0 - 1 law was extensively used by P. Levy, W. Feller, etc. to obtain important properties of some variables derived from sequences of independent random variables. It was therefore natural to start the investigation of asymptotic events of other sequences of random variables of interest and the next important results in this respect have been the Hewitt-Savage 0 - 1 law for symmetrical dependent sequences of random variables [23] and Blackwell's characterization of invariant events of Markov chains [4](the invariant events constitute an important class of asymptotic events).

Nowadays, there is a sizeable literature concerning asymptotic events of random variables, especially for Markov chains. It seems to us that the time is ripe for an account of the basic theory of asymptotic events of Markov chains and the main aim of this paper is to attempt such an account.

Some of the approaches and results given here are new, others are extensions of the known results to more general settings, but we shall also present many known results which, in our view, are basic for the asymptotic events theory.

No applications are included here, although the general theory presented draws heavily on ideas and methods occurring in papers dealing with asymptotic events for various types of Markov chains. The applications, which are numerous and important, will be taken up elsewhere.

In the early papers on Potential Theory it has been noticed that any nonhomogeneous Markov chain can be thought of as a homogeneous one in the modified context of a space-time chain. However, such an approach was not often pursued, due probably to the fact that the state space of a space-time chain seemed to be untractable. We intend to show here that, as far as the theory of tail and invariant events is concerned, the space-time approach provides a unified method of dealing with both homogeneous and nonhomogeneous chains.

The connection between the invariant σ-field and the Martin boundary theory which appeared in Blackwell's paper [4] was noticed by Doob [16]. Further Neveu [33] and Jamieson and Orey [27] showed that some notions of Martin boundary theory as space-time harmonic functions can be related to the tail σ-field of a Markov chain. The relation between the Martin boundary theory and asymptotic (tail and invariant) σ-fields will be shown here to go much further and the Martin boundary theory will provide a unified approach to the asymptotic σ-fields theory.

The paper is devided into five chapters. The first chapter is the Introduction. The second chapter introduces some notions related to asymptotic σ- fields and basic properties to be used in the sequel are derived. The third chapter contains representations for asymptotic events and random variables by means of harmonic and space-time harmonic functions, as well as in terms of some almost surely convergent sequences of sets. As consequences, criteria for triviality of asymptotic σ-fields are derived. The fourth chapter investigates the connection between the Martin boundary theory and asymptotic σ-fields. It is shown that the basic almost surely convergence result of the Martin boundary theory implies new results as well as most of the results previously obtained by sundry methods in the asymptotic σ-fields theory. The last chapter contains some structure theorems for asymptotic σ-fields as descriptions of atomic and nonatomic events, and global characterizations of asymptotic σ-fields.

In the choice of the material presented here, I might have been biased by my own interests and research and it is possible that insufficient attention has been paid to some contributions to the asymptotic σ-fields theory. I would like to emphasize that I did not intend to pass judgements on the importance of various contributions to the field and that I am aware of the fact that the present survey reflects my own viewpoint on some topics that have preoccupied me for many years.

2. PRELIMINARY RESULTS.

2.1 DEFINITIONS AND NOTATIONS. Let $(S,\mathcal{B})$ be a measurable space, $N=\{0,1,\dots\}$ and λ,μ two finite measures on $(S,\mathcal{B})$. We shall denote by $\lambda\otimes\mu$ the product measure of λ and μ on $(S\times S,\mathcal{B}\otimes\mathcal{B})$ and by $\|\lambda-\mu\|$ the total variation norm of $\lambda-\mu$, i.e. $\|\lambda-\mu\| = (\lambda-\mu)^+(S) + (\lambda-\mu)^-(S)$, where $(\lambda-\mu)^+$ and $(\lambda-\mu)^-$ are the positive and negative part of $\lambda-\mu$ in its Hahn-Jordan decomposition. In the case $S=N$ and $\mathcal{B} = \mathcal{P}(N)$, where $\mathcal{P}(N)$ is the class of all subsets of N, we get $\|\lambda-\mu\| = \sum_{i\varepsilon N} |\lambda(i)-\mu(i)|$. Further, Λ^c will stand for the complementary set of Λ, $\Lambda_1\Delta\Lambda_2$ for the symmetric difference of Λ_1 and Λ_2, Z for the set of integers and R for the set of real numbers. Two measures λ and μ are called singular with respect to each other (denoted $\lambda\perp\mu$) if there is a set $B\varepsilon\mathcal{B}$ such that $\lambda(B) = 0$ and $\mu(B^c) = 0$; and λ will be said to be absolutely continuous with respect to μ (denoted $\lambda \ll \mu$) if $\mu(B) = 0$ implies $\lambda(B) = 0$. Given μ, every λ can be decomposed into a sum $\lambda_1+\lambda_2$ where $\lambda_1 \ll \mu$ and $\lambda_2\perp\mu$. Moreover, there exists a set H in $\mathcal{B}$, called Hahn set, such that $\lambda(B) = \lambda(B\cap H) + \lambda(B\cap H^c)$ where $\lambda(\cdot\cap H) = \lambda_1(\cdot)$ and $\lambda(\cdot\cap H^c) = \lambda_2(\cdot)$.

A kernel N is a mapping from $S\times\mathcal{B}$ into $(-\infty,\infty]$ such that

(i) for every x in S the mapping $A\to N(x,A)$ (denoted $N(x,\cdot)$) is a measure on $\mathcal{B}$,

(ii) for every A in $\mathcal{B}$ the mapping $x\to N(x,A)$ (denoted $N(\cdot,A)$) is a measurable function with respect to $(S,\mathcal{B})$.

A kernel N is said to be positive if its range is in $[0,\infty]$; it is said to be proper if S is the union of an increasing sequence $\{S_n: n\geq 0\}$ of subsets of S such that $N(\cdot,S_n)$ are bounded. A kernel for which $N(x,S)=1$ for all $x\varepsilon S$ is said to be a transition probability kernel.

Let $\Omega = S\times S\times\dots$, $\mathcal{F} = \mathcal{B}\otimes\mathcal{B}\otimes\dots$ and for each $\omega = (x_0,x_1,\dots,x_n,\dots)\varepsilon\Omega$ let $X_n(\omega) = x_n$. Then it is well known ([15]) that given a probability measure ν

on $\mathcal{B}$ and a sequence of transition probability kernels $(P_n)_{n\varepsilon N}$, there exists a probability P_ν on $\mathcal{F}$ under which the sequence of random variables $\{X_n(\omega):n \geq 0\}$ forms a nonhomogeneous Markov chain and

$$P_\nu(X_o \varepsilon A) = \nu(A)$$
$$P_\nu(X_{n+1} \varepsilon A | X_n = x) = P_n(x,A) \tag{2.1}$$

The probability P_ν is uniquely determined by its finite dimensional marginals defined as follows

$$P_\nu(X_o \varepsilon B_o, X_1 \varepsilon B_1, \ldots, X_k \varepsilon B_k) =$$
$$\int_{B_o} \nu(dx_o) \int_{B_1} P_1(x_o,dx_1) \ldots \int_{B_k} P_k(x_{k-1},dx_k) \ . \tag{2.2}$$

The measure ν is called starting measure or initial probability distribution of the chain. If $\nu = \delta(x)$ where δ stands for the Dirac measure, we shall abbreviate P_x for $P_{\delta(x)}$.

Denote by $\mathcal{F}_n$ the σ-algebra generated by $X_o,\ldots,X_n$ and by $\mathcal{F}^n$ the σ-algebra generated by $X_n, X_{n+1}, \ldots$. The transition probability after n steps $P^{m,m+n}(x,B)$ denotes the probability that X_{m+n} is in B given that $X_m = x$ and is defined inductively as

$$P^{m,m+1}(x,B) = P_m(x,B)$$

$$P^{m,m+n}(x,B) = \int P^{m,m+n-1}(x,dy) P_{m+n-1}(y,B) \ .$$

If $P_n = P$ for all n , $P^{m,m+n}$ depends only on n and will be denoted by $P^{(n)}$. When S is countable we can easily check that $P^{(n)} = P^n$.

$\{X_n : n \geq 0\}$ will be said to be a homogeneous Markov chain or a Markov chain with stationary transition probabilities if for any integers m,n with $m < n$ and for any $B \varepsilon \mathcal{B}$

$$P_\nu(X_n \in B | \mathcal{F}_m) = P^{(n-m)}(X_m, B) \qquad P_\nu \text{ a.s.} \tag{2.3}$$

The σ-field $\mathcal{T} = \bigcap_{n=0}^{\infty} \mathcal{F}^n$ will be said to be the tail σ-field of the chain. A $\mathcal{T}$ measurable random variable (event in $\mathcal{T}$) will be called tail variable (event).

We shall next consider the shift operator θ which maps Ω into Ω by $\theta\omega = \omega'$ where $\omega = (x_0, x_1, \ldots, x_n, \ldots)$ and $\omega' = (x_1, x_2, \ldots, x_{n+1}, \ldots)$. $\theta\Lambda$ will stand for the set $\{\theta\omega : \omega \in \Lambda\}$, $\theta^{-1}\Lambda$ for $\{\omega : \theta\,\omega \in \Lambda\}$ and $\theta^0\Lambda$ for Λ . θ^n will denote the nth iterate of θ . If Y is a function on Ω , θY is defined by $\theta Y(\omega) = Y(\theta\omega)$ and $\theta^n Y$ as its nth iterate. A random variable Y for which $\theta Y(\omega) = Y(\omega)$ for all $\omega \in \Omega$ will be said to be invariant. It is easy to see that $X_n(\theta\omega) = X_{n+1}(\omega), X_n(\theta^p\omega) = X_{n+p}(\omega)$ and $\theta^{-p}\{X_n \in B\} = \{X_{n+p} \in B\}$. A set $\Lambda \in \mathcal{F}$ will be said to be invariant if $\theta^{-1}\Lambda = \Lambda$ The class of all invariant sets, denoted by $\mathcal{J}$ is a σ-field, called the invariant σ-field.

A set Λ in a σ-field $\mathcal{G}$ is called atomic with respect to $\mathcal{G}$ if $P_\nu(\Lambda) > 0$ and Λ does not contain two disjoint subsets of positive probability belonging to $\mathcal{G}$. A set Λ in $\mathcal{G}$ is called completely nonatomic with respect to $\mathcal{G}$ if $P_\nu(\Lambda) > 0$ and Λ does not contain any atomic subsets belonging to $\mathcal{G}$. It is well known (see e.g. [38] p. 81-82) that Ω may be represented as $\Omega = \bigcup_{n=0}^{\infty} \Lambda_n$ where Λ_j's for $j \geq 1$ or Λ_0 may be absent, but, if present, Λ_0 is completely nonatomic and $\Lambda_1, \Lambda_2, \ldots$ are atomic sets with respect to $\mathcal{G}$. If Λ_0 is present we shall say that $\mathcal{G}$ is non-atomic, whereas if Λ_0 is absent we shall say that $\mathcal{G}$ is atomic. Further, $\mathcal{G}$ will be said to be finite if Λ_0 is absent and there are only a finite number of atomic sets. Finally, if $\Lambda_1 = \Omega$, $\mathcal{G}$ will be said to be trivial.

Denote by 1_Λ the indicator of Λ , i.e. the function which takes on the value 1 for $\omega \in \Lambda$ and 0 for $\omega \in \Lambda^c$. We shall say that $\Lambda_1 = \Lambda_2$ a.s. if $1_{\Lambda_1} = 1_{\Lambda_2}$ a.s. and that $\lim_{n\to\infty} \Lambda_n = \Lambda$ a.s. if $\lim_{n\to\infty} 1_{\Lambda_n} = 1_\Lambda$ a.s.

2.2 θ's ACTION ON $\mathcal{T}$. It is easy to see that both θ and θ^{-1} map sets of $\mathfrak{F}$ into sets of $\mathfrak{F}$ and are countably additive. Also, we can easily check that θ^{-1} preserves the disjointness of sets and commutes with complementation and countable intersections. These properties of θ^{-1}, unpossessed by θ, are probably accountable for the use of θ^{-1} in the definition and manipulations involving invariant sets from the very beginning of the ergodic theory. However, in all the examples available, the failure of such properties for θ is due to the relevance of the first coordinate of ω which is removed by the action of θ. We shall now see that if we restrict our attention to the action of θ on the sets of $\mathcal{T}$ we can show that all the properties of θ^{-1} mentioned above are also possessed by θ. We prove first

PROPOSITION 1. θ maps $\mathcal{T}$ one-to-one and onto.

PROOF. In view of the already mentioned properties of θ^{-1} we can easily show that $\theta^{-1}\mathfrak{F}^{n} = \mathfrak{F}^{n+1}$ for all $n \geq 0$. Further for any set $\Lambda \subseteq \Omega$, $\theta(\theta^{-1}\Lambda) = \Lambda$ and so it follows that $\theta\mathfrak{F}^{n+1} = \mathfrak{F}^{n}$ and an upshot of these considerations is $\theta\mathcal{T} = \mathcal{T}$. To show that θ is one-to-one, suppose that for $\Lambda_1, \Lambda_2 \in \mathcal{T}, \theta\Lambda_1 = \theta\Lambda_2$. This means that there exist $\omega_1 \in \Lambda_1$ and $\omega_2 \in \Lambda_2$ such that $\theta\omega_1 = \theta\omega_2$. Because Λ_1 and Λ_2 are both in $\mathcal{T}$ we can assume that $X_o(\omega_1) = X_o(\omega_2)$. Therefore the first coordinates of ω_1 and ω_2 are identical and since $\theta\omega_1 = \theta\omega_2$, so are the other coordinates. We get $\omega_1 = \omega_2$. Thus $\Lambda_1 = \Lambda_2$ and the proof is complete.

PROPOSITION 2. $\mathcal{J} = \{\Lambda \in \mathcal{T} : \theta\Lambda = \Lambda\}$.

PROOF. If $\Lambda \in \mathcal{J}$, then $\Lambda \in \mathfrak{F}^{o}$ and therefore for any integer $n \geq 0$, $\Lambda = \theta^{-n}\Lambda \in \mathfrak{F}^{n}$. Thus $\Lambda \in \mathcal{T}$. Furthermore, $\Lambda = \theta(\theta^{-1}\Lambda) = \theta\Lambda$ which implies $\mathcal{J} \subseteq \{\Lambda \in : \theta\Lambda = \Lambda\}$. The reverse inclusion follows directly from the assertion of Proposition 1 that θ is one-to-one over $\mathcal{T}$.

The above given Propositions 1 and 2 are due to Abrahamse [1].

