Z. Wahrscheinlichkeitstheorie verw. Gebiete
48, 81–96 (1979)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

On the Invariant Events of a Markov Chain

Harry Cohn
University of Melbourne, Department of Statistics, Parkville, Victoria 3052, Australia

1. Introduction and Summary

Let S be a countable set of integers, $N=\{0,1,\ldots\}$ and $\Omega=S^N$. Define the variables $\{X_n(\omega): n\geqq 0\}$ on Ω by $X_n(\omega)=\omega_n$, where $\omega=(\omega_0,\omega_1,\ldots,\omega_n,\ldots)$. Let $\mathscr{F}$ be the σ-field generated by the variables $\{X_n(\omega): n\geqq 0\}$. An initial probability vector $\mu^{(0)}=(\mu_i^{(0)}; i\in S)$ and a (1-step) transition probability matrix $P=(P(i,j); i,j\in S)$ determine a probability measure $\mathscr{P}$ on $\mathscr{F}$ and a temporally homogeneous Markov chain $\{X_n(\omega): n\geqq 0\}$ on $(\Omega,\mathscr{F},\mathscr{P})$ such that $\mu_i^{(0)}=\mathscr{P}(X_0=i)$ and $P(i,j)=\mathscr{P}(X_{n+1}=j|X_n=i)$ provided that $\mathscr{P}(X_n=i)>0$. Denote by $\{\mu^{(n)}: n\geqq 0\}$ the sequence of the absolute probability vectors, where $\mu^{(n)}=(\mu_i^{(n)}; i\in S)$, $\mu_i^{(n)}=\mathscr{P}(X_n=i)$ and let $P^n=(P^{(n)}(i,j); i,j\in S)$ be the n-step transition probability matrix. Throughout the paper our results will refer to a Markov chain for which the initial probability vector is strictly positive (i.e. $\mu_i^{(0)}>0$ for all $i\in S$) and $\mathscr{P}$ will correspond to such an initial probability vector. In the proofs we shall sometimes consider Markov chains assuming the same transition probability matrix P but a different initial probability vector, (say) λ, and in this case $\mathscr{P}_\lambda$ will stand for the corresponding probability of the chain. We shall abbreviate $\mathscr{P}_i$ for $\mathscr{P}_{\delta(i)}$ where δ stands for the Dirac measure.

Let $\mathscr{F}_n$ be the σ-field generated by X_n, and $\mathscr{F}_n^\infty$ the σ-field generated by X_n, $X_{n+1},\ldots$, $\mathscr{T}=\bigcap_{n=0}^{\infty}\mathscr{F}_n$ will be said to be the *tail σ-field* of the chain. A set A in a σ-field $\mathscr{G}$ is called *atomic* with respect to $\mathscr{G}$ if $\mathscr{P}(A)>0$ and A does not contain two disjoint subsets of positive probability belonging to $\mathscr{G}$. A set A in $\mathscr{G}$ is called *completely nonatomic* with respect to $\mathscr{G}$ if $\mathscr{P}(A)>0$ and A does not contain any atomic subset belonging to $\mathscr{G}$. It is well known that, in general, Ω may be represented as $\Omega=\bigcup_{n=0}^{\infty}A_n$, where A_0 is completely non-atomic and $A_1,A_2,\ldots$ are atomic sets with respect to $\mathscr{G}$. If $A_1=\Omega$, $\mathscr{G}$ will be said to be *trivial*.

If λ and ν are two finite measures on a measurable space $(X,\mathscr{X})$ we denote by $\|\lambda-\nu\|$ the total variation of $\lambda-\nu$ i.e. $\|\lambda-\nu\|=(\lambda-\nu)^+(X)+(\lambda-\nu)^-(X)$, where $(\lambda-\nu)^+$ and $(\lambda-\nu)^-$ are the positive and negative parts of $\lambda-\nu$ in its Jordan decomposition. It is easy to see that if $X=S$ and $\mathscr{X}$ is the class of all subsets of S, $\|\lambda-\nu\|=\sum_{i\in S}|\lambda(i)-\nu(i)|$. Further A^c will stand for the complementary set of A, $A_1\Delta A_2$

0044-3719/79/0048/0081/$03.20

for the symmetric difference of A_1 and A_2, Z for the set of the integers and R for the set of the real numbers. A shift function $T: \Omega \to \Omega$ is defined by setting $T(\omega_0, \omega_1, \ldots) = (\omega_1, \omega_2, \ldots)$. We shall write $TA = \{T\omega: \omega \in A\}$, $T^{-1}A = \{\omega: T\omega \in A\}$ and $T^0 A = A$. A set $A \in \mathscr{F}$ is said to be *invariant* if $T^{-1}A = A$. The class of all invariant sets, denoted by $\mathscr{I}$ is a σ-field, called the invariant σ-field. It is easy to see that both T and T^{-1} are countably additive maps from $\mathscr{F}$ into $\mathscr{F}$. Besides, T^{-1} preserves the disjointness of sets and commutes with complementation and countable intersections. These properties of T^{-1}, not possessed by T, are probably accountable for the use of T^{-1} in the definition and the investigations of the invariant sets from the very beginning of the ergodic theory.