PROPOSITION 3. Suppose that $\Lambda, \Lambda_1, \Lambda_2, \ldots$ belong to $\mathcal{T}$. Then

(i) $\theta\Lambda^c = (\theta\Lambda)^c$

(ii) $\theta\bigcap_{n=1}^{\infty}\Lambda_n = \bigcap_{n=1}^{\infty}\theta\Lambda_n$

(iii) $\theta^{m+n}\Lambda = \theta^m\theta^n\Lambda$ for $m, n \in Z$

PROOF. By Proposition 1, θ and θ^{-1} are interchangeable, when applied to the sets of $\mathcal{T}$. Let us apply θ^{-1} to both sides of (i); we get $\Lambda^c = \theta^{-1}(\theta\Lambda)^c$. Since θ^{-1} commutes with complementation $\theta^{-1}(\theta\Lambda)^c = \theta^{-1}\theta\Lambda^c = \Lambda^c$ and we got an equality. But $\theta^{-1}\Lambda' = \theta^{-1}\Lambda''$ means $\Lambda' = \Lambda''$ and the proof of (i) is complete. (ii) and (iii) can be proved in the same way.

Proposition 1-3 show that there is no reason to use θ^{-1} instead of θ in the definition of an invariant set. Since the Markov assumption was not used in the above proofs, such an observation holds for the invariant sets of a σ-field generated by an arbitrary sequence of random variables.

2.3 SMALL SETS. A set Λ in $\mathcal{F}$ will be said to be a null set if $P_\nu(\Lambda) = 0$ and positive otherwise. If for all $n \in Z$, $P_\nu(\theta^n\Lambda) = 0$, Λ will be called a small set (see [1]). Obviously, any small set is a null set, but not all null sets are small sets. Indeed, if we take $\Lambda = \{x\} \times S \times \ldots$ such that $P_\nu(X_o = x) = 0$ we get $P_\nu(\Lambda) = 0$. However $\theta\Lambda = \Omega$ and therefore Λ is not a small set. Less trivial examples can be given for sets Λ in $\mathcal{T}$ in the case of an improperly homogeneous chain which will be defined below.

Examples of small sets: (1) any set Λ for which $P_x(\Lambda) = 0$ for all $x \in S$, (2) any invariant null set (because $T^n\Lambda = \Lambda$ for all $n \in Z$).

We shall further identify a class of Markov chains, called properly homogeneous, for which all the null sets of $\mathcal{T}$ are small sets.

Denote $\nu_n(B) = P_\nu(X_n \in B)$ for $B \in \mathcal{B}$ and let H_{n-1} be the Hahn set occurring in the Lebesgue decomposition of ν_n with respect to ν_{n-1}. A homogeneous Markov chain for which $\nu_1 \ll \nu_o$ and $\lim_{n\to\infty} \nu_n(H_n - H_{n+1}) = 0$ will be said to be properly homogeneous and improperly homogeneous otherwise.

To justify this definition we need to elucidate the implications of the conditions it imposes on the chain. Notice first that $\nu_1 << \nu_o$ implies $\nu_n << \nu_{n-1}$ for all n. Indeed $\nu(B) = 0$ implies $\nu_1(B) = 0$. But

$$\nu_1(B) = \int P(x,B)\,\nu(dx) .$$

It follows that $P(x,B) = 0$ for almost all x both with respect to ν and ν_1. Since

$$\nu_2(B) = \int P(x,N)\,\nu_1(dx) ,$$

we get $\nu_2(B) = 0$, and so on.

Consider next the equality

$$\nu_n(B) = \nu_n(B \cap H_{n-1}) + \nu_n(B \cap H^c_{n-1})$$

where $\nu_{n-1}(H^c_{n-1}) = 0$. The absolute continuity of ν_n with respect to ν_{n-1} implies that $\nu_n(H^c_{n-1}) = 0$ and $\nu_n(H_{n-1} \cap H_n) = \nu_n(H_n)$. If $\nu_{n-1}(H_{n-1} - H_n) = 0$, $\{\nu_n\}$ are equivalent measures. Suppose that $\nu_{n-1}(H_{n-1} - H_n) > 0$ then $P_\nu(X_n \varepsilon H_n | X_{n-1} \varepsilon H_{n-1}) = 1$. This implies

$$\int_{H_{n-1}} P(x,H_n)\,\nu_{n-1}(dx) = \nu_{n-1}(H_{n-1}) = 1 ,$$

which yields $P(x,H_n) = 1$ for almost all x with respect to ν_{n-1}. Further $P(x,H_n) = 1$ is also true for almost all x with respect to ν_n and this implies $P(X_{n-1} \varepsilon H_n | X_{n-2} \varepsilon H_{n-1}) = 1$. Now we can easily deduce that $\lim_{n\to\infty} \{X_n \varepsilon H_{n+1}\} = \{X_o \varepsilon H_1\}$ a.s. with respect to P_ν and therefore $\lim_{n\to\infty} \{X_n \varepsilon H_n - H_{n+1}\} = \{X_o \varepsilon H_o - H_1\}$ a.s. with respect to P_ν. Thus if a chain is properly homogeneous $P_\nu(\limsup_{n\to\infty}\{X_n \varepsilon (H_n - H_{n+1})\}) = 0$. Since $\nu_{n+1}(H_n - H_{n+1}) = 0$ implies $\nu_{n+k}(H_n - H_{n+1}) = 0$ for any $k > 0$, we can see that if a chain is improperly homogeneous, i.e. if $P_\nu(\limsup_{n\to\infty}\{X_n \varepsilon (H_n - H_{n+1})\}) > 0$, then the temporary homogeneity of its transition probabilities is of no use for the sequence of sets $\{H_n - H_{n+1} ; n = 0,1,\ldots\}$ which are of no relevance to the chain after time n (n is the last time $(H_n - H_{n+1})$ occurs with positive probability, i.e. $n = \sup_k\{(H_k - H_{k+1}) : P(H_k - H_{k+1}) > 0\})$.

The notion of a properly homogeneous chain for countable chains was introduced in [11].

PROPOSITION 1. If $\{X_n : n \geq 0\}$ is a properly homogeneous chain, then any null set in $\mathcal{T}$ is a small set.

PROOF. Suppose that $\Lambda \in \mathcal{T}$ and $P_\nu(\Lambda) = 0$. Since $P_\nu(\theta^{-1}\Lambda | X_n = x) = P_\nu(\Lambda | X_{n-1} = x)$ we can write

$$P_\nu(\theta^{-1}\Lambda) = \int P(\Lambda | X_{n-1} = x)\nu_n(dx) .$$

But

$$P_\nu(\Lambda) = \int P(\Lambda | X_{n-1} = x)\nu_{n-1}(dx) = 0$$

and using $\nu_n \ll \nu_{n-1}$ we get $P_\nu(\theta^{-1}\Lambda) = 0$. Inductively, we can prove that $P_\nu(\theta^{-n}\Lambda) = 0$ for any $n \geq 0$. We show now that $P_\nu(\theta\Lambda) = 0$. Indeed,

$$P_\nu(\theta\Lambda) = \int_{H_n - H_{n-1}} P(\theta\Lambda | X_n = x)\nu_n(dx)$$

$$\leq P_\nu(X_n \in (H_n - H_{n+1}))$$

and since $\lim_{n\to\infty} P_\nu(X_n \in (H_n - H_{n+1})) = 0$, $P_\nu(\theta\Lambda) = 0$ and the proof can be easily completed.

Proposition 3(iii), §2.2 and the above Proposition 1 together imply

COROLLARY 1. If $\{X_n : n \geq 0\}$ is a properly homogeneous chain and Λ is a positive set, then $P(\theta^n\Lambda) > 0$ for all $n \in Z$.

The following result is due to Abrahamse [1].

PROPOSITION 2. Suppose that $\Lambda \in \mathcal{T}$ and that $\Lambda \,\Delta\, \theta\Lambda$ is a small set. Then $\Lambda' = \bigcup_{n=-\infty}^{\infty} \theta^n\Lambda$ is an invariant set and $P_\nu(\Lambda' \,\Delta\, \Lambda) = 0$ for all starting probabilities ν .

PROOF. One can easily check that $\bigcup_{n=-\infty}^{\infty} \theta^n\Lambda$ is invariant. Further, for any

$n \in Z$

$$\Lambda \Delta \theta^n \subseteq (\Lambda \Delta \theta\Lambda) \cup (\theta\Lambda \Delta \theta^2\Lambda) \cup \ldots \cup (\theta^{n-1}\Lambda \Delta \theta^n\Lambda) \ .$$

Applying Proposition 3, §2.2 we get $\theta^m(\Lambda \Delta \theta\Lambda) = \theta^m\Lambda \Delta \theta^{m+1}\Delta$ for $m = 0,1,\ldots,n-1$ It follows that $P_\nu(\theta^n\Lambda \Delta \Lambda) = 0$ for all $n \in Z$ and the proof is readily completed on noticing that

$$P_\nu(\Lambda' \Delta \Lambda) = P_\nu\Big(\bigcup_{n=-\infty}^{\infty} \theta^n\Lambda \Delta \Lambda\Big) \leq \sum_{n=-\infty}^{\infty} P_\nu(\theta^n\Lambda \Delta \Lambda) = 0 \ .$$

2.4 SPACE-TIME CHAINS. Denote by $\mathcal{N}$ the family of all subsets of N . A process $\{(X_n,T_n) : n \geq 0\}$ with T_n taking values in $(N,\mathcal{N})$, $n = 0,1,\ldots$ is called a space-time chain (associated with $\{X_n : n \geq 0\}$) if $T_{n+1} = T_n + 1$. In what follows we suppose that $\{X_n : n \geq 0\}$ is a nonhomogeneous Markov chain and confine our attention to the space-time chains for which $T_o = k$ for a certain k in N , i.e. the chain $\{(X_n,n+k) : n \geq 0\}$.

Let $\tilde{\Omega} = (N \times S) \times (N \times S) \times \ldots$ and $\tilde{\mathfrak{F}} = (\mathcal{N} \otimes \mathcal{B}) \otimes (\mathcal{N} \otimes \mathcal{B}) \otimes \ldots$. The main reason for the usefulness of the space-time chain concept is given by the following:

PROPOSITION 1. $\{(X_n,n+k) : n \geq 0\}$ for any $k \in N$ can be thought of as a homogeneous Markov chain on the probability space $(\tilde{\Omega},\tilde{\mathfrak{F}},\tilde{P}_{\tilde{\nu}})$ where $\tilde{P}_{\tilde{\nu}}$ is determined by the transition probability function

$$\tilde{P}((x,m+k) \ ; B \times \{n+k\}) = \begin{cases} P_n(x,B) & \text{for } x \in S \ , \ B \in \mathcal{B}, \ m = n-1, \ n \in N \\ 0 & \text{otherwise} \end{cases} \qquad (2.4)$$

and the starting probability $\tilde{\nu}(k,\cdot) = \nu(\cdot)$.

PROOF. It is easy to see that the s step transition probability function of the space-time chain is

$$\tilde{P}^{(s)}((x,m+k);(B\times\{n+k\})) = \begin{cases} P^{m,m+s}(x,B) & \text{for } x\varepsilon S,\ B\varepsilon\mathcal{B},\ n=m+s,\ m=0,1,\ldots \\ 0 & \text{otherwise} \end{cases} \tag{2.5}$$

Further

$$P_\nu((X_n,n+k)\ \varepsilon\ B\times\{n+k\}\mid\mathfrak{F}_m) = P_\nu^{m,n}(X_m,B) \qquad P_\nu \text{ a.s.} \tag{2.6}$$

Now combining (2.5), (2.6) and taking into account the relationship between $\tilde{P}_{\tilde{\nu}}$ and P_ν we get

$$\tilde{P}^{(n-m)}((X_m,m+k)\ ;\ B\times\{n+k\}) = \tilde{P}_{\tilde{\nu}}((X_n,n+k)\ \varepsilon\ B\times\{n+k\}\mid\tilde{\mathfrak{F}}_m) \qquad \tilde{P}_{\tilde{\nu}} \text{ a.s.}$$

Thus the formula (2.3) defining an homogeneous chain is verified and the proof is complete.

Proposition 1 has been known for a long time in connection with the Potential theory (see e.g. Doob [16]).

REMARK. The above Proposition asserts that a space-time chain turns a nonhomogeneous chain into a homogeneous one. This is, however, done at the expense of complicating the state-space of the chain. Also, since any state (n,x) of this new chain appears only once (at time $n-k$) the absolute probabilities $\tilde{\nu}_n(\tilde{B}) = \tilde{P}_{\tilde{\nu}}((X_n,n+k)\ \varepsilon\ \tilde{B})$ are mutually singular. Thus, such a chain is improperly homogeneous in the sense of the definition given in 2.2. However, we shall see further on, that there are still many properties of homogeneous chains which applied to space-time chains yield relevant properties of the original chain $\{X_n : n\geq 0\}$, even in the case when the original chain is homogeneous. The following Proposition 2 is one of this kind.

PROPOSITION 2. The formulas

$$f_n(X_n,X_{n+1},\ldots) = f((X_n,n+k)\ ,\ (X_{n+1},n+k+1),\ldots)$$

for $n=0,1,\ldots$ with $f_n(X_n,X_{n+1},\ldots) = 1_\Lambda$, $n=0,1,\ldots$ set up a one-to-one correspondence between the events Λ of the tail σ-field $\mathcal{T}$ and the events

$\tilde{\Lambda} = \{f((X_n,n+k),(X_{n+1},n+k+1),\ldots\}$ for $n=0,1,\ldots$ of the invariant σ-field $\tilde{\mathcal{J}}$ for any $k \in N$. This correspondence preserves the probability, i.e. $P_\nu(\Lambda) = \tilde{P}_{\tilde{\nu}}(\tilde{\Lambda})$.

PROOF. Suppose that $\Lambda \in \mathcal{J}$. Then there exists a real function f_o on $(S^\infty, \mathcal{B}^\infty)$ such that $Y = 1_\Lambda = f_o(X_o,X_1,\ldots)$. If we further require that $\Lambda \in \mathcal{T}$ then there must exist a sequence of measurable functions on $(S^\infty, \mathcal{B}^\infty)$, say $\{f_n : n \geq 0\}$, such that

$$Y = f_n(X_n,X_{n+1},\ldots) \quad \text{for } n=0,1,\ldots \tag{2.7}$$

If $\Lambda \in \mathcal{J}$ we have $\theta^n Y = Y$ for $n=1,2,\ldots$ and in such a case there exists a function f such that

$$Y = f(X_n,X_{n+1},\ldots) \quad \text{for } n=0,1,\ldots \tag{2.8}$$

Reciprocally, if a set Λ satisfies (2.7) (or (2.8)) then $\Lambda \in \mathcal{T}$ (or $\Lambda \in \mathcal{J}$).

Suppose now that $\tilde{Y}$ is the indicator of a set $\tilde{\Lambda}$ in $\tilde{\mathcal{J}}$. Then, according to what we have seen before, $\tilde{Y}$ can be represented as

$$\tilde{Y} = \tilde{f}((X_n,n+k),(X_{n+1},n+k+1),\ldots) \quad \text{for } n=0,1,\ldots$$

But $\tilde{f}((X_n,n+k),(X_{n+1},n+k+1),\ldots) = f_n(X_n,X_{n+1},\ldots)$ $n=0,1,\ldots$ and it is easily seen that such equalities set up a one-to-one correspondence between $\mathcal{T}$ and $\tilde{\mathcal{J}}$. Finally, $P_\nu(\Lambda) = \tilde{P}_{\tilde{\nu}}(\tilde{\Lambda})$ follows easily from the definition of $\tilde{P}_{\tilde{\nu}}$.

Proposition 2 is essentially due to Jamieson and Orey [27] (see also [36]).

3. REPRESENTATIONS FOR ASYMPTOTIC RANDOM VARIABLES AND EVENTS.

A transition probability kernel P defines a linear mapping on the set of positive and $\mathcal{B}$-measurable functions into itself by

$$Pf(x) = \int P(x,dy)f(y) \tag{3.1}$$

If for any $x \in S$, $Ph(x) = h(x)$, h will be said to be a P-harmonic function.

Consider a sequence of transition probability kernels $(P_n ; n \geq 1)$. $h(x,n)$ with $x \in S$ and $n \in N$ will be said to be a $\tilde{P}$-harmonic (or space-time harmonic) function if $P_n h(x,n) = h(x,n-1)$ for all $x \in S$ and $n \in N$.

We shall write h_n for $h(n,\cdot)$ and agree to suppress the qualifiers P and $\tilde{P}$ when referring to harmonic and space-time harmonic functions.

We notice easily that a state-space harmonic function is a harmonic function corresponding to the space-time transition probability kernel $\tilde{P}$ associated to the space-time chain $\{(X_n,n) : n \geq 0\}$, where the original chain $\{X_n : n \geq 0\}$ is a nonhomogeneous Markov chain with transition probabilities functions $\{P_n : n \geq 1\}$.

In what follows we shall confine our attention to the bounded positive harmonic (space-time harmonic) functions and we shall see that there is an important connection between such functions and the invariant σ-fields (tail σ-fields).

We notice first that whatever the starting measure ν , $\{h(X_n), \mathfrak{F}_n : n \geq 0\}$ defines a martingale with respect to the probability space $(\Omega, \mathfrak{F}, P_\nu)$. Indeed, since h is positive and bounded, the martingale property $E_\nu(|h(X_n)|) < \infty$ is satisfied, whereas the second property

$$E_\nu(h(X_n) \mid \mathfrak{F}_{n-1}) = h(X_{n-1}) \quad P_\nu \text{ a.s. },$$

is a consequence of the Markov property and (3.1).

Since $\{h(X_n), \mathfrak{F}_n ; n \geq 0\}$ is a bounded martingale, the martingale convergence theorem ([31] p. 398) implies that

$$\lim_{n\to\infty} h(X_n) = X \tag{3.2}$$

exists P_ν a.s. . Thus, to each bounded and positive harmonic function there corresponds a tail random variable X . We can further check that X is P_ν a.s.

equal to an invariant random variable (say) X'. Indeed, define $X'(\omega) = \liminf_{n\to\infty} h(X_n(\omega))$. Because $\theta h(X_n(\omega)) = h(X_{n+1}(\omega))$ we get that $\theta X(\omega) = X(\omega)$ for $\omega \in \{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}$, whereas if ω belongs to the set $\{\omega : \liminf_{n\to\infty} X_n(\omega) \neq X(\omega)\}$, then $\theta \liminf_{n\to\infty} X_n(\omega) = \liminf_{n\to\infty} X_{n+1}(\omega) = \liminf_{n\to\infty} X_n(\omega)$. Hence $X'(\omega)$ is invariant.

Reciprocically, if $X(\omega)$ is a bounded, positive and invariant random variable, $h(x) = E_x(X)$ is a harmonic function. Indeed, the Markov property, the measurability of X with respect to $\mathcal{J}$ and the invariance of X yield

$$\begin{aligned} PE_x(X) &= \int P(x,dy)E_y(X) \\ &= \int P(x,dy)E(\theta^{-1}X|X_1 = y) \\ &= \int P(x,dy)E(X|X_1 = y) \\ &= E_x(X) \end{aligned}$$

for all $x \in S$. If we agree to call equivalent two invariant variables X and X' for which $P_\nu(X \neq X') = 0$ for any starting probability ν, then we can easily see that to any variable Z from an equivalent class, there corresponds the same harmonic function h. On the other hand, if two harmonic functions h and h' are not identical, i.e. there exists x in S such that $h(x) \neq h'(x)$ then the variable X corresponding to h and X' corresponding to h' are not equivalent, since taking $\nu = \delta(x)$ we get $E_x(X) \neq E_x(X')$. Thus, we have proved the following basic result of Blackwell [4] (see also [8]).

THEOREM 1. (i) Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain. The formula

$$h(x) = E_x(X)$$

set up a one-to-one correspondence between equivalent classes of positive, bounded, invariant random variables X and positive, bounded harmonic functions

h .

(ii) $\{X_n : n \geq 0\}$, h and X are related by the formula

$$\lim_{n\to\infty} h(X_n) = X \qquad P_\nu \text{ a.s.}$$

for any starting probability ν .

Suppose that we associate to the coordinate variables $\{X_n\}$ defined on $(\Omega, \mathfrak{F}^n)$ a nonhomogeneous Markov chain assuming the starting measure ν and the sequence of transition probabilities $\{P_k, k \geq n\}$. Denote by P_ν^n the probability measure on $\mathfrak{F}^n$ determined by ν and $\{P_k : k \geq n\}$. We shall denote by $E(Y|X_n = x)$ the mathematical expectation of the random variable Y with respect to P_x^n , where P_x^n stands for P_ν^n with $\nu = \delta(x)$. Two tail variables Y and Y' will be said to be equivalent if $P_\nu^n(Y \neq Y') = 0$ for $n = 0,1,\ldots$ and any starting probability ν

Theorem 1 has a parallel result for space-time harmonic functions and tail σ-fields, expressed by the following

THEOREM 2. (i) Suppose that $\{X_n : n \geq 0\}$ is a nonhomogeneous Markov chain. The formulas

$$h(n,x) = E(Y|X_n = x) \qquad n = 0,1,\ldots \quad x \in S$$

set up a one-to-one correspondence between equivalent classes of positive, bounded tail random variables Y and positive, bounded space-time harmonic functions $\{h(n,x)\}$.

(ii) $\{P_\nu^m\}$, $\{h(n,x)\}$ and Y are related by the formula

$$\lim_{n\to\infty} h(n,X_n) = Y \qquad P_\nu^m \text{ a.s.}$$

for $m = 0,1,\ldots$ and any starting probability ν .

PROOF. Consider the space-time chain $\{(X_n, n+k) : n \geq 0\}$, associated to the chain $\{X_n : n \geq 0\}$ assuming the probability measure P_ν^m . Since according to Proposition 1, §2.4, $\{(X_n, n+k), n \geq 0\}$ can be thought of as a homogeneous chain on a certain probability space $(\tilde{\Omega}, \tilde{\mathfrak{F}}, \tilde{P}_{m,\nu})$, $h(n,x)$ is easily seen to be harmonic with respect to the transition matrix $\tilde{P}$ defined by (2.4) and a fortiori with respect to any transition matrix $\tilde{P}_m$ associated to the measure P_ν^m . Further, the harmonicity of h(x,n) yields

$$h(n,x) = E(h(n+1,X_{n+1})|X_n = x) .$$

Thus $\{h(n, X_n) : \mathfrak{F}_n : n \geq m\}$ is the convergent martingale corresponding to $\{h(X_n) : \mathfrak{F}_n : n \geq 0\}$ in the previous Theorem 1.

Now Proposition 2 §2.4 and Theorem 1 given above provide the remaining part of the proof.

Theorem 2 was given by Neveu [33] (p. 154). The proof given here is new.

COROLLARY 1. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain. Then the following two conditions are equivalent

(i) All positive, bounded, harmonic functions are constant.

(ii) The invariant σ-field $\mathfrak{J}$ is trivial under any starting measure ν .

PROOF. Suppose that there exists a non-constant, positive, bounded, harmonic function. Then there are two points x_1 and x_2 such that $h(x_1) \neq h(x_2)$ Assume now that we take the starting measure to be $\nu = \frac{1}{2}(\delta(x_1) + \delta(x_2))$. Then according to Theorem 1 there exists a random variable X such that $h(x) = E_x(X)$ and $E_{x_1}(X) \neq E_{x_2}(X)$. But if such a situation occurs, X cannot be P_ν a.s. constant since in that case $E_{x_1}(X) = E_{x_2}(X_2) = c$ where c is a constant with $P_\nu(X = c) = 1$ and we would get a contradiction. The converse assertion is a straightforward consequence of Theorem 1.

Analoguously, Theorem 2 yields

COROLLARY 2. Suppose that $\{X_n : n \geq 0\}$ is a nonhomogeneous Markov chain. Then the following two conditions are equivalent

(i) All positive, bounded, space-time harmonic functions are constant.

(ii) The tail σ-field $\mathcal{T}$ is trivial under any probability P_ν^n, $n = 0,1,\ldots$ and any starting probability ν.

Corollary 1 was proved by Blackwell [4]. Corollary 2 was given in Jamieson and Orey [27] for homogeneous chains.

These Corollaries have some important consequences to Martin boundary theory in connecting the harmonic (space-time harmonic) functions theory to the theory of the asymptotic σ-fields of the chain. This connection will be more fully explored in the next chapter.

We shall next deal with representations for invariant and tail events. It is assumed that ν is fixed and we suppress the qualifier P_ν when referring to a.s. statements or null sets.

A set C in $\mathcal{B}$ will be said to be almost closed if $\lim_{n\to\infty}\{X_n \in C\}$ exists a.s. and $P_\nu(\lim\sup\{X_n \in C\}) > 0$. B will be said to be a transient set if $\lim\sup_{n\to\infty}\{X_n \in B\}$ is a null set. Denote by $\mathcal{C}$ the class of all almost closed and transient sets by $\mathcal{U}$ the class of all transient sets and by $\mathcal{N}$ the class of sets in $\mathcal{J}$ which are null sets. It is easy to see that $\mathcal{C}$ is a boolean algebra and $\mathcal{U}$ is an ideal in $\mathcal{C}$. Denote by $\mathcal{J}/\mathcal{N}$ and $\mathcal{C}/\mathcal{U}$ the quotient boolean algebras obtained by factorizing $\mathcal{J}$ and $\mathcal{C}$ by $\mathcal{N}$ and $\mathcal{U}$ respectively. The following result exhibits the relationship between the elements of $\mathcal{J}/\mathcal{N}$ and $\mathcal{C}/\mathcal{U}$.

THEOREM 3. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain. Then to each invariant set Λ there corresponds a transient or almost closed set B such that $\Lambda = \lim_{n\to\infty}\{X_n \in B\}$ a.s. according as Λ is a null set or not. This correspondence is an isomorphism from $\mathcal{J}/\mathcal{N}$ onto $\mathcal{C}/\mathcal{U}$.

PROOF. Suppose that Λ is invariant and introduce the martingale $\{P_\nu(\Lambda|\mathcal{F}_n), \mathcal{F}_n ; n \geq 0\}$. Since $\Lambda \in \mathcal{T}$, the Markov property implies

$P_\nu(\Lambda | \mathfrak{F}_n) = P_{X_n}(\Lambda)$, a.s. . The martingale convergence theorem applied to this bounded martingale yields $\lim_{n\to\infty} P_{X_n}(\Lambda) = 1_\Lambda$ a.s. . The case $P_\nu(\Lambda)$ can be easily disposed by taking $B = \phi$. Suppose that $P_\nu(\Lambda) > 0$ and define now $C = \{x : P_x(\Lambda) \geq 0.5\}$. Then $\lim_{n\to\infty} 1_{\{X_n \varepsilon C\}} = 1_\Lambda$ a.s., which yields $\lim_{n\to\infty}\{X_n \varepsilon C\} = \Lambda$, a.s. . Reciprocically, suppose that $\lim_{n\to\infty}\{X_n \varepsilon C\} = \Lambda$ exists a.s. for any starting measure ν . Then $\liminf_{n\to\infty}\{X_n \varepsilon C\} = \bigcup_{n=o}^{\infty} \bigcap_{m=n}^{\infty}\{X_n \varepsilon C\}$ is an invariant event and $P_\nu(\lim_{n\to\infty}\{X_n \varepsilon C\}$ a.s. $\Delta \liminf_{n\to\infty}\{X_n \varepsilon B\}) = 0$.

The remaining part of the theorem is rather straightforward and will be left to the reader as an exercise. Theorem 3 is due to Blackwell [4].

Theorem 3 like Theorem 1 has an analogue for nonhomogeneous Markov chains and tail σ-fields which can be obtained by applying Theorem 3 to the space-time chain.

Denote by $\mathcal{J}$ the class of all sequences $A = (A_o, A_1, \ldots, A_n, \ldots)$ such that $\lim_{n\to\infty}\{X_n \varepsilon A_n\}$ exists P_ν a.s. and $P_\nu(\limsup_{n\to\infty}\{X_n \varepsilon A_n\}) > 0$ and of all sequences $A = (A_o, A_1, \ldots, A_n, \ldots)$ such that $\limsup_{n\to\infty}\{X_n \varepsilon A_n\}$ is a null set. Write $\mathcal{V}$ for the class of all sequences $A = (A_o, A_1, \ldots, A_n, \ldots)$ such that $\limsup_{n\to\infty}\{X_n \varepsilon A_n\}$ is a null set and $\mathfrak{M}$ for the class of all events in $\mathcal{T}$ which are null. For $A = (A_o, A_1, \ldots, A_n, \ldots)$ and $B = (B_o, B_1, \ldots, B_n, \ldots)$ we shall define $A^c = (A_o^c, A_1^c, \ldots, A_n^c, \ldots)$, $A \cup B = (A_o \cup B_o,\ A_1 \cup B_1, \ldots,\ A_n \cup B_n, \ldots)$, $\theta A = (A_1, A_2, \ldots, A_{n+1}, \ldots)$ and $\theta^{-1}A = (S, A_o, \ldots, A_{n-1}, \ldots)$. It is easy to check that $\mathcal{J}$ is a boolean algebra and $\mathcal{V}$ is an ideal in $\mathcal{J}$. Further $\mathcal{T} / \mathfrak{M}$ and $\mathcal{J} / \mathcal{V}$ will denote the quotent boolean algebras obtained by factorizing $\mathcal{T}$ and $\mathcal{J}$ by $\mathfrak{M}$ and $\mathcal{V}$ respectively.

THEOREM 4. Assume that $\{X_n : n \geq 0\}$ is a nonhomogeneous Markov chain. Then to each tail event Λ there corresponds a sequence $(B_0, B_1, \dots, B_n, \dots)$ in $\mathcal{V}$ or $\mathcal{Y} - \mathcal{V}$ such that $\lim_{n\to\infty}\{X_n \in B_n\} = \Lambda$ P_ν a.s. for any starting measure ν, according as Λ is in $\mathcal{M}$ or $\mathcal{T} - \mathcal{M}$. This correspondence is an isomorphism from $\mathcal{Y}/\mathcal{V}$ onto $\mathcal{T}/\mathcal{M}$.