In a paper concerning the structure of $\mathscr{T}$, Abrahamse [1] has shown that if T is restricted to the sets of $\mathscr{T}$, then it proves tractable and useful. He has first proved that T maps $\mathscr{T}$ one-to-one onto itself and $\mathscr{I} = \{A \in \mathscr{T}: TA = A\}$ (Theorem 1, [1]). This result implies that an invariant set can also be defined as a set with the property $TA = A$. To the further "rehabilitation" of T we remark that making use of the above mentioned result of [1] we can prove that T restricted to $\mathscr{T}$ has also other desirable properties, which will be needed in what follows, expressed by the following

Proposition 1. *Suppose that* $A, A_1, A_2, \ldots$ *belong to* $\mathscr{T}$. *Then*

(i) $TA^c = (TA)^c$,

(ii) $T \bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} TA_n$,

(iii) $T^{m+n}A = T^m T^n A$ *for* $m, n \in Z$.

We remark that these results hold in general, the Markov property being not used in their derivation. We shall say that A is a *null set* if $\mathscr{P}(A) = 0$. If $\mathscr{P}(T^n A) = 0$ for all $n \in Z$, A will be said to be a *small set*. It is easy to see that if A is a null set then $\mathscr{P}(T^{-n}A) = 0$ for all $n \in N$. Indeed, $\mathscr{P}(T^{-n}A \mid X_n = i) = \mathscr{P}(A \mid X_0 = i) = 0$ for all $i \in \{j: \mu_j^{(n)} > 0\}$. However $\mathscr{P}(TA)$ is not necessarily null for any null set A and therefore not all the null sets are small sets. In Sect. 4 we identify a class of small sets that will prove useful in some applications.

Suppose that $\mathscr{G}$ and $\mathscr{H}$ are two σ-fields such that $\mathscr{G} \subset \mathscr{H}$. We shall say that $\mathscr{G} = \mathscr{H}$ a.s. if the sets of $\mathscr{H}$ are the sets of $\mathscr{G}$ modulo small sets and $\mathscr{G} \subset \mathscr{H}$ a.s. otherwise.

Let $A = (A_0, A_1, \ldots)$ be a sequence of subsets of S. We shall say that $\lim_{n\to\infty} \{X_n \in A_n\} = A$ a.s. if

$$\mathscr{P}(A \Delta \liminf_{n\to\infty} \{X_n \in A_n\}) = \mathscr{P}(A \Delta \limsup_{n\to\infty} \{X_n \in A_n\}) = 0.$$

A subset C of S will be said to be *almost closed* if $\lim_{n\to\infty} \{X_n \in C\}$ exists a.s. and assumes positive probability. C will be said to be a *transient set* if $\limsup_{n\to\infty} \{X_n \in C\}$ is a null set. Denote by $\mathfrak{C}$ the class of all almost closed and transient sets, by $\mathfrak{B}$ the class of all transient sets and by $\mathscr{N}$ the class of all null sets in $\mathscr{I}$. It is easy to see that $\mathfrak{C}$ is a boolean algebra and $\mathfrak{B}$ is an ideal in $\mathfrak{C}$. The following basic result due to Blackwell [3] (see also Chung [5], Theorem 1, Sect. 17) exhibits the relationship between the elements of $\mathfrak{C}$ and $\mathscr{I}$.

Theorem A. *To each invariant set A there corresponds a transient or almost closed set C such that* $A=\lim_{n\to\infty}\{X_n\in C\}$ a.s., *according as A is a null set or not. This correspondence is an isomorphism from $\mathscr{I}/\mathscr{N}$ onto $\mathfrak{C}/\mathfrak{B}$.*

Abrahamse, in the already mentioned paper [1] has shown that an isomorphism of the type referred to in Theorem A can also be established between some sequences of sets $A=(A_0, A_1, \ldots)$ such that $\lim_{n\to\infty}\{X_n\in A_n\}$ exists a.s. and the sets of $\mathscr{T}$, with the difference that the rôle of the null sets is played here by the small sets. In analogy to the invariant sets case discussed above, we shall say that A is a *totally transient sequence* if $\limsup_{n\to\infty}\{X_n\in A_n\}$ is a small set and A will be said to be a *tail sequence* if it is not a totally transient sequence and if $\limsup_{n\to\infty}\{X_n\in A_n\} - \liminf_{n\to\infty}\{X_n\in A_n\}$ is a small set. For $A=(A_0, A_1, \ldots)$ and $B=(B_0, B_1, \ldots)$ we shall define $A^c=(A_0^c, A_1^c, \ldots)$, $A\cup B=(A_0\cup B_0, A_1\cup B_1, \ldots)$ and $A\cap B =(A_0\cap B_0, A_1\cap B_1, \ldots)$, $TA=(A_1, A_2, \ldots)$ and $T^{-1}A=(S, A_0, \ldots)$. If we denote by $\mathfrak{S}$ the class of all totally transient and tail sequences and by $\mathfrak{D}$ the class of all totally transient sets, then we can easily check that $\mathfrak{S}$ is a boolean algebra and $\mathfrak{D}$ is an ideal in $\mathfrak{S}$. Denote by $\mathscr{M}$ the class of all small sets in $\mathscr{T}$. We shall say that $\Lambda\,\Delta\lim_{n\to\infty}\{X_n\in A_n\}$ a.s. is a small set if $\Lambda\,\Delta\limsup_{n\to\infty}\{X_n\in A_n\}$ and $\Lambda\,\Delta\liminf_{n\to\infty}\{X_n\in A_n\}$ are small sets.

The following result is due to Abrahamse (Theorem 5, [1]).

Theorem B. *To each set $\Lambda\in\mathscr{T}$ there corresponds a totally transient or a tail sequence $A=(A_0, A_1, \ldots)$ such that $\Lambda\,\Delta\lim_{n\to\infty}\{X_n\in A_n\}$ a.s. is a small set, according as Λ is a small set or not. This correspondence is an isomorphism from $\mathscr{T}/\mathscr{M}$ onto $\mathfrak{S}/\mathfrak{D}$, and commutes with T.*

The first criterion on the structure of the invariant σ-field is due to Blackwell, who in the already mentioned paper [3] showed that a necessary and sufficient condition for the triviality of $\mathscr{I}$ is that every bounded solution ϕ of the equation

$$\phi(i)=\sum_{j\in S}P(i,j)\,\phi(j) \tag{1}$$

be constant. Breiman [4] gave a characterization for some kind of atomic sets Λ of $\mathscr{I}$ in terms of the bounded solutions of the inequation

$$\phi(i)\leqq\sum_{j\in C}P(i,j)\,\phi(j)$$

where C is an almost closed set corresponding to an invariant set Λ.

Recently, Derriennic [9] proved for an arbitrary state space S, that $\mathscr{I}$ is trivial under any initial probability if and only if

$$\lim_{n\to\infty}\frac{1}{n}\left\|\sum_{j=1}^{n}P^{(j)}(x,\cdot)-P^{(j)}(y,\cdot)\right\|=0, \qquad x,y\in S.$$

Derriennic's proof leans heavily on the properties of the contractions on a Banach space.

In the present paper we are concerned both with the structure of the invariant σ-field and with its relation to the tail σ-field. In Sect. 2 we give some necessary and sufficient conditions for $\mathscr{I}=\mathscr{T}$ a.s. Our conditions are related to the quantities

$$\alpha(x)=\lim_{n\to\infty}\|P^{(n)}(x,\cdot)-P^{(n-1)}(x,\cdot)\|, \qquad x\in S \tag{2}$$

which were first considered by Ornstein and Sucheston [15], who proved that under certain assumptions $\alpha(x)$ is either 0 or 2, a property that became known as "the 0–2 law". Subsequently, Derriennic [9] proved a very general 0–2 law and showed that if $\mathscr{I}=\mathscr{T}$ a.s. under any initial probability $\mu^{(0)}$, then $\alpha(x)=0$ for all x and $\sup_{x\in S}\alpha(x)=2$ otherwise. Ornstein and Sucheston's proof is based on L_1-operators theory, whereas Derriennic used a combined martingale and operator-theoretic approach.

Our approach is based on the martingale convergence theorem and does not use the notion of operator. Besides, one of the equivalent conditions for $\mathscr{I}=\mathscr{T}$ a.s. is expressed by means of an a.s. convergent sequence, which proves adequate in some applications involving recurrence conditions.

In Sect. 3 we give a result characterizing both the atomic and the completely non-atomic sets of $\mathscr{I}$, which parallels the results given for the tail σ-field in [6] and [12].

In the final Section we study the invariant sets attached to a normed sequence of random variables which converges almost surely and explore their relation to the σ-field generated by the limiting random variable. As an application, classes of invariant events are identified for some supercritical branching processes.

2. The Case $\mathscr{I}=\mathscr{T}$ a.s.

For any state i such that $\mu_i^{(1)}>0$ we shall define the random variables

$$z_n(\omega)=\begin{cases}\dfrac{P^{(n)}(i,\omega_n)}{P^{(n-1)}(i,\omega_n)} & \text{if } P^{(n-1)}(i,\omega_n)>0\\ 1 & \text{if } P^{(n-1)}(i,\omega_n)=0.\end{cases}$$

The random variables $\{z_n(\omega)\}$, defined in a slightly different way, were considered in [7] where they were used to give a unified martingale approach to some results of the tail σ-field theory. We found out recently that similar random variables were considered before, in connection with Martin boundary theory, where their convergence was derived by using the space-time harmonic function theory (see e.g. [10]).

We shall further show that the random variables $\{z_n(\omega)\}$ can be used to derive a criterion for $\mathscr{I}=\mathscr{T}$ a.s.

Theorem 1. *The following three statements are equivalent:*

(i) $\mathscr{I}=\mathscr{T}$ a.s.

(ii) $\lim_{n\to\infty} z_n(\omega)=1$ a.s. *for all i such that* $\mu_i^{(1)}>0$,

(iii) $\lim_{n\to\infty}\|P^{(n)}(i,\cdot)-P^{(n-1)}(i,\cdot)\|=0$ *for all* $i\in S$.