REMARK. The isomorphism stated by Theorems 3 and 4 as well as the one to be considered in the sequel cannot be extended from Boolean algebras to σ-algebras, as can be seen from the following example: Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain assuming only transient states. Then $P(X_n = i.i.o) = 0$ for any $i \in S$, whereas $P(\liminf_{n\to\infty}\{X_n \in S\}) = P(\limsup_{n\to\infty}\{X_n \in S\}) = 1$

We next confine our attention to the tail σ-field of a homogeneous Markov chain and we shall show that an isomorphism of the type alluded to in Theorem 4 can be shown to commute with θ for homogeneous chains if the null sets considered in the statement of Theorem 4 are replaced by small sets.

A sequence $A = (A_0, A_1, \dots)$ will be said to be totally transient if $\limsup_{n\to\infty}\{X_n \in A_n\}$ is a small set and totally non-transient if $P_\nu(\limsup\{X_n \in A_n\}) > 0$ and $\limsup_{n\to\infty}\{X_n \in A_n\} \,\Delta\, \liminf_{n\to\infty}\{X_n \in A_n\}$ is a small set. We shall say that $\Lambda \,\Delta\, \lim_{n\to\infty}\{X_n \in A_n\}$ a.s. is a small set if both $\Lambda \,\Delta\, \liminf_{n\to\infty}\{X_n \in A_n\}$ and $\Lambda \,\Delta\, \limsup_{n\to\infty}\{X_n \in A_n\}$ are small sets. Denote by $\mathcal{L}$ the class of all sets in $\mathcal{T}$ which are small sets, by $\mathcal{W}$ and $\mathcal{R}$ the classes of all totally transient and totally transient as well as totally non-transient sequences respectively. $\mathcal{R}/\mathcal{W}$ and $\mathcal{T}/\mathcal{L}$ will denote the quotent boolean algebra obtained by factorizing $\mathcal{R}$ and $\mathcal{T}$ by $\mathcal{W}$ and $\mathcal{L}$ respectively.

The following Theorem 5 extends a result established by Abrahamse [1] for countable chains.

THEOREM 5. Assume that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain. Then to each set Λ in $\mathcal{T}$ there corresponds a totally transient or a totally non-transient sequence $A=(A_o,A_1,\dots)$ such that $\Lambda \,\Delta\, \lim_{n\to\infty}\{X_n \in A_n\}$ a.s. is a small set according as Λ is in $\mathcal{L}$ or in $\mathcal{T}-\mathcal{L}$. This correspondence is an isomorphism from $\mathcal{T}/\mathcal{L}$ onto $\mathcal{R}/\mathcal{W}$ and commutes with θ.

PROOF. We can easily check that $\mathcal{R}$ is a boolean algebra and $\mathcal{W}$ an ideal in $\mathcal{R}$ on using elementary measure and set operations properties. Recall further the martingale $\{P_\nu(\Lambda|\mathcal{F}_n), \mathcal{F}_n : n \geq 0\}$ used in the proof of Theorem 3. Under the assumptions of the theorem, we get that $P_\nu(\Lambda|\mathcal{F}_n) = P_{X_n}(\theta^n\Lambda)$. Thus if we denote $A_n = \{x : P_x(\theta^n\Lambda) > 0.5\}$ then $\lim_{n\to\infty}\{X_n \in A_n\} = \Lambda$ a.s. . Further, if instead of Λ we consider the set $\theta\Lambda$, the same martingale argument as above yields $\lim_{n\to\infty}\{X_n \in A_{n+1}\} = \theta\Lambda$ a.s. and thus the correspondence $\Lambda \to (A_o,A_1,\dots)$ commutes with θ. But the same argument can be applied to $\theta^k\Lambda$ for any $k \in Z$ to yield $\lim_{n\to\infty}\{X_n \in A_{n+k}\} = \theta^k\Lambda$ a.s. and now using Proposition 3, §2.2 we get that $\Lambda\,\Delta \lim_{n\to\infty}\{X_n \in A_n\}$ a.s. is a small set. Reciprocically suppose that $(A_o,A_1,\dots,\dots) \in \mathcal{R}$. Take $\Lambda = \lim_{n\to\infty}\inf\{X_n \in A_n\}$; then $\Lambda\,\Delta \lim_{n\to\infty}\{X_n \in A_n\}$ a.s. is easily checked to commute with θ and to be a small set. Notice finally that the totally transient sequences and the small sets if added or removed from the $\mathcal{R}$ and $\mathcal{T}$ respectively, do not alter the above established correspondence and the proof is now complete.

REMARK. The isomorphism stated in Theorem 2 is the restriction to the subclasses $\mathcal{J}$ and $\mathcal{C}$ respectively of the isomorphism stated by Theorem 4 Indeed, to see this it is sufficient to notice that any null invariant set is a small set and that for any C such that $P_\nu(\lim_{n\to\infty}\sup\{X_n \in C\}) = 0$, $\lim_{n\to\infty}\{X_n \in C\}$ a.s. is a small set.

4. MARTIN BOUNDARY THEORY AND ASYMPTOTIC σ-FIELDS OF MARKOV CHAINS.

Suppose that $\{X_n : n \geq 0\}$ is a countable Markov chain assuming the state space S and denote by $P^n(i,j)$ the n step transition probability from i to j. Assume that the chain is transient and consider the Green function*

$$G(i,j) = \sum_{n=o}^{\infty} P^n(i,j)$$

where $P^o(i,j) = \delta_{i,j}$, $\delta_{i,j}$ being the Kronecker symbol. Define the Martin exit boundary kernel K by

$$K(i,j) = \frac{G(i,j)}{\sum_{n=o}^{\infty} \nu^{(n)}(j)}$$

and consider the metric

$$d(x_1,x_2) = \sum_{i\varepsilon S} |K(i,x_1) - K(i,x_2)| 2^{-i} U_i(1)$$

where $U_i(1)$ is the probability that a path from i ever reaches 1. The space S is completed by adding limit points and so completed is a compact metric space. Let S' be the set consisting of the limit points of metrized S in the completed space. The set S' is called the Martin exit boundary of S. A harmonic function h* is said to be minimal if for any harmonic function h such that $h(i) \leq h^*(i)$ for all $i \varepsilon S$, there exists a constant c such that $h = c h^*$. A point ξ in S' is called minimal if $K(\cdot,\xi)$ is a minimal harmonic function.

The main object of the Martin boundary theory is the identification of the class of all harmonic functions associated to a transition probability kernel and for this it suffices to identify the minimal harmonic functions. Indeed, if we denote by S_e the set of all minimal boundary points, then there is a

* For clear surveys of Martin boundary theory for countable chains, the reader can consult Neveu [34] or Kemeny Snell and Knapp [28].

representation theorem for harmonic function, called the Martin-Doob-Hunt integral representation, asserting that any harmonic function h can be represented as

$$h(i) = \int_{S_e} K(i,\xi)V(d\xi)$$

V being a probability measure on the borelian subsets of S_e which is uniquely determined by h .

There is a useful criterion for minimality of a harmonic function, based on examining the Martin boundary of the h-process associated to a harmonic function h . An h-process is a Markov chain assuming the transition probabilities

$$Q(i,j) = \begin{cases} \dfrac{P(i,j)h(j)}{h(i)} & \text{if } 0 < h(i) < \infty \\ \\ 0 & \text{otherwise} \end{cases}$$

If 1 is a minimal harmonic function for the h-process, h is minimal for the original chain. Equivalently, if the only bounded, positive harmonic function for the h-processare constant, h is a minimal harmonic function. According to Corollary 1, §3 this happens if and only if $\mathcal{J}$ is trivial for the h-process and therefore the identification of harmonic functions is essentially connected with the structure of the invariant σ-field.

If we consider the space-time chain derived from a nonhomogeneous chain $\{X_n : n \geq 0\}$, we get a rather simpler Green function:

$$G((m,i) : (n,j)) = P^{m,n}(i,j)$$

where $P^{m,n}(i,j) = P(X_n = j \mid X_m = i)$ with $i , j \in S$ and $m , n \in N$, and the same arguments as before applied to the space-time chain, as well as the Corollary 2, §2 show that the identification of the space-time harmonic functions is essentially connected with the structure of the tail σ-field $\mathcal{T}$. Thus results concerning the Martin boundary theory for some types of chains as, for example, those given by Lamperty and Snell [20] or Blackwell and Kendall [5], etc. can be interpreted as assertions about the tail σ-fields of the chains.

The above mentioned results in the Martin boundary theory refer to countable Markov chains. Some of these properties have been extended to more general cases. However, the Martin-Doob-Hunt representation as well as the most relevant properties of h-processes have not (at least not yet) been extended beyond the countable case.

The connection between the Martin boundary theory and the theory of invariant events developed by Blackwell in [4] has been remarked by Doob in 1959 [16]. The connection between the space-time Markov chains and the tail σ-field has been discovered only in 1967 by Jamieson and Orey [27] and rediscovered by Abrahamse in 1969 [1]. Many authors of papers which appeared in the meantime have been unaware of the fact that a result concerning the Martin boundary of a particular chain was the same as a result formulated in the language of the tail σ-field in another paper and even recently some authors seem unaware of this connection.

Moreover, there is more to gain by applying the Martin boundary theory to asymptotic σ-fields of a Markov chain and the object of the remainder of this section is to point out some applications of this kind. Namely, we shall investigate some consequences of the basic almost surely convergence theorem in the Martin boundary theory to the structure of the tail and invariant σ-field of a Markov chain.

Let λ and μ be two probability measures. The Radon-Nykodim derivative of the restriction of λ to the sub σ-algebra $\mathcal{A}$ with respect to the restriction of

μ to $\mathcal{A}$ will be denoted by $(d\lambda/d\mu)\Big|_{\mathcal{A}}$. Suppose that $g(x,A)$ defined by

$$g(x,A) = \sum_{n=o}^{\infty} P_x(X_n \varepsilon A)$$

is a regular kernel and define the measures

$$g_\mu(A) = \sum_{n=o}^{\infty} P_\mu(X_n \varepsilon A)$$

and

$$g_\nu(A) = \sum_{n=o}^{\infty} P_\nu(X_n \varepsilon A) .$$

Then both g_μ and g_ν are σ-finite measures on $\mathcal{B}$. Write now

$$g_\mu(dy) = K_\nu(\mu,y)g_\nu(dy) + s_\mu(dy)$$

for the Lebesgue decomposition of g_μ with respect to g_ν . Here s_μ and g_ν are mutually singular on $\mathcal{B}$. $K_\nu(\mu,x)$ is called Martin boundary kernel.

The basic almost sure convergence result in Martin boundary theory is the following

THEOREM 1. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain. Then

$$\lim_{n\to\infty} K_\nu(\mu,X_n) = \frac{dP_\mu}{dP_\nu}\Big|_{\mathcal{J}} \qquad P_\nu \text{ a.s.}$$

Theorem 1 is basically due to Abrahamse [2] (see also Revuz [39]). It is based on an idea used in the countable case by Hunt [24]. For an extension to non-regular kernels based on Chacon-Ornstein ergodic theorem see Derriennic [13]. Denote by Γ the set of all probability measures on $\mathcal{B}$. We shall next confine our attention to the case when $\mathcal{J}$ is trivial with respect to any ν in Γ .

THEOREM 2. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain. Then the following three statements are equivalent

(i) $\mathcal{J}$ is trivial with respect to any starting probability ν

(ii) The probability measures $(P_\nu)\nu \in \Gamma$ agree on $\mathcal{J}$

(iii) $\lim_{n\to\infty} K_\nu(\mu, X_n) = 1 \quad P_\nu$ a.s.

for any μ, $\nu \in \Gamma$

PROOF. Suppose that (i) holds. Then $\left.\frac{dP_\mu}{dP_\nu}\right|_{\mathcal{J}}$ must be P_ν a.s. constant, since it is $\mathcal{J}$-measurable. Assume that μ and ν are singular. Then there exists a set H in $\mathcal{J}$ such that $P_\mu(H) = 1$ and $P_\nu(H) = 0$. If we consider the starting probability $\lambda = \frac{1}{2}(\mu+\nu)$ then $P_\mu \ll P_\lambda$ and $P_\nu \ll P_\lambda$. The only case that does not contradict the singularity of P_μ and P_ν is when $P_\lambda(H) > 0$ and $P_\lambda(H^c) > 0$ but such a situation is excluded by the triviality of $\mathcal{J}$ with respect to P_λ. Thus (i) $\to$ (ii). Suppose now that (ii) holds. Then $\frac{dP_\mu}{dP_\nu} = 1 \quad P_\nu$ a.s. and (iii) follows from Theorem 1. Finally, assume that (iii) holds and $\mathcal{J}$ is not trivial with respect to a certain starting probability ν. Then there would exist two disjoint invariant sets I_1 and I_2 such that $P_\nu(I_1) > 0$ and $P_\nu(I_2) > 0$. By the martingale convergence theorem $\lim_{n\to\infty} P_\nu(I_1 | X_n) = 1_{I_1} \quad P_\nu$ a.s. . Thus there exists x in S such that $P_x(I_1) > P_\nu(I_1)$. But by (iii) one has $P_\nu(I_1) = P_x(I_1)$ and this contradiction completes the proof.

Write now

$$P_\mu(X_n \in dy) = \tilde{K}_\nu(\mu,y) P_\nu(X_n \in dy) + \tilde{s}_\mu(X_n \in dy)$$

for the Lebesgue decomposition of P_μ with respect to P_ν. Here P_ν and $\tilde{s}_\mu$ are mutually singular on $\mathcal{B}$. It is easy to see that $\tilde{K}_\nu(\mu,y)$ is the Martin boundary kernel of the space-time chain $\{(X_n, n+k) : n \geq 0\}$ with $k \in N$. Theorem 1 has an analogue for the tail σ-fields and nonhomogeneous Markov chains expressed by the following

THEOREM 3. Suppose that $\{X_n : n \geq 0\}$ is a nonhomogeneous Markov chain. Then

$$\lim_{n\to\infty} \tilde{K}_\nu(\mu, X_n) = \frac{dP_\mu}{dP_\nu}\Big|_{\mathcal{T}} \qquad P_\nu \text{ a.s.}$$

The proof of this Theorem follows easily from Theorem 1 and Proposition 2 §2.4.

The convergence of $\tilde{K}_\nu(\mu, X_n)$ in the countable case was proved by Doob [16], [17] but the limit was not identified as in Theorem 3. The possibility of extending Doob's result to homogeneous chains with separable state space and assuming transition probability densities was mentioned by Orey [36]. Theorem 3 contains all these results as particular cases and will be further seen to yield a large number of results concerning the tail σ-field of nonhomogeneous chains.

Also, Theorem 2 has an analogue expressed by the following

THEOREM 4. Suppose that $\{X_n : n \geq 0\}$ is a nonhomogeneous Markov chain. Then the following three statements are equivalent

(i) $\mathcal{T}$ is trivial with respect to any probability measure P_ν with $\nu \in \Gamma$

(ii) The probability measures $(P_\nu)_{\nu \in \Gamma}$ agree on $\mathcal{T}$

(iii) $\lim_{n\to\infty} \tilde{K}_\nu(\mu, X_n) = 1 \quad P_\nu$ a.s.