Proof. Suppose that (i) holds and define the random variables

$$\alpha_n(X_n)=\frac{\mathscr{P}(X_1=i|X_n)}{\mu_i^{(1)}}-\frac{\mathscr{P}(X_0=i|X_n)}{\mu_i^{(0)}}. \tag{3}$$

We shall prove that $\lim_{n\to\infty}\alpha_n(X_n)=0$ a.s. if and only if $\lim_{n\to\infty}z_n(\omega)=1$ a.s.

By a well known property of Markov chains we get

$$\alpha_n(X_n)=\frac{\mathscr{P}(X_1=i|\mathscr{F}_n^\infty)}{\mu_i^{(1)}}-\frac{\mathscr{P}(X_0=i|\mathscr{F}_n^\infty)}{\mu_i^{(0)}}.$$

Now the martingale convergence theorem (see e.g. [14] p. 409) yields

$$\lim_{n\to\infty}\alpha_n(X_n)=\frac{\mathscr{P}(X_1=i|\mathscr{T})}{\mu_i^{(1)}}-\frac{\mathscr{P}(X_0=i|\mathscr{T})}{\mu_i^{(0)}} \quad \text{a.s.} \tag{4}$$

By elementary calculations we can deduce that unless $\mathscr{P}(X_0=i|X_n)=0$, $z_n(\omega)$ is the ratio of the quantities $\mathscr{P}(X_1=i|X_n)/\mu_i^{(1)}$ and $\mathscr{P}(X_0=i|X_n)/\mu_i^{(0)}$ which appear on the right hand side of (3) and in the case $\mathscr{P}(X_0=i|X_n)=0$, $z_n(\omega)$ is defined as being equal to 1.

As in [7] we get

$$\lim_{n\to\infty}z_n(\omega)=\frac{\mathscr{P}(X_1=i|\mathscr{T})}{\mathscr{P}(X_0=i|\mathscr{T})}\frac{\mu_i^{(0)}}{\mu_i^{(1)}} \tag{5}$$

for almost all $\omega\in\{\omega:\mathscr{P}(X_0=i|\mathscr{T})>0\}$.

Unlike the $\{z_n(\omega)\}$, the random variables $\{\alpha_n(X_n)\}$ are defined without any modifications and their limit (4) is established without the restriction: "for almost all $\omega\in\{\omega:\mathscr{P}(X_0=i|\mathscr{T})>0\}$" imposed for the validity of (5). Therefore, to complete the proof of the fact that $\lim_{n\to\infty}\alpha_n(X_n)=0$ a.s. if and only if $\lim_{n\to\infty}z_n(\omega)=1$ a.s. it will be sufficient to show that there exists a sequence $\{C_n: n\geqq 0\}$ such that $\lim_{n\to\infty}\{X_n\in C_n\}=\Lambda_0=\{\omega:\mathscr{P}(X_0=i|\mathscr{T})=0\}$ a.s. and that $P^{(n-1)}(i,j)=P^{(n)}(i,j)=0$ for $j\in C_{n-1}\cup C_n$, $n=1,2,\ldots$. Indeed, we know that Λ_0 differs from a set in $\mathscr{I}$ at most by a null set and according to Theorem A there exists a set C such that $\lim_{n\to\infty}\{X_n\in C\}=\Lambda_0$ a.s. It is easy to see that we can take $C_n=C-D_n$, where $D_n=\{j:\mu_j^{(n)}>0, \mathscr{P}(\Lambda_0|X_n=j)=0\}$, $n=0,1,\ldots$. Suppose now that for a certain k, $P^{(k)}(i,j)>0$ with $j\in C_k$. Then by the Chapman-Kolmogorov formula $\mathscr{P}(\Lambda_0|X_0=i)\geqq\mathscr{P}(\Lambda_0|X_k=j)P^{(k)}(i,j)>0$. But from the definition of Λ_0 we obtain $\mathscr{P}(\Lambda_0|X_0=i)=0$ which is a contradiction. Therefore $P^{(n)}(i,j)=0$ for $j\in C_n$, $n=1,2,\ldots$. Further $\mathscr{P}(\Lambda_0|X_1=i)=\mathscr{P}(\Lambda_0|X_0=i)=0$ and therefore $P^{(n-1)}(i,j)=0$ for $j\in C_{n-1}$, $n=1,2,\ldots$. Hence $\lim_{n\to\infty}\alpha_n(X_n)=0$ a.s. if and only if $\lim_{n\to\infty}z_n(\omega)=1$ a.s.

We show now that $\lim_{n\to\infty}\alpha_n(X_n)=0$ a.s. is implied by (i). Suppose the contrary, i.e. that there exists a set Λ_1 such that for $\omega\in\Lambda_1$, $\alpha(\omega)=\lim_{n\to\infty}\alpha_n(X_n)\neq 0$. We may take for definiteness $\alpha(\omega)>0$ for $\omega\in\Lambda_1$. If we integrate (4) over Λ_1, after elementary calculations we get

$$\mathscr{P}(\Lambda_1|X_1=i)>\mathscr{P}(\Lambda_1|X_0=i)$$

which contradicts the assumption (i) and the implication (i)→(ii) is proved.

Suppose now that (ii) holds. Then the dominated convergence theorem applied to the sequence $\{\alpha_n(X_n)\}$ yields

$$\lim_{n\to\infty}\sup_{B\subset S}|\mathscr{P}(X_n\in B|X_1=i)-\mathscr{P}(X_n\in B|X_0=i)|\leqq \lim_{n\to\infty}\int_S|\alpha_n(X_n)|\,d\mathscr{P}=0$$

which implies

$$\lim_{n\to\infty}\|P^{(n)}(i,\cdot)-P^{(n-1)}(i,\cdot)\|=0. \tag{6}$$

We recall that the states i for which (6) was derived are subjected to the restriction $\mu_i^{(1)}>0$, which was assumed in the definition of $\{\alpha_n(X_n)\}$. We show now that this restriction can be removed. Because $\mu_i^{(0)}$ was supposed positive for all $i\in S$ we need only consider the case $\mu_i^{(0)}>0$ and $\mu_i^{(1)}=0$. Denote $S'=\{j:\mu_j^{(1)}>0\}$, and let $a^+=\max(a,0)$. Then, for any $i\in S$

$$\begin{aligned}&\|P^{(n)}(i,\cdot)-P^{(n-1)}(i,\cdot)\|\\&\quad=\tfrac{1}{2}\sum_{k\in S}\Big[\sum_{j\in S}P(i,j)P^{(n-1)}(j,k)-\sum_{j\in S}P(i,j)P^{(n-2)}(j,k)\Big]^+.\end{aligned} \tag{7}$$

Further (6) and (7) imply

$$\begin{aligned}&\lim_{n\to\infty}\|P^{(n)}(i,\cdot)-P^{(n-1)}(i,\cdot)\|\\&\quad\leqq\lim_{n\to\infty}\sum_{j\in S'}P(i,j)\|P^{(n-1)}(j,\cdot)-P^{(n-2)}(j,\cdot)\|=0\end{aligned}$$

and the implication (ii)→(iii) is proved.

Suppose now that (iii) holds and assume that $\mathscr{I}\subset\mathscr{T}$ a.s. In this case there exists a set $\Lambda\in\mathscr{T}$ such that $\mathscr{P}(T\Lambda\Delta\Lambda)>0$ and we can suppose without loss of generality that Λ and $T\Lambda$ are disjoint. Indeed, if Λ and $T\Lambda$ are not disjoint then in view of Proposition 1 we can arrange to have such a situation by taking $\Lambda\cap(T\Lambda)^c$ instead of Λ.

By a well known procedure which goes back to Blackwell [3], we know that if we take $B_n=\{j:\mathscr{P}(\Lambda|X_n=j)>\delta\}$ with $\delta>\frac{1}{2}$ then $\{B_n\}\in\mathfrak{S}$ is the sequence corresponding to Λ in the isomorphism alluded to in Theorem B (see also [1]). Further, according to the same Theorem B, this isomorphism commutes with T and therefore $\lim_{n\to\infty}\{X_n\in B_{n+1}\}=T\Lambda$ a.s. Now, if $i\in B_m$ and n is sufficiently large

$$\mathscr{P}(X_{n+m}\in B_{n+m}|X_m=i)\geqq\delta \quad\text{and}\quad \mathscr{P}(X_{n+m+1}\in B_{n+m+1}|X_m=i)\geqq\delta$$

and taking into account that B_n and B_{n+1} are disjoint for all n, we get $\|P^{(n)}(i,\cdot)-P^{(n-1)}(i,\cdot)\|\geqq 2\delta-1>0$ which is a contradiction and the proof is complete.

Remark. As in the case of the tail σ-field considered in [7], a property like (ii) may prove useful when dealing with recurrence type conditions, whereas (iii) expressing a "global property" seems less adequate in handling such situations with probabilistic methods. A criterion of this type can be used to derive the triviality of the tail σ-field, by proving first, what is often simpler, the triviality of $\mathscr{I}$ and then $\mathscr{I} = \mathscr{T}$ a.s. We notice that in this way we can give an alternative proof of the triviality of the tail σ-field for a Markov chain assuming a trivial $\mathscr{I}$ and satisfying the property $P(\omega: X_{n+1}(\omega) = X_n(\omega) \text{ i.o.}) = 1$, a result which was first established by Küchler in [13], where he gave a full description of the tail σ-field structure of a birth and death process. It is possible to show that in this case $\lim_{n\to\infty} z_n(\omega) = 1$ a.s. in the same way as in the alternative proof of the Blackwell and Freedman 0–1 law given in [7].

3. The Structure of $\mathscr{I}$

We consider now the vector chain $\{Y_n: n \geqq 0\}$ with $Y_n = (X_n^{(1)}, X_n^{(2)})$, $n \geqq 0$, defined on the probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathscr{P}})$; $\{X_n^{(1)}: n \geqq 0\}$ and $\{X_n^{(2)}: n \geqq 0\}$ being two independent copies of $\{X_n: n \geqq 0\}$, $\tilde{\Omega} = \Omega \times \Omega$, $\tilde{\mathscr{F}} = \mathscr{F} \otimes \mathscr{F}$ and $\tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{P}$.