The proof of this Theorem follows easily from Theorem 3 and Proposition 2 §2.4.

The following result gives a "0 - 2 law" for nonhomogeneous Markov chains.

THEOREM 5. Suppose that $\{X_n : n \geq 0\}$ is a nonhomogeneous Markov chain and denote

$$\alpha(x,y,m) = \lim_{n\to\infty} \| P^{m,n}(x,\cdot) - P^{m,n}(y,\cdot) \| .$$

Then

(i) $$\sup\{\lim_{n\to\infty} \| P^{m,n}_{\nu_m} - P^{m,n}_{\nu'_m} \| ; \ \nu_m, \nu'_m \in \Gamma, \ m = 0,1,\ldots\}$$

$$= \sup_{x,y\in S, m\in N} \alpha(x,y,m) = (0 \text{ or } 2)$$

(ii) $\sup_{x,y \in S, m \in N} \alpha(x,y,m) = 0$ is a necessary and sufficient condition for the triviality of $\mathcal{T}$ with respect to any probability measure P_ν^m, $m = 0,1,\ldots$ and $\nu \in \Gamma$.

PROOF. We shall apply Theorem 4 to the nonhomogeneous Markov chain assuming the probability measure P_x^m and take $\mu = \delta(y)$. Thus if $\mathcal{T}$ is trivial with respect to P_x^m

$$\lim_{n\to\infty} \tilde{K}_x(y,X_n) = 1 \qquad P_x \text{ a.s.} \tag{4.1}$$

But

$$\sup_{A \in \mathcal{B}} |P^{m,n}(x,A) - P^{m,n}(y,A) \leq \sup_{A \in \mathcal{B}} [\int_A |1 - \tilde{K}_x(y,z)| P_x(X_n \in dz) \tag{4.2}$$

$$+ \int_A |\tilde{s}_y(X_n \in dz)|] .$$

(4.1) together with Theorem 4(ii) can be used in (4.2) to yield

$$\lim_{n\to\infty} \| P^{m,n}(x,\cdot) - P^{m,n}(y,\cdot) \| = 0 .$$

Thus $\alpha(x,y,m) = 0$ for $x,y \in S$ and $m \in N$. Notice now that

$$P_\nu^{m,n}(A) - P_{\nu'}^{m,n}(A) = \int P_x^{m,n}(A)\nu(dx) - \int P_y^{m,n}(A)\nu'(dy)$$

$$= \int [\int (P_x^{m,n}(A) - P_y^{m,n}(A))\nu(dx)]\nu'(dy)$$

which entails

$$\lim_{n\to\infty} \| P_\nu^{m,n} - P_{\nu'}^{m,n} \| \leq \lim_{n\to\infty} \int [\int \| P_x^{m,n} - P_y^{m,n} \| \nu(dx)]\nu'(dy) = 0$$

and the first part of the Theorem is proved.

Suppose now that there exists a probability measure ν such that $\mathcal{T}$ is not trivial with respect to P_ν. Then there would exist two disjoint sets in $\mathcal{T}$, say T_1 and T_2 such that $P_\nu(T_1) > 0$, $P_\nu(T_2) > 0$ and $P_\nu(T_1 \cup T_2) = 1$. Further, by the martingale convergence theorem (see [31]) $\lim_{n\to\infty} P_\nu(T_2 | X_2) = 1_{\Lambda_2}$ P_ν a.s.. Assume now that ε is a number with $0 < \varepsilon < 1$ and denote $B_n^1 = \{x : P(T_1 | X_n = x) > 1 - \frac{\varepsilon}{2}\}$ and $B_n^2 = \{x : P(T_2 | X_n = x) > 1 - \frac{\varepsilon}{2}\}$. Then, we can easily check that B_n^1 and B_n^2 are disjoint for all n and that $\lim_{n\to\infty} \{X_n \varepsilon B_n^1\} = T_1$ P_ν a.s., $\lim_{n\to\infty}\{X_n \varepsilon B_n^2\} = T_2$ P_ν a.s.. Since $\lim_{n\to\infty} P(X_n \varepsilon B_n^1 | X_m = x) = P(T_1 | X_m = x)$, $\lim_{n\to\infty} P(X_n \varepsilon B_n^2 | X_m = x) = P(T_2 | X_m = x)$ for all $x \varepsilon S$ we get for $x \varepsilon B_m^1$ and $y \varepsilon B_m^2$

$$\begin{aligned}\lim_{n\to\infty} \| P^{m,n}(x,\cdot) - P^{m,n}(y,\cdot) \| &\geq \limsup_{n\to\infty} (P^{m,n}(x,B_n^1) - P^{m,n}(y,B_n^1)) \\ &\quad + \limsup_{n\to\infty} (P^{m,n}(y,B_n^2) - P^{m,n}(x,B_n^2)) \\ &\geq 2 - \varepsilon\end{aligned}$$

and the proof is done.

As a corollary, we get the following "0 - 2 law" for homogeneous Markov chains

COROLLARY. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain and denote

$$\beta(x,y) = \lim_{n\to\infty} \| P^n(x,\cdot) - P^n(y,\cdot) \| .$$

Then

(i) $\sup\{\lim_{n\to\infty} \| P_\nu^n - P_{\nu'}^n \| ; \nu, \nu' \varepsilon \Gamma\} = \sup_{x,y \varepsilon S} \beta(x,y) = (0 \text{ or } 2)$

(ii) $\sup_{x,y\varepsilon S} \beta(x,y) = 0$ is necessary and sufficient condition for the triviality of $\mathcal{T}$ with respect to any probability measure P_ν with $\nu \varepsilon \Gamma$.

The equivalence between $\sup_{x,y\varepsilon S} \beta(x,y) = 0$ and the triviality of $\mathcal{T}$ under any initial distribution was given by Jamieson and Orey [27] generalizing a result due to Blackwell and Freedman [6] , see also [20] and [35]. For an extension to a continuous parameter chain see Duflo and Revuz [18], and to the nonhomogeneous Markov chains see Iosifescu [25], [26]. The remaining part of the Corollary is due to Derriennic [14] who used a combined martingale and operator theory* approach to prove the entire Corollary. The proof given here is new.

Let $\mathcal{G}$ and $\mathcal{H}$ be two sub σ-fields of $\mathcal{F}$ such that $\mathcal{G} \subseteq \mathcal{H}$. We shall say that $\mathcal{G} = \mathcal{H}$ P_ν a.s. if for any set Λ in $\mathcal{H}$ there exists a set Λ' in $\mathcal{G}$ such that $P_\nu(\Lambda \,\Delta\, \Lambda') = 0$.

The following "0 - 2 law" gives a criterion for $\mathcal{J} = \mathcal{T}$ P_ν a.s. for any $\nu \in \Gamma$.

THEOREM 6. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain and denote

$$\gamma(x) = \lim_{n\to\infty} \| P^{(n)}(x, \cdot) - P^{(n+1)}(x, \cdot) \| .$$

Then

(i) $$\sup\{\lim_{n\to\infty} \| P_\nu^{(n)} - P_\nu^{(n+1)} \| , \nu \in \Gamma\} = \sup_{x\in S} \gamma(x) = (0 \text{ or } 2) .$$

(ii) $\sup_{x\in S} \gamma(x) = 0$ is a necessary and sufficient condition for $\mathcal{J} = \mathcal{T}$ a.s. with respect to any starting measure ν with $\nu \in \Gamma$.

PROOF. If we proved that $\lim_{n\to\infty} \tilde{K}_x(P_x, X_n) = 1$ P_x a.s. for any $x \in S$, then as in the proof of Theorem 5 we can show that $\gamma(x) = 0$ for any $x \in S$. Also to prove that we can replace x by an arbitrary measure ν we can proceed as in the proof of Theorem 5.

---

* For basic methods of the operator theory pertinent to Markov chains see Foguel [19].

Suppose that $\lim_{n\to\infty} \tilde{K}_x(P_x, X_n) \neq 1$ P_x a.s. for some x and denote

$\Lambda = \{\omega : \left.\frac{dP_{P_x}}{dP_x}\right|_{\mathcal{T}} > 1\}$. Then, according to Theorem 3 we have $P_x(\Lambda) > 0$ and $P_{P_x}(\Lambda) > P_x(\Lambda)$ which entails $P_x(\theta^{-1}\Lambda) > P_x(\Lambda)$. But $\mathcal{J} = \mathcal{T}$ P_ν a.s. for all $\nu \in \Gamma$ and therefore there exists an invariant set Λ' such that $P_x(\Lambda \Delta \Lambda') = P_{P_x}(\Lambda \Delta \Lambda') = 0$. Thus $P_{P_x}(\Lambda) = P_x(\theta^{-1}\Lambda) = P_x(\Lambda)$ and we have got a contradiction that proves the "0" part of the theorem.

Suppose now that there exist a starting probability ν and a set Λ in $\mathcal{T}$ such that $P_\nu(\Lambda) > 0$ and $P_\nu(\Lambda \Delta \theta^{-1}\Lambda) > 0$. We assume without loss of generality that Λ and θ^{-1} are disjoint, since otherwise in view of Proposition 2, §2.2 we can arrange to have such a situation by taking $\Lambda \cap (\theta^{-1}\Lambda)^c$ instead of Λ. Suppose now that we choose a number ε with $0 < \varepsilon < 1$ and denote $A_n = \{x : P(\Lambda | X_n = x) > 1 - \frac{\varepsilon}{2}\}$. Then by an already familiar reasoning $\lim_{n\to\infty}\{X_n \in A_n\} = \Lambda$ P_ν a.s. . Further since $P(\theta^{-1}\Lambda | X_{n+1} = x) = P(\Lambda | X_n = x)$ and in view of the disjointness of Λ and $\theta^{-1}\Lambda$ one must have $A_n \cap A_{n+1} = \emptyset$ for all n. Finally, as in the proof of Theorem 5 we get

$$\begin{aligned}\lim_{n\to\infty} \| P^{(n)}(x, \cdot) - P^{(n+1)}(x, \cdot) \| &\geq \lim_{n\to\infty}(P^{(n)}(x, A_n) - P^{(n+1)}(x, A_n)) \\ &\quad + \lim_{n\to\infty}(P^{(n+1)}(x, A_{n+1}) - P^{(n)}(x, A_{n+1})) \\ &\geq 2 - \varepsilon\end{aligned}$$

and the proof is complete.

A result of the type of Theorem 6, called "0 - 2 law" was first given by Ornstein and Sucheston [37]. Theorem 6 was given by Derriennic [14]. A related result was obtained independently by McDonald [32]. The proof given here is new.

There is yet another "0 - 2 law" due to Derriennic [14] which gives a criterion for the triviality of $\mathcal{J}$ with respect to P_ν for any $\nu \in \Gamma$; namely

THEOREM 7. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain and denote

$$\delta(x,y) = \lim_{n\to\infty} \frac{1}{n} \left\| \sum_{i=1}^{n} (P^{(i)}(x,\cdot) - P^{(i)}(y,\cdot)) \right\| .$$

Then

(i) $\sup\{\lim \frac{1}{n} \| \sum_{i=1}^{n} (P_\nu^{(i)} - P_{\nu'}^{(i)}) \| ; \nu , \nu' \in \Gamma\} = \sup_{x,y\in S} \delta(x,y) = \{0 \text{ or } 2\}$

(ii) $\sup_{x,y\in S} \delta(x,y) = 0$ is a necessary and sufficient condition for the triviality of $\mathcal{J}$ with respect to any probability measure P_ν with $\nu \in \Gamma$.

At first sight, the assertion of Theorem 7 seems unexpected, since unlike Theorems 5 and 6, the total variation property appearing in it does not look like a consequence of a previously given almost sure convergence result. However, we shall now see that Theorem 7 is related to Theorem 1, as well as to an almost sure convergence property based on Chacon-Ornstein ergodic theorem, established by Derriennic in [13].

Theorem 1 was given under the assumption that g was a regular kernel. However, under more general conditions (see Derriennic [13]), it can be shown that if we denote $g_\nu^{(n)}(A) = \sum_{i=o}^{n} P_\nu(X_i \in A)$, $g_\mu^{(n)}(A) = \sum_{i=o}^{n} P_\mu(X_i \in A)$ and write

$$g_\mu^{(n)}(dy) = K_\nu^n(\mu,y) g_\nu^{(n)}(dy) + s_\mu^{(n)}(dy)$$

for the Lebesgue decomposition of $g_\mu^{(n)}$ with respect to $g_\nu^{(n)}$, then $\lim_{n\to\infty} K_\nu(\mu,X_n) = \lim_{n\to\infty} K_\nu^n(\mu,X_n)$ P_ν a.s. . If $\mathcal{J}$ is further assumed to be trivial with respect to P_ν , then $\lim_{n\to\infty} K_\nu^n(\mu,X_n) = 1$ P_ν a.s. and

$$\lim_{n\to\infty} \frac{1}{n} \left\| \sum_{i=1}^{n} (g_\mu^{(i)} - g_\nu^{(i)}) \right\| = \lim_{n\to\infty} \frac{2}{n} \sup_{A\varepsilon B} (g_\mu^{(n)}(A) - g_\nu^{(n)}(A))$$

$$= \lim_{n\to\infty} \int |K_\nu^{(n)}(\mu,y) - 1| \frac{g_\nu^{(n)}(dy)}{n} \quad .$$

Here $\frac{g_\nu^{(n)}}{n} \leqslant 1$ and since $\lim_{n\to\infty} |K_\nu^{(n)}(\mu,X_n) - 1| = 0 \quad P_\nu$ a.s. we can see that Theorem 7 is equivalent to an assertion that a certain integral converges to 0 when the integrand tends to 0 with respect to some measure P_ν .

Because $K_\nu^{(n)}(\mu,S) = \infty$ and the measure is not P_ν but $\frac{g^{(n)}}{n}$, such a result is not a consequence of a theoretical result from the Integration theory, although it is likely to be obtainable directly.

Suppose that for any x in S , $P(x,\cdot)$ is absolutely continuous with respect to a measure m , i.e. that $P(x,dy) = p(x,y)m(dy)$. Then for any $n \geqslant 1$, $P^{(n)}(x,dy) = p^{(n)}(x,y)m(dy)$ and $P_\nu(X_n \varepsilon\, dy) = p_\nu^{(n)}(y)m(dy)$, $P_\mu(X_n \varepsilon\, dy) = p_\mu^{(n)}(y)m(dy)$ where

$$p_\nu^{(n)}(y) = \int p^{(n)}(x,y)\nu(dx)$$

and

$$p_\mu^{(n)}(y) = \int p^{(n)}(x,y)\mu(dx) \ .$$

Denote $\bar{K}_\nu^n(\mu,y) = \sum_{i=1}^{n} p_\mu^{(i)}(y)/p_\nu^{(i)}(y)$. Then, it is easy to see that $\bar{K}_\nu^n(\mu,y) = K_\nu^n(\mu,y)$.