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

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

Denote by I_0 the completely nonatomic set and by $I_1, I_2, \ldots$ the atomic sets occurring in the representation of Ω corresponding to $\mathscr{I}$.

The assertions "for almost all $\tilde{\omega}$" or "a.s." in the statement of the following Theorem will be understood to hold with respect to $\tilde{\mathscr{P}}$.

Theorem 2. (i) *There exists the limits*

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

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

(ii) $\beta(\tilde{\omega}) = 2$ *for almost all* $\tilde{\omega} \in I_0 \times I_0 \cup \bigcup_{u \neq u'} I_u \times I_{u'}$,

$\beta(\tilde{\omega}) = 0$ *for almost all* $\tilde{\omega} \in I_u \times I_u$, $u = 1, 2, \ldots$.

Proof. It is easy to see that $\frac{1}{n}\left\| \sum_{l=1}^{n} (P^{(l)}(i, \cdot) - P^{(l)}(j, \cdot)) \right\|$ converges (see e.g. Derriennic [9] p. 115) since if we denote $f(n) = \left\| \sum_{l=1}^{n} (P^{(l)}(i, \cdot) - P^{(l)}(j, \cdot)) \right\|$, then $f(n)$ can be shown to be a subadditive function i.e. $f(m+n) \leqq f(m) + f(n)$ for all $m, n \in N$ and therefore $\lim_{n\to\infty} f(n)/n = \inf_{n \geqq 1} f(n)/n$. Hence $\lim_{n\to\infty} \beta_{m,n}(\tilde{\omega}) = \beta_m(\tilde{\omega})$ exists for all $\tilde{\omega} \in \tilde{\Omega}$.

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

Because I_0 is completely nonatomic, for any $\varepsilon>0$ we can find $n(\varepsilon)$ disjoint sets $I(1), \ldots, I(n(\varepsilon))$, such that $I_0=I(1)\cup I(2)\cup\ldots\cup I(n(\varepsilon))$ and $0<\mathscr{P}(I(s))<\varepsilon/4$ for $1\leqq s\leqq n(\varepsilon)$. Let $C(s)=\{j: \mathscr{P}(I(s)|X_0=j)>1-\varepsilon/4\}$. As we have seen before $\lim_{n\to\infty}\{X_n\in C(s)\}=I(s)$ a.s. with respect to $\mathscr{P}$ for $s=1,2,\ldots,n(\varepsilon)$. It follows that $\lim_{l\to\infty} P^{(l)}(i, C(s))=\mathscr{P}(I(s)|X_0=i)$. Take now $i_1\in C(s)$ and $i_2\in C(s')$ with $s\neq s'$. Then

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\left\|\sum_{l=1}^{n}(P^{(l)}(i_1,\cdot)-P^{(l)}(i_2,\cdot)\right\| &\geqq \lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}(P^{(l)}(i_1,C(s))-P^{(l)}(i_2,C(s)))\\ &\quad+\lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}(P^{(l)}(i_2,C(s')-P^{(l)}(i_1,C(s')))\\ &\geqq 2-\varepsilon.\end{aligned}$$

It follows that $\liminf_{n\to\infty}\beta_m(\tilde{\omega})>2-\varepsilon$ for almost all $\tilde{\omega}\in I(s)\times I(s')$. However, we can further split each of the sets $I(s)$, $s=1,\ldots,n(\varepsilon)$ into disjoint subsets whose probabilities are smaller than $\varepsilon'/4$, for any preassigned ε' smaller than ε. Using the same reasoning as above we get, in particular, that $\liminf_{n\to\infty}\beta_m(\tilde{\omega})>2-\varepsilon'$ for almost all $\tilde{\omega}\in I(s)\times I(s')$ and because we can apply the same reasoning to any subsets of $I(s)$, $s=1,\ldots,n(\varepsilon)$ we deduce that $\beta(\tilde{\omega})=2$ for almost all $\tilde{\omega}\in I_0\times I_0$.

The proof of $\beta(\tilde{\omega})=2$ for almost all $\tilde{\omega}\in\bigcup_{u\neq u'} I_u\times I_{u'}$ is easier and will be left to the reader.

We shall prove now that $\beta(\tilde{\omega})=0$ for almost all $\tilde{\omega}\in I_u\times I_u$, $u=1,2,\ldots$ According to Theorem 1 of [9], for any i_1, $i_2\in S$

$$\lim_{n\to\infty}\frac{1}{n}\left\|\sum_{l=1}^{n}(P^{(l)}(i_1,\cdot)-P^{(l)}(i_2,\cdot))\right\|\leqq 2\sup_{A\in\mathscr{I}(i_1,i_2)}(\mathscr{P}_{i_1}(A)-\mathscr{P}_{i_2}(A)) \tag{8}$$

where $\mathscr{P}_i(A)=\mathscr{P}_\nu(A|X_0=i)$ for any initial probability ν such that $\nu_i^{(0)}>0$ and $\mathscr{I}(i_1,i_2)$ is the invariant σ-field of the Markov chain assuming the initial probability $\lambda=\frac{1}{2}(\delta(i_1)+\delta(i_2))$.