In [13] Derriennic has proved a general result which can be applied to $\bar{K}_\nu^n(\mu,y)$ to yield $\lim_{n\to\infty} \bar{K}_\nu^n(\mu,X_n) = 1$ if $\mathcal{J}$ is trivial with respect to P_ν . The conditions of Derriennic's result include both instances of the dissipative case (when g is a regular kernel) as well as of the conservative case (when g is infinite with positive probability). Thus, such results are connected with Theorem 7.

There is another aspect, worthwhile to be mentioned, in relation to $\bar{K}^n_\nu(\mu,y)$. If we apply the space-time chains considerations to $\bar{K}^n_\nu(\mu,y)$ in the same way as we did in the proof of Theorem 3, we get that

$$\lim_{n\to\infty} \frac{p_\mu^{(n)}(X_n)}{p_\nu^{(n)}(X_n)} = 1 \quad P_\nu \text{ a.s. if and only if } \mathfrak{T} \text{ is trivial with respect to } P_\nu .$$

Then, as in Theorem 6 this can lead to a "0 - 2 law" for expressions like $\| P^{(n)}(x,\cdot) - P^{(n+1)}(x,\cdot)\|$ under the assumption that the densities $\{p(x,y) , x \in S\}$ exist. In other words the Chacon-Ornstein ergodic theorem [7] used in [13] turns out to imply, in the long run, the Ornstein-Sucheston "0 - 2 law" [37]!

5. STRUCTURE RESULTS FOR ASYMPTOTIC σ-FIELDS.

We shall consider the vector chain $\{\hat{X}_n : n \geq 0\}$ with $X_n = (X_n^{(1)}, X_n^{(2)})$, $n \geq 0$ defined on the probability space $(\tilde{\Omega}, \tilde{\mathfrak{F}}, \hat{P})$, $\{X_n^{(1)} : n \geq 0\}$ and $\{X_n^{(2)} : n \geq 0\}$ being two independent copies of $\{X_n : n \geq 0\}$ and $\hat{P} = P_\nu \otimes P_\nu$.

Throughtout this section we shall assume that $\mathfrak{B}$ is separable, i.e. is generated by a countable collection of sets $A_1, A_2, \ldots$.

Let f_n be a real measurable function of 2n variables and define for any $x,y \in S$ the set function

$$\psi_n(x,y;A) = f_n(P(x,A),\ldots,P^{(n)}(x,A),P(y,A),\ldots,P^{(n)}(y,A)) .$$

In what follows we shall need the following Lemma (see e.g. [37]).

LEMMA. For any $n = 0,1,\ldots$ the total variation of ψ_n, $\|\psi_n\|$ is a measurable function with respect to $\mathfrak{B} \otimes \mathfrak{B}$.

PROOF. Denote by $\mathfrak{B}_k$ the σ-field generated by the sets $A_1, A_2, \ldots, A_k$, $k = 1,2,\ldots$ and by $\psi_n^k(x,y\,;\cdot)$ the restriction of $\psi_n(x,y:\cdot)$ to the σ-field

$\mathcal{B}_k$. By a version of martingale convergence theorem ([31])

$$\lim_{k\to\infty} \frac{d\psi_n^k(x,y;\cdot)}{d|\psi_n(x,y;\cdot)|} = \frac{d\psi_n(x,y;\cdot)}{d|\psi_n(x,y;\cdot)|}$$

a.s. with respect to $|\psi_n(x,y;\cdot)|$ and $L_1(|\psi_n(x,y;\cdot)|)$. Hence, the sequence of measurable functions

$$\|\psi_n^k(x,y;\cdot)\| = \int \left| \frac{d\psi_n^k(x,y;\cdot)}{d|\psi_n(x,y;\cdot)|} \right| d|\psi_n(x,y;\cdot)|$$

converges as $k \to \infty$ to

$$\int \left| \frac{d\psi_n(x,y;\cdot)}{d|\psi_n(x,y;\cdot)|} \right| d|\psi_n(x,y;\cdot)| = \|\psi_n(x,y;\cdot)\|$$

which is therefore measurable.

Let us now consider, for any nonnegative integers m and n , the random variable

$$\alpha_{m,n}(\tilde{\omega}) = \| P^{m,n}(X_m^{(1)},\cdot) - P^{m,n}(X_m^{(2)},\cdot)\|$$

and denote by T_o the completely nonatomic set and by $T_1, T_2, \ldots$ the atomic sets occurring in the P_ν representation of Ω corresponding to $\mathcal{T}$.

Consider the following condition

CONDITION (C). Let $\{X_n : n \geq 0\}$ be a nonhomogeneous Markov chain. Then for any $m = 0,1,\ldots$ and $u = 1,2,\ldots$ $\nu_m\{(x,y)|T_u$ is either atomic for $\mathcal{T}$ or a null set with respect to $P^m_{\frac{1}{2}(\delta(x)+\delta(y))}\} = 1$.

REMARK. The separability of $\mathcal{B}$ as well as Condition (C) are satisfied by any countable nonhomogeneous Markov chain. Indeed, the separability of $\mathcal{B}$ is trivially satisfied, whereas Condition (C) is a consequence of the absolute continuity of the measures $P^m_{\frac{1}{2}(\delta(x)+\delta(y))}$ with respect to $\nu^{(m)}$ whenever $\nu^{(m)}(x) > 0$ and $\nu^{(m)}(y) > 0$.

The following result will give a characterization of the completely nonatomic and atomic sets of $\mathcal{T}$ by means of $\{\alpha_{m,n}(\tilde{\omega})\}$.

THEOREM 1. Suppose that $\{X_n : n \geq 0\}$ is a nonhomogeneous Markov chain satisfying condition (C) and $\mathcal{B}$ is separable. Then

(i) There exist the limits

$$\lim_{n\to\infty} \alpha_{m,n}(\tilde{\omega}) = \alpha_m(\tilde{\omega})$$

$$\lim_{n\to\infty} \alpha_m(\tilde{\omega}) = \alpha(\tilde{\omega}) \qquad \hat{P}_\nu \quad \text{a.s.}$$

(ii) $\alpha(\tilde{\omega}) = 2$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \in T_o \times T_o \cup \bigcup_{u \neq u'} T_u \times T_{u'}$,

$\alpha(\tilde{\omega}) = 0$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \in T_u \times T_u$, $u = 1,2,\ldots$

PROOF. We shall first show that for any fixed m, x and y, $\| P^{m,n}(x,\cdot) - P^{m,n}(y,\cdot)\|$ is nondecreasing with respect to n. Indeed, if we denote $\tilde{\mathcal{B}}' = \{A \times B \mid A \in \mathcal{B},\ B \in \mathcal{B}\}$, then

$$\| P^{m,n+1}(x,\cdot) - P^{m,n+1}(y,\cdot)\| \leq 2 \sup_{A \in \tilde{\mathcal{B}}'} |P_x^m((X_n, X_{n+1}) \in A) - P_y^m((X_n, X_{n+1}) \in A)|$$

$$\leq 2 \sup_{\tilde{A} \in \tilde{\mathcal{B}}'} [\,|\int_A P^{m,n}(x,dt)\int_B P^{n,n+1}(t,dz) - \int_A P^{m,n}(y,dt)\int_B P^{n,n+1}(t,dz)|\,]$$

$$\leq 2 \sup_{\tilde{A} \in \tilde{\mathcal{B}}'} [P^{m,n}(x,A) - P^{m,n}(y,A)] = \| P^{m,n}(x,\cdot) - P^{m,n}(y,\cdot)\| .$$

This implies that $\lim_{n\to\infty} \alpha_{m,n}(\tilde{\omega}) = \alpha_m(\tilde{\omega})$ exists for all $\tilde{\omega}$. The existence of $\alpha(\tilde{\omega})$ will be proved in the course of the proof of (ii).

Since T_o is completely nonatomic, for any $\varepsilon > 0$ we can find $n(\varepsilon)$ disjoint sets $T(1),\ldots,T(n(\varepsilon))$ in $\mathcal{T}$ such that $T_o = T(1) \cup T(2) \cup \ldots \cup T(n(\varepsilon))$ and $0 < P_\nu(T(s)) < \varepsilon/4$ for $1 \leq s \leq n(\varepsilon)$, (see e.g. [38] p. 81). Let $B_n(s) = \{x : P(T(s) \mid X_n = s > 1 - \varepsilon/4\}$. As we have seen before (in the proof of Theorem 2, §3) we can get that $\lim_{n\to\infty}\{X_n \in B_n(s)\} = T(s)$ a.s. with respect to P_ν

for $s = 1,2,\ldots,n(\varepsilon)$. Since for m sufficiently large $B_m(s)$ and $B_m(s')$ with $s \neq s'$ are not empty we can find $x \in B_m(s)$ and $y \in B_m(s')$ such that

$$\begin{aligned}
&\lim_{n\to\infty} \| P^{m,n}(x,\cdot) - P^{m,n}(y,\cdot) \| \geq \\
&\lim_{n\to\infty} (P^{m,n}(x,B_n(s)) - P^{m,n}(y,B_n(s))) + \\
&\lim_{n\to\infty} (P^{m,n}(y,B_n(s')) - P^{m,n}(x,B_n(s'))) \\
= \ & P(T(s) | X_m = x) - P(T(s) | X_m = y) + \\
&P(T(s') | X_m = y) - P(T(s') | X_m = x) \\
\geq \ & 1 - {}^{\varepsilon}/4 - {}^{\varepsilon}/4 + 1 - {}^{\varepsilon}/4 - {}^{\varepsilon}/4 \\
= \ & 2 - \varepsilon \ .
\end{aligned}$$

Since by the above Lemma $\{\alpha_{m,n}(\tilde{\omega})\}$ are random variables, $\liminf_{m\to\infty} \alpha_m(\tilde{\omega})$ is also a random variable and $\liminf_{n\to\infty} \alpha_m(\tilde{\omega}) \geq 2 - \varepsilon$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \in I(s) \times I(s')$. However, we can split each of the sets $T(s)$, $s = 1,2,\ldots,n(\varepsilon)$ into disjoint subsets whose probabilities are smaller that ${}^{\varepsilon'}/4$ for any preassigned ε' smaller than ε . Using the same reasoning as above we get, in particular, that $\liminf_{n\to\infty} \alpha_n(\tilde{\omega}) > 2 - \varepsilon'$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \in T(s) \times T(s')$ and because we can apply this to any subset of $T(s)$, $s = 1,\ldots,n(\varepsilon)$ we deduce that $\alpha(\tilde{\omega}) = 2$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \in T_o \times T_o$.

The proof of $\alpha(\tilde{\omega}) = 2$ for P_ν almost all $\tilde{\omega} \in \bigcup_{u\neq u'} T_u \times T_{u'}$ is easier and will be left to the reader as an exercise.

We shall now prove that $\alpha(\tilde{\omega}) = 0$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \in T_u \times T_u$, $u = 1,2,\ldots$. Denote $\hat{A}_m = \{(x,y) | T_u$ is atomic in $\mathcal{T}$ or null, with respect to $P^m_{\frac{1}{2}(\delta(x) + \delta(y))}\}$ and notice that for any $(x,y) \in \hat{A}_m$, $dP^m_x / dP^m_y \Big|_{T_u} = c_n(x,y)$ is a constant and $0 < c_m(x,y) < \infty$. Indeed T_u must be atomic or null with respect to P_x^m and P_y^m since these two probability measures are absolutely continuous with respect to $P^m_{\frac{1}{2}(\delta(x)+\delta(y))}$. Further, if $c_m(x,y)$ were 0 or ∞ P_x^m and P_y^m would be singular on T_u, i.e. there would exist a set $H \subset T_u$ with $H \varepsilon \mathcal{T}$ such that $P_x^m(H) = 0$ and $P_y^m(H) = 1$. But in such a case H and H^c would be sets in $\mathcal{T}$ with $P^m_{\frac{1}{2}(\delta(x)+\delta(y))}(H) > 0$ and $P^m_{\frac{1}{2}(\delta(x)+\delta(y))}(H^c) > 0$ which contradicts the assumption that $(x,y) \varepsilon \hat{A}_m$. Thus $c_m(x,y) = \left.\frac{dP_x^m}{dP_y^m}\right|_{T_u} = \frac{P_x^m(T_u)}{P_y^m(T_u)}$. Define further $A_n(u) = \{x : P(T_u | X_n = x) > 1 - \varepsilon\}$ and take $(x,y) \varepsilon (A_n(u) \times A_n(u)) \cap \hat{A}_n$. Making use of Theorem 3, we get

$$\lim_{n\to\infty} \tilde{K}_y^m(x,X_n) = c_m(x,y) \qquad P_y^m \text{ a.s. on } T_u . \tag{5.1}$$

Since for any A in

$$P^{m,n}(x,A) - P^{m,n}(y,A) \leqslant \int_A |1 - \tilde{K}_y^m(x,z)| P_y(X_n \varepsilon\, dz)$$

we are led to the inequalities

$$\begin{aligned} &\limsup_{n\to\infty} \sup_{A\varepsilon\mathcal{B}} (P^{m,n}(x,A) - P^{m,n}(y,A)) \\ &\quad \lim_{n\to\infty} \int_{A_n(u)} |1 - \tilde{K}_y^m(x,X_n)| P_y(X_n \varepsilon\, dz) \\ &\qquad + \limsup_{n\to\infty} P^{m,n}(x, A_n^c(u)) \\ &\qquad = |1 - c_m(x,y)| P(T_u) + \varepsilon . \end{aligned} \tag{5.2}$$

If we take into account that $P_x^m(T_u) > 1 - \varepsilon$, $P_y^m(T_u) > 1 - \varepsilon$ and $1 - \varepsilon < c_m(x,y) < \frac{1}{1-\varepsilon}$, (5.2) yields

$$\lim_{n\to\infty} \frac{1}{2} \| P^{m,n}(x,\cdot) - P^{m,n}(y,\cdot) \| \leqslant \max(\varepsilon, \frac{\varepsilon}{1-\varepsilon}) P(T_u) + \varepsilon .$$

Since the quantity on the right side of (5.1) can be made arbitrarily small by choosing ε sufficiently small we get $\alpha(\tilde{\omega}) = 0$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \varepsilon T_u \times T_u$, $u = 1,2,\ldots$ and the proof is complete.

A characterization of completely nonatomic sets of $\mathcal{T}$ for a countable chain was given in [10].

Define now for any nonnegative integers m and n, the random variable

$$\beta_{m,n}(\omega) = 1 - \frac{1}{2} \| \nu^{(n)}(\cdot) - P^{m,n}(X_n, \cdot) \| .$$

The following result gives a characterization of the tail σ-field structure by means of the sequence $\{\beta_{m,n}(\omega)\}$.

THEOREM 2. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain satisfying Condition (C) and $\mathfrak{B}$ is separable. Then

(i) There exists the limits

$$\lim_{n\to\infty} \beta_{m,n}(\omega) = \beta_m(\omega)$$

$$\lim_{m\to\infty} \beta_m(\omega) = \beta(\omega) \qquad P_\nu \quad \text{a.s.}$$

(ii)
$$\beta(\omega) = 0 \quad \text{for } P_\nu \text{ almost all } \omega \varepsilon T_o$$

$$\beta(\omega) = P(T_u) \text{ for } P_\nu \text{ almost all } \omega \varepsilon T_u \ , \ u \geq 1 .$$

PROOF. As in the proof of Theorem 1 we can use simple inequalities to prove that $\|\nu^{(n)}(\cdot) - P^{m,n}(x, \cdot)\|$ is nondecreasing with respect to n . Indeed, we can start off by writing

$$\|\nu^{(n+1)}(\cdot) - P^{m,n+1}(x, \cdot)\| \leq \sup_{\tilde{A}\varepsilon\mathfrak{B}'} |P_\nu((X_n,X_{n+1}) \varepsilon A) - P_x^m((X_n,X_{n+1}) \varepsilon A)|$$

$$\leq [|\int_A \nu^{(n)}(dy) \int_B P^{n,n+1}(y,dz) - \int_A P^{m,n}(x,dy) \int_B P^{n,n+1}(y,dz)|]$$

and complete the proof in the same way as in the proof of Theorem 1. Here $\tilde{A}$ and $\mathfrak{B}'$ are the same as defined in the proof of Theorem 1. Thus $\lim_{n\to\infty} \beta_{m,n}(\omega) = \beta_m(\omega)$ exists for all $\omega \varepsilon \Omega$. The existence of $\beta(\omega)$ will be proved in the course of the proof of (ii).

Notice now that we can write

$$\nu^{(n)}(A) - P^{m,n}(x,A) = \int (P^{m,n}(y,A) - P^{m,n}(x,A))\nu^{(m)}(dy)$$

for any A in $\mathcal{B}$. Thus

$$\beta_{m,n}(\omega) = 1 - \sup_{A\varepsilon\mathcal{B}} \int (P^{m,n}(y,A) - P^{m,n}(x,A))\nu^{(m)}(dy) \;.$$

Choose further a sequence $\{B_n(u) : n \geq 0\}$ such that $\lim_{n\to\infty}\{X_n \varepsilon B_n(u)\} = T_u$ P_ν a.s. for $u = 0,1,\ldots$ and write $\beta_{m,n}$ as

$$\beta_{m,n}(\omega) = 1 - \sup_{A\varepsilon\mathcal{B}} [\int_{B_m(u)} (P^{m,n}(y,A) - P^{m,n}(X_m,A))\nu^{(m)}(dy)$$

$$+ \int_{B_m^c(u)} (P^{m,n}(y,A) - P^{m,n}(X_m,A))\nu^{(m)}(dy)] \;.$$

If we take into consideration that

$$\lim_{m\to\infty}\lim_{n\to\infty} \int_{B_m(u)} P^{m,n}(y,B_n^{(c)}(u))\nu^{(m)}(dy) = \lim_{m\to\infty}\lim_{n\to\infty} \int_{B_m^c(u)} P^{m,n}(y,B_n(u))\nu^{(m)}(dy)$$

$$= 0$$

we get that

$$\beta(\omega) = \lim_{m\to\infty}\lim_{n\to\infty} \beta_{m,n}(\omega) = 1 - \lim_{m\to\infty}\lim_{n\to\infty} \sup_{A\subset B_m(u)} \int_{B_m(u)} (P^{m,n}(y,A) - P^{m,n}(X_m,A))\nu^{(m)}(dy)$$

$$- \lim_{m\to\infty}\lim_{n\to\infty} \sup_{A\subset B_m^c(u)} \int_{B_m^c(u)} (P^{m,n}(y,A) - P^{m,n}(X_m,A))\nu^{(m)}(dy) \tag{5.3}$$

P_ν a.s. provided that the limits appearing on the right side of (5.3) exist P_ν a.s. . But a scrutiny of the proof of Theorem 1 above reveals that these limits exist and that

$$\lim_{m\to\infty}\lim_{n\to\infty}\sup_{A\subset B_m(u)}\int_{B_m(u)}(P^{m,n}(y,A)-P^{m,n}(X_m,A))\nu^{(m)}(dy)$$

$$=\begin{cases}P_\nu(T_u) & P_\nu \text{ a.s. if } u=0\\ \\ 0 & \text{if } u\geq 1\end{cases}$$

and

$$\lim_{m\to\infty}\lim_{n\to\infty}\sup_{A\subset B_m^c(u)}\int_{B_m^c(u)}(P^{m,n}(y,A)-P^{m,n}(X_n,A))\nu^{(m)}(dy)=1-P_\nu(T_u)$$

P_ν a.s. for $u=0,1,\ldots$ and the proof is complete.

A result of the type of Theorem 2 as well as the Corollaries given below were given by Cohn [9], the transition probabilities appearing in it being replaced by the backward transition probabilities. Griffeath has used a "coupling" method in [21] to show that a dual result can be obtained using instead the forward transition probabilities. For a martingale proof of this result see [10]. Further Griffeath has extended this result to the case when $\mathcal{B}$ is polish in [22]. Here we have used a different approach to further extend this result to the case when $\mathcal{B}$ is only assumed to be separable.

COROLLARY 1. (i) $\beta(\omega)=1$ P_ν a.s. if and only if $\mathcal{T}$ is trivial with respect to P_ν.

(ii) $\beta(\omega)>\delta$ P_ν a.s. with $0<\delta\leq\frac{1}{2}$ if and only if $\mathcal{T}$ is finite with respect to P_ν.

(iii) $\beta(\omega)>0$ P_ν a.s. and (i) and (ii) do not hold if and only if $\mathcal{T}$ is atomic with respect to P_ν.

(iv) $P_\nu(\beta(\omega)=0)>0$ if and only if $\mathcal{T}$ is nonatomic with respect to P_ν.

PROOF. The Corollary is an immediate consequence of Theorem 2 the only point needing a proof being that if $\beta(\omega) > \frac{1}{2}$ P_ν a.s., then $\beta(\omega) = 1$ P_ν a.s. Indeed, if $\beta(\omega) > \delta$ P_ν a.s. with $\delta > 0$, $\mathcal{T}$ is at most finite and in such a case there must exist an atomic set T_i such that $P_\nu(T_i) \leq \frac{1}{2}$.

Conditions for the finiteness of $\mathcal{T}$ were previously given by Bartfai and Revesz [3] and Iosifescu [26].

COROLLARY 2. If $P_\nu(\beta(\omega) > 0) > 0$ and we denote the probability distribution of β by

$$\beta : \begin{pmatrix} 0 & x_1 & x_2 & \cdots\cdots \\ \\ P_0 & P_1 & P_2 & \cdots\cdots \end{pmatrix}$$

then $\mathcal{T}$ contains P_i/x_i atomic sets having probability x_i, $i \geq 1$.

PROOF. Let us denote by $\{T_k^i ; k = 1,\ldots k_i\}$ the atomic set of $\mathcal{T}$ which have the same probability. Then by Theorem 2, $P_\nu(T_k^i) = x_i$, $k = 1,\ldots,k_i$. Thus the number k_i is equal to P_i/x_i and the proof is complete.

The following result is a straightforward consequence of Theorem 2.

COROLLARY 3. The tail σ-field of the sequence $\{X_n : n \geq 0\}$ has the same structure as the tail σ-field of any of its subsequences $\{X_{n_k} : k \geq 0\}$.

We shall further show that we can establish results analogous to Theorems 1 and 2 given above for the invariant σ-field $\mathcal{J}$ of a homogeneous Markov chain.

In analogy to Condition (C) we shall consider

CONDITION (C'). Let $\{X_n : n \geq 0\}$ be a homogeneous Markov chain. Then for any $m = 0,1,\ldots$ and $u = 1,2,\ldots$

$$\hat{\nu}_m\{(x,y) \mid I_u \text{ is either atomic for } \mathcal{J} \text{ or null with respect to } P_{\frac{1}{2}(\delta(x)+\delta(y))}\} = 1 .$$

Let us define, for any integers m and n, the random variable

$$\gamma_{m,n}(\tilde{\omega}) = \frac{1}{n}\left\|\sum_{i=1}^{n}(P^{(i)}(X_m^{(1)},\cdot) - P^{(i)}(X_m^{(2)},\cdot))\right\|$$

and denote by I_o the completely nonatomic set and by $I_1, I_2, \ldots$ the atomic sets occurring in the P_ν-representation of Ω corresponding to $\mathcal{J}$.

The following result will give a characterization of the completely nonatomic set and atomic sets of $\mathcal{J}$ by means of $\{\gamma_{m,n}(\tilde{\omega})\}$.

THEOREM 3. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain satisfying Condition (C') and $\mathcal{B}$ is separable. Then

(i) There exist the limits

$$\lim_{n\to\infty} \gamma_{m,n}(\tilde{\omega}) = \gamma_m(\tilde{\omega})$$

$$\lim_{m\to\infty} \gamma_m(\tilde{\omega}) = \gamma(\tilde{\omega}) \qquad P_\nu \text{ a.s.}$$

(ii) $\gamma(\tilde{\omega}) = 2$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \in I_o \times I_o \cup \bigcup_{u \neq u'} I_u \times I_{u'}$,

$\gamma(\tilde{\omega}) = 0$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \in I_u \times I_u$, $u = 1, 2, \ldots$.

PROOF. It is easy to see that $\frac{1}{n}\left\|\sum_{i=1}^{n}(P^{(i)}(x,\cdot) - P^{(i)}(y,\cdot))\right\|$ converges for any x and y fixed, as n goes to ∞ (see e.g. [14] p. 115) since if we denote $f(n) = \left\|\sum_{i=1}^{n}(P^{(i)}(x,\cdot) - P^{(i)}(y,\cdot))\right\|$ then f(n) can be shown to be a subadditive function i.e. $f(m+n) \leq f(m) + f(n)$ for all m, $n \in N$ and therefore $\lim_{n\to\infty} f(n)/_n = \inf_{n\geq 1} f(n)/_n$. Hence $\lim_{n\to\infty} \gamma_{m,n}(\tilde{\omega}) = \gamma_m(\tilde{\omega})$ exists for all $\tilde{\omega} \in \tilde{\Omega}$.

The existence of $\gamma(\tilde{\omega})$ will be proved in the course of the proof of (ii).

Since I_o is completely nonatomic, for any $\varepsilon > 0$ we can find $n(\varepsilon)$ disjoint sets $I(1), \ldots, I(n(\varepsilon))$ in $\mathcal{J}$ such that $I_o = I(1) \cup I(2) \cup \ldots \cup I(n(\varepsilon))$ and $0 < P(I(s)) < \varepsilon/4$ for $1 \leq s \leq n(\varepsilon)$. Let $C(s) = \{x : P_x(I(s)) > 1 - \varepsilon/4\}$. As we have seen in the proof of Theorem 3, §3, we get $\lim_{n\to\infty}\{X_n \in C'(s)\} = I(s) \quad P_\nu$ a.s.

for $s = 1,2,\dots,n(\varepsilon)$. It follows that $\lim_{n\to\infty} P^{(n)}(x,C(s)) = P_x(I(s))$. Take now $x \varepsilon C'(s)$ and $y \varepsilon C'(s')$ with $s \neq s'$.

Then

$$\lim_{n\to\infty} \frac{1}{n} \left\| \sum_{i=1}^{n} (P^{(i)}(x,\cdot) - P^{(i)}(y,\cdot) \right\| \geq$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (P^{(i)}(x,C(s)) - P^{(i)}(y,C(s))) +$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (P^{(i)}(y,C(s')) - P^{(i)}(x,C(s')))$$

$$\geq 2 - \varepsilon \ .$$

Thus $\liminf_{n\to\infty} \gamma_m(\tilde{\omega}) > 2 - \varepsilon$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \varepsilon I(s) \times I(s')$ and as in the proof of Theorem 1 we can conclude that $\gamma(\tilde{\omega}) = 2$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \varepsilon I_o \times I_o$.

The proof of $\gamma(\tilde{\omega}) = 2$ for $\hat{P}_\nu$ almost all $\tilde{\omega} \varepsilon \bigcup_{u\neq u'} I_u \times I_{u'}$ is easier and will be left to the reader as an exercise.

We shall now prove that $\gamma(\tilde{\omega}) = 0$ for almost all $\tilde{\omega} \varepsilon I_u \times I_u$, $u = 1,2,\dots$. According to Theorem 1 of [14] for any x , $y \varepsilon S$

$$\lim_{n\to\infty} \frac{1}{n} \left\| \sum_{i=1}^{n} (P^{(i)}(x,\cdot) - P^{(i)}(y,\cdot)) \right\| \leq 2 \sup_{\Lambda\varepsilon \mathcal{J}(x,y)} (P_x(\Lambda) - P_y(\Lambda)) \tag{5.4}$$

where $\mathcal{J}(x,y)$ is the invariant σ-field of the Markov chain assuming the starting measure $\frac{1}{2}(\delta(x) + \delta(y))$.

According to Condition (C') there is a set $\hat{A}_m$ of points (x,y) with $\hat{\nu}_m(\hat{A}_m) = 1$ such that I_u is atomic or null with respect to $P_{\frac{1}{2}(\delta(x)+\delta(y))}$ for $m = 0,1,\dots$. Further (5.4) implies

$$\lim_{n\to\infty} \frac{1}{n} \left\| \sum_{i=1}^{n} (P^{(i)}(x,\cdot) - P^{(i)}(y,\cdot)) \right\| \tag{5.5}$$

$$\leq 2|P_x(I_u) - P_y(I_u))| + 2 \sup_{\substack{\Lambda \varepsilon \mathcal{J}(x,y) \\ \Lambda \subset I_u^c}} (P_x(\Lambda) - P_y(\Lambda)) .$$

Choose now $C_u(\varepsilon) = \{x : P_x(I_u) > 1 - \varepsilon/4\}$ and take x , $y \varepsilon C_u$. Then (5.5) implies

$$\lim_{n\to\infty} \frac{1}{n} \left\| \sum_{i=1}^{n} (P^{(i)}(X_m^{(1)} = x, \cdot\,) - P^{(i)}(X_m^{(2)} = x, \cdot\,)) \right\| \leq 2\,\frac{\varepsilon}{4} + 2\,\frac{\varepsilon}{4} = \varepsilon$$

for any $(x,y) \varepsilon \hat{A}_m(u) \cap (C_u(\varepsilon) \times C_u(\varepsilon))$. But

$$\lim_{m\to\infty} P_\nu\{(X_m^{(1)}, X_m^{(2)}) \varepsilon \hat{A}_m(u) \cap (C_u(\varepsilon) \times C_u(\varepsilon))\} = P_\nu^2(I_u)$$

and the proof is complete.

Define now, for any nonnegative integers m and n , the random variable

$$\delta_{m,n}(\omega) = 1 - \frac{1}{2n} \left\| \sum_{i=1}^{n} (\nu^{(i+m)}(\cdot) - P^{(i)}(X_m, \cdot\,)) \right\| .$$

The following result gives a characterization of the invariant σ-field structure by means of the sequence $\{\delta_{m,n}(\omega)\}$.