We show now that if I_u is an atomic set of $\mathscr{I}$ and if at least one of the inequalities $\mathscr{P}_{i_1}(I_u)>0$ and $\mathscr{P}_{i_2}(I_u)>0$ holds, then I_u is also atomic with respect to $\mathscr{I}(i,j)$. Indeed, suppose the contrary; then according to Theorem A there exist two disjoint almost closed sets C'_u and C''_u such that $\lim_{n\to\infty}\{X_n\in C'_u\}=A_1$ a.s. and $\lim_{n\to\infty}\{X_n\in C''_u\}=A_2$ a.s. with respect to $\mathscr{P}_\lambda$ and $\mathscr{P}_\lambda(A_1)\mathscr{P}_\lambda(A_2)>0$. Further, because $\liminf_{n\to\infty}\{X_n\in C'_u\}$ and $\liminf_{n\to\infty}\{X_n\in C''_u\}$ are both invariant and disjoint sets, then we must have either $\mathscr{P}(\liminf_{n\to\infty}\{X_n\in C'_u\})=0$ or $\mathscr{P}(\liminf_{n\to\infty}\{X_n\in C''_u\})=0$. But μ_0 was supposed positive for all $i\in S$ and therefore this entails either $\mathscr{P}_i(\liminf_{n\to\infty}\{X_n\in C'_u\})=0$ for all $i\in S$ or $\mathscr{P}_i(\liminf_{n\to\infty}\{X_n\in C''_u\})=0$ for all $i\in S$ and the inequality $\mathscr{P}_\lambda(A_1)\mathscr{P}_\lambda(A_2)>0$ is contradicted. Therefore I_u is an atomic set with respect to $\mathscr{I}(i,j)$.

Notice now that if we denote $S_n=\{i: \mu_i^{(n)}>0\}$, $n=0,1,\dots$ then $S_0 \supseteq S_1 \supseteq \dots \supseteq S_n \supseteq \dots$. Indeed, we can prove this by induction if we take into account that $\mu_i^{(0)}$ was supposed positive for all $i \in S$. Take now $i_1, i_2 \in S_m$. Then for any I_u, $u \geqq 1$ we get $\mathscr{P}(I_u|X_0=i_1)=\mathscr{P}(I_u|X_1=i_1)=\dots=\mathscr{P}(I_u|X_m=i_1)$ and similarly $\mathscr{P}(I_u|X_m=i_2)=\mathscr{P}(I_u|X_0=i_2)$. Further (8) yields

$$\lim_{n\to\infty} \frac{1}{n}\left\| \sum_{l=1}^{n} (P^{(l)}(X_m^{(1)}=i_1,\cdot)-P^{(l)}(X_m^{(2)}=i_2,\cdot))\right\|$$
$$\leqq 2\,|\mathscr{P}(I_u|X_m=i_1)-\mathscr{P}(I_u|X_m=i_2)|+2 \sup_{\substack{A\in\mathscr{I}(i,j)\\ A\subset I_u^c}} (\mathscr{P}_{i_1}(A)-\mathscr{P}_{i_2}(A)). \tag{9}$$

Choose now $C_u(\varepsilon)=\{i: \mathscr{P}_i(I_u)>1-\varepsilon/4\}$ and take i_1, $i_2 \in C_u$. Then (9) implies

$$\lim_{n\to\infty} \frac{1}{n}\left\| \sum_{l=1}^{n} (P^{(l)}(X_m^{(1)}=i_1,\cdot)-P^{(l)}(X_m^{(2)}=i_2,\cdot))\right\| \leqq 2\frac{\varepsilon}{4}+2\frac{\varepsilon}{4}=\varepsilon.$$

But $\lim_{m\to\infty} \{X_m \in C_u(\varepsilon)\}=I_u$ a.s. with respect to $\mathscr{P}$ for any $\varepsilon>0$ and the Theorem follows.

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

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

To study $\{\gamma_{m,n}(\omega)\}$ we shall need the following

Lemma 1. *For any m, n and i*

$$\left\| \sum_{l=1}^{n} (\mu^{(l+m)}(\cdot)-P^{(l)}(i,\cdot))\right\| \geqq \left\| \sum_{l=2}^{n+1} (\mu^{(l+m)}(\cdot)-P^{(l)}(i,\cdot))\right\|.$$

Proof. We recall that if ν is a signed measure with $\nu(S)=0$ then, $\|\nu\|=2\sup_{A\subset S}\nu(A)$. Further

$$\left\| \sum_{l=2}^{n+1} (\mu^{(l+m)}(\cdot)-P^{(l)}(i,\cdot))\right\|$$
$$\leqq 2\sup_{B\subset S^2} \sum_{l=1}^{n} [\mathscr{P}((X_{l+m}, X_{l+m+1})\in B)-\mathscr{P}((X_{l+m}, X_{l+m+1})\in B|X_m=i)]$$
$$=2\sup_{B\subset S^2} \sum_{l=1}^{n} \sum_{(j_1,j_2)\in B} P(j_1,j_2)(\mu_{j_1}^{(l+m)}-P^{(l)}(i,j_1))$$
$$\leqq \left\| \sum_{l=1}^{n} (\mu^{(l+m)}(\cdot)-P^{(l)}(i,\cdot))\right\|.$$

Theorem 3. (i) *There exist the limits*

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

$$\lim_{m\to\infty} \gamma_m(\omega)=\gamma(\omega) \quad \text{a.s.}$$

(ii) *$\gamma(\omega)=0$ for almost all $\omega\in I_0$,*
$\gamma(\omega)=P(I_u)$ for almost all $\omega\in I_u$ $u=1,2,\ldots$.