THEOREM 4. Suppose that $\{X_n : n \geq 0\}$ is a homogeneous Markov chain satisfying Condition (C') and $\mathcal{B}$ is separable. Then

(i) There exist the limits

$$\lim_{n\to\infty} \delta_{m,n}(\omega) = \delta_m(\omega)$$

$$\lim_{m\to\infty} \delta_m(\omega) = \delta(\omega) \qquad P_\nu \text{ a.s.}$$

(ii)

$$\delta(\omega) = 0 \qquad \text{for } P_\nu \text{ almost all } \omega \varepsilon I_o$$

$$\delta(\omega) = P(I_u) \text{ for } P_\nu \text{ almost all } \omega \varepsilon I_u \text{ , } u = 1,2,\ldots$$

PROOF. As in the proof of Theorem 3 the convergence of the quantity $\frac{1}{n} \left\| \sum_{i=1}^{n} (\nu^{(i+m)}(\cdot) - P^{(i)}(x, \cdot\,) \right\|$ for any x fixed is a consequence of the subadditivity of $\left\| \sum_{i=1}^{n} (\nu^{(i+m)}(\cdot) - P^{(i)}(x, \cdot\,)) \right\|$ which can be shown to be entailed by the triangle inequality and the inequality

$$\Big\| \sum_{i=2}^{n+1} (\nu^{(i+m)}(\cdot) - P^{(i)}(x,\cdot))\Big\| \leq$$

$$2 \sup_{\tilde{A}\in \mathcal{B}'} \sum_{i=1}^{n} [P_\nu((X_{i+m},X_{i+m+1}) \in A) - P_\nu((X_{i+m},X_{i+m+1}) \in A | X_m = x)]$$

$$\leq \Big\| \sum_{i=1}^{n} (\nu^{(i+m)}(\cdot) - P^{(i)}(x,\cdot))\Big\|$$

has a proof analogous to the parallel result for $\{\beta_{m,n}\}$ given at the beginning of the proof of Theorem 2. In rest the proof can be carried out as in the case of Theorem 2, if one takes into account the changes occurring when considering invariant sets rather than tail sets which were manifest in the proof of Theorem 3. Theorems 3 and 4 extend the results established for countable chains in [11].

COROLLARY 1. (i) $\delta(\omega) = 1$ P_ν a.s. if and only if $\mathcal{J}$ is trivial with respect to P_ν.

(ii) $\delta(\omega) > \delta$ P_ν a.s. with $0 < \delta \leq \frac{1}{2}$ if and only if $\mathcal{J}$ is finite to P_ν.

(iii) $\delta(\omega) > 0$ P_ν a.s. and (i) and (ii) do not hold if and only if $\mathcal{J}$ is atomic with respect to P_ν.

(iv) $P_\nu(\delta(\omega) = 0) > 0$ if and only if $\mathcal{J}$ is nonatomic with respect to P_ν.

COROLLARY 2. If $P_\nu(\delta(\omega) > 0) > 0$ and we denote the probability distribution of δ by

$$\delta : \begin{pmatrix} 0 & x_1 & x_2 \cdots\cdots \\ P_0 & P_1 & P_2 \cdots\cdots \end{pmatrix}$$

then $\mathcal{J}$ contains P_i/x_i atomic sets with respect to P_ν having probability x_i,

$i \geq 1$.

The proof of these Corollaries can be carried out as those of the corresponding Corollaries after Theorem 2.

We notice that Corollary 3 after Theorem 2 does not have a counterpart for invariant σ-fields following from Theorem 4. In fact a result of this type does not hold for the invariant σ-field.

We shall next consider another characterization of the atomic sets of the tail σ-field of a homogeneous Markov chain.

Suppose that T is an atomic set of $\mathcal{T}$. Then for any $k \geq 1$, $P_\nu(T \cap \theta^k T)$ is either 0 or equal to $P_\nu(T)$. We are therefore led to consider the quantity d defined as

$$d = \inf\{k > 0 : \theta^k T = T \quad \text{a.s. with respect to } P_\nu\}$$

and agree to take $d = \infty$ if there is no integer k such that $\theta^k T = T \quad P_\nu$ a.s. d will be said to be the asymptotic period of the set T .

THEOREM 5. Suppose that $\{X_n : n \geq 0\}$ is a properly homogeneous Markov chain and d the asymptotic period of the atomic set T of $\mathcal{T}$. Then

(i) If $d = \infty$, then there exists a sequence $\{B_n : n \geq 0\}$ of mutually disjoint sets of $\mathcal{B}$ such that $\lim_{n\to\infty}\{X_n \in B_n\} = T \quad P_\nu$ a.s.

(ii) If $d < \infty$, then there exists a sequence $\{B_n : n \geq 0\}$ of sets of $\mathcal{B}$ with $B_{nd} = B_0$, $B_{nd+1} = B_1, \ldots, B_{(n+1)d-1} = B_{d-1}$, for $n = 0,1,\ldots$, where $\{B_0, B_1, \ldots, B_{d-1}\}$ are mutually disjoint sets, such that $\lim_{n\to\infty}\{X_n \in B_n\} = T \quad P_\nu$ a.s.

(iii) In either case $\bigcup_{k=o}^{d} B_k$ is an atomic set of $\mathcal{J}$.

PROOF. A consequence of Proposition 1, §2.3 is that if Λ is a set in $\mathcal{T}$ with $P_\nu(T) > 0$, then $P_\nu(\theta^k \Lambda) > 0$ for all $k \in Z$. Further if T is an atomic set of $\mathcal{T}$, then the sets $\theta^k T$ with $k \in Z$ will also be atomic sets of $\mathcal{T}$. Indeed

$P_\nu(\theta^k T) > 0$ for all $k \in Z$ and if we suppose that $\theta^k T$ for a certain k is not atomic then there would exist two disjoint sets in $\mathcal{T}$, say, T' and T'', $P_\nu(T') > 0$, $P_\nu(T'') > 0$ and $\theta^k T = T' \cup T''$. But according to Proposition 3, § 2.2 we get $\theta^{-k}(\theta^k T) = T = \theta^{-k} T' \cup \theta^{-k} T''$. Since θ^{-k} preserves the disjointness of sets and by Corollary 1 after Proposition 1, §2.3 we get $P_\nu(\theta^{-k} T') > 0$, $P_\nu(\theta^{-k} T'') > 0$ and the atomicity of $\mathcal{T}$ would be contradicted. Thus $\theta^k T$ with $k \in Z$ are all atomic sets of $\mathcal{T}$ and are either such that $P_\nu(\theta^m T \cap \theta^n T) = 0$ for $m \neq n$ with $m, n \in Z$, in which case $d = \infty$, or there exists a number k with $T = \theta^k T$ P_ν a.s., in which case $d < \infty$.

Suppose now that we choose the sets $\{B_n\}$ such that $\lim_{n\to\infty}\{X_n \in B_n\} = T$ P_ν a.s. as defined in the proof of Theorem 5, §3 i.e. $B_n = \{x : P(T\ X_n = x) > \frac{1}{2}\}$ for $n = 0, 1, \ldots$. We have $\lim_{n\to\infty}\{X_n \in B_{n+k}\} = \theta^k T$ P_ν a.s. on applying Theorem 5, §3. Since $P_\nu(T \cap \theta^k T) = 0$ we get

$$\begin{aligned} 1 \geq P_\nu(T \cup \theta^k T \mid X_n = x) &= P_\nu(T \mid X_n = x) + P(\theta^k T \mid X_n = x) \\ &= P_\nu(T \mid X_n = x) + P(T \mid X_{n+k} = x)\ . \end{aligned} \tag{5.6}$$

If we assume that there exists x in $B_n \cap B_{n+k}$ then (5.6) yields $1 \geq P(T \mid X_n = x) + P(T \mid X_{n+k} = x) > \frac{1}{2} + \frac{1}{2} = 1$ which is absurd. Thus $P_\nu(T \cap \theta^k T) = 0$ implies the disjointness of B_n and B_{n+k} for all n and k and (i) and (ii) are proved.

To prove (iii) it will suffice to show that

$$\bigcup_{n=-\infty}^{\infty} \theta^n T = \lim_{n\to\infty}\{X_n \in \bigcup_{k=0}^{\infty} B_k\} \qquad P_\nu \quad \text{a.s.}$$

It is clear that $I = \bigcup_{n=-\infty}^{\infty} \theta^n T$ belongs to $\mathcal{J}$. In addition, we can prove that

I is an atomic set of $\mathcal{J}$. Indeed, if we admit the contrary, since T is an atomic set of $\mathcal{T}$ and I can be expressed as a union of mutually disjoint sets, we would get that a subset of I (say), I' in $\mathcal{J}$ with $0 < P_\nu(I') < P_\nu(I)$ will entail the existence of a subset of T (say), T' in $\mathcal{T}$ with $0 < P_\nu(T') < P_\nu(T)$ which is absurd.

Consider now the set $C = \{x : P(I|X_o = x) > \frac{1}{2}\}$. Then $C \supseteq \bigcup_{k=o}^{d} B_k$. This entails $I \supseteq \{X_n \in \bigcup_{k=o}^{d} B_k \text{ i.o.}\}$. Furthermore

$$\begin{aligned} P_\nu(I) &= P_\nu\Big(\bigcup_{k=-\infty}^{\infty} \theta^k\{X_n \in B_n \quad \text{all large } n\}\Big) \\ &= P_\nu(X_n \in B_{n+k} \quad \text{all large } n \text{ , some } k) \\ &\leq P_\nu(X_n \in \bigcup_{k=o}^{\infty} B_k \quad \text{all large } n) \end{aligned}$$

and the proof is complete.

Theorem 5 was established by Abrahamse [1] in the case of a countable chain assuming a restriction on the chain which is satisfied by the properly homogeneous chains considered above.

REFERENCES

1. Abrahamse, A.F. The Tail σ-Field of a Markov Chain. Ann. Math. Statist. 40 (1969) 127-136.

2. Abrahamse, A.F. The Martin Potential Kernel for Improperly Essential Chains. Z. Wahrscheinlichkeits. 18 (1971) 218-226.

3. Bartfai, P. and P. Revesz. On a Zero-One Law. Z. Wahrscheinlichkeits. 7 (1967) 43-47.

4. Blackwell, D. On Transient Markov Processes with a Countable Number of States and Stationary Transition Probabilities. Ann. Math. Statist. 26 (1955) 654-658.

5. Blackwell, D. and D. Kendall. The Martin Boundary for Polya's Urn Scheme and an Application to Stochastic Population Growth. J. Appl. Probab. 1 (1964) 284-296.

6. Blackwell, D. and D. Freedman. The Tail σ-Field of a Markov Chain and a Theorem of Orey. Ann. Math. Statist. 35 (1964) 1291-1295.

7. Chacon, R.V. and D.S. Ornstein. A General Ergodic Theorem. Illinois J. Math. 4 (1960) 153-160.

8. Chung, K.L. Markov Chains with Stationary Transition Probabilities, 2nd Ed., Springer, New York, 1967.

9. Cohn, H. On the Tail Events of a Markov Chain. Z. Wahrscheinlichkeits. 29 (1974) 65-72.

10. Cohn. H. On the Tail σ-Field of the Countable Markov Chains. Rev. Roumaine Math. Pures Appliq. 21 (1976) 667-675.

11. Cohn, H. On the Invariant Events of Markov Chains. Z. Wahrscheinlichkeits. 48 (1979) 81-96.

12. Cohn, H. Countable Nonhomogeneous Markov Chains : Asymptotic Behaviour. Adv. Appl. Probab. 9 (1977) 542-552.

13. Derriennic, Y. Sur la Frontiere de Martin de Processus de Markov a temps Discret. Ann. Inst. H. Poincare, Sec. B, XI, 3 (1973) 233-258.

14. Derriennic, Y. Lois ≪Zero ou Deux≫ pour les Processus de Markov. Applications aux Marches Aleatoire. Ann. Inst. H. Poincare, Sec. B, XII, 2 (1976) 111-129.

15. Doob, J.L. Stochastic Processes. Wiley, New York, 1953.

16. Doob, J.L. Discrete Potential Theory and Boundaries. J. Math. & Mech. 8 (1959) 433-458.

17. Doob, J.L. A Markov Chain Theorem. Probability and Statistics. (H. Cramer memorial volume) ed. U. Grenander. Almquist and Wiksel, Stockholm & New York (1959) 50-57.

18. Duflo, M. and D. Revusz. Proprietes Asymptotiques des Probabilites de Transition des Processus de Markov Recurrents. Ann. Inst. H. Poincare 5 (1969) 233-244.

19. Foguel, S.R. The Ergodic Theory of Markov Processes. Van Nostrand, New York, 1969.

20. Freedman, D. Markov Chains. Holden Day, San Francisco, 1971.

21. Griffeath, D. Partial Coupling and Loss of Memory for Markov Chains. Ann. Probab. 4 (1976) 850-858.

22. Griffeath, D. Coupling Methods for Markov Processes. Adv. in Math. Suppl. Studies, Stud. in Probab. & Ergodic Theory 2 (1978) 1-43.

23. Hewitt, E. and L.J. Savage. Symmetric Measures on Cartesian Products. Trans. Amer. Math. Soc. 80 (1955) 470-501.

24. Hunt, G.A. Markov Chains and Martin Boundaries. Illinois J. Math. 4 (1960) 313-340.

25. Iosifescu, M. On Two Recent Papers on Ergodicity in Non-homogeneous Markov Chains. Ann. Math. Statist. 43 (1972) 1732-1736.

26. Iosifescu, M. On Finite Tail σ-Algebras. Z. Wahrscheinlichkeits. 24 (1972) 159-166.

27. Jamieson, B. and S. Orey. Markov Chains Recurrent in the Sense of Harris. Z. Wahrscheinlichkeits. 8 (1967) 41-48.

28. Kemeny, J.G., J.L. Snell and A.W. Knapp. Denumerable Markov Chains. Springer, New York, 1976.

29. Kolmogorov, A.N. Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer, Berlin, 1933.

30. Lamperty, J. and J.L. Snell. Martin Boundaries for Certain Markov Chains. J. Math. Soc. Japan 15 (1963) 113-128.

31. Loeve, M. Probability Theory. 3rd Ed. Van Nostrand, Princeton, 1963.

32. McDonald, D. On Semi-Markov and Semi-Regenerative Processes I. Z. Wahrscheinlichkeits. 42 (1978) 261-277.

33. Neveu, J. Bases Mathematiques du Calcul de Probabilites. Masson, Paris, 1964.

34. Neveu, J. Chaines de Markov et Theorie du Potential. Annal. Fac. Sci. Clermont-Ferrand 24 (1964) 37-89.

35. Orey, S. An Ergodic Theorem for Markov Chains. Z. Wahrscheinlichkeits. 1 (1962) 174-186.

36. Orey, S. Limit Theorems for Markov Chains Transition Probabilities. Van Nostrand, Reinhold, London, 1971.

37. Ornstein, D. and D. Sucheston. An Operator Theorem on L_1 Convergence to Zero with Applications to Markov Kernels. Ann. Math. Statist. 41 (1970) 1631-1639.

38. Renyi, A. Probability Theory. North-Holland Publ. Co., Amsterdam, 1970.

39. Revuz, D. Markov Chains. North-Holland Publ. Co., Amsterdam, 1975.