Proof. First notice that by the triangle inequality and by the above Lemma 1

$$\left\|\sum_{l=1}^{n+p}(\mu^{(l+m)}(\cdot)-P^{(l)}(i,\cdot))\right\| \leqq \left\|\sum_{l=1}^{n}(\mu^{(l+m)}(\cdot)-P^{(l)}(i,\cdot))\right\|+\left\|\sum_{l=1}^{p}(\mu^{(l+m)}(\cdot)-P^{(l)}(i,\cdot))\right\|.$$

Therefore, as in the case of the expression considered in Theorem 2

$$f(n)=\left\|\sum_{l=1}^{n}(\mu^{(l+m)}(\cdot)-P^{(l)}(i,\cdot))\right\|$$

is a subadditive function and hence $\lim_{n\to\infty}\gamma_{m,n}(\omega)=\gamma_m(\omega)$ exists for all m.

The existence of $\gamma(\omega)$ will be obtained in the course of the proof of (ii).

To prove that $\gamma(\omega)=0$ for almost all $\omega\in I_0$ we can proceed in the same way as in the proof of Theorem 2 and we leave it to the reader to work out the details.

To prove that $\gamma(\omega)=P(I_u)$ for almost all $\omega\in I_u$, $u\geqq 1$ we notice first that

$$\gamma_{m,n}(\omega)=1-\frac{1}{n}\sum_{j\in S}\left[\sum_{l=1}^{n}\sum_{k\in C(u)}\mu_k^{(m)}(P^{(l)}(k,j)-P^{(l)}(X_m,j)) + \sum_{l=1}^{n}\sum_{k\notin C(u)}\mu_k^{(m)}(P^{(l)}(k,j)-P^{(l)}(X_m,j))\right]^+. \tag{10}$$

We choose further $C(u)=\{j:\mathscr{P}(I_u|X_0=j)>1-\varepsilon/4\}$ for a certain preassigned $\varepsilon>0$.

We remark now that the first sum of the right hand side of (10) can be neglected since it is bounded by

$$\sum_{k\in C(u)}\mu_k^{(m)}\left\|\frac{1}{n}\sum_{l=1}^{n}(P^{(l)}(k,\cdot)-P^{(l)}(X_m,\cdot))\right\|$$

which by Theorem 2 tends to 0 for almost all $\omega\in I_u$ as n and m go to infinity.

We shall now deal with the second sum of the right hand side of (10). It is easy to see that

$$\lim_{m\to\infty}\lim_{n\to\infty}\frac{1}{n}\left\|\sum_{l=1}^{n}\sum_{k\notin C(u)}\mu_k^{(m)}(P^{(l)}(k,\cdot)-P^{(l)}(X_m,\cdot))\right\| - \lim_{m\to\infty}\lim_{n\to\infty}\frac{1}{n}\left\|\sum_{l=1}^{n}\sum_{k\in C'(u)}\mu_k^{(m)}(P^{(l)}(k,\cdot)-P^{(l)}(X_m,\cdot))\right\|=0 \tag{11}$$

where $C'(u)=\{j:\mathscr{P}(I_u^c|X_0=j)>1-\varepsilon/2\}$. Indeed, the left hand side of (11) is bounded by $\lim_{m\to\infty}\mu^{(m)}(C^c(u)\varDelta C'(u))$ which equals 0 since

$$\lim_{m\to\infty}\{X_m\in C^c(u)\}=\lim_{m\to\infty}\{X_m\in C'(u)\}=I_u^c \quad \text{a.s.}$$

Therefore, it remains to prove that

$$\lim_{m\to\infty}\lim_{n\to\infty}\left(1-\frac{1}{2n}\left\|\sum_{l=1}^{n}\sum_{k\in C'(u)}\mu_k^{(m)}(P^{(l)}(k,\cdot)-P^{(l)}(X_m,\cdot))\right\|\right)=\gamma(\omega) \tag{12}$$

for almost all $\omega\in I_u$.

Note further that

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{2n}\left\|\sum_{l=1}^{n}\sum_{k\in C'(u)}\mu_k^{(m)}(P^{(l)}(k,\cdot)-P^{(l)}(X_m,\cdot))\right\|\\
&\quad\geqq \lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n}\sum_{k\in C'(u)}\mu_k^{(m)}(P^{(l)}(k,C'(u))-P^{(l)}(X_m,C'(u)))\\
&\quad\geqq(1-\varepsilon)\mathscr{P}(X_m\in C'(u)).
\end{aligned} \tag{13}$$

provided that $X_m\in C(u)$.

Now the limit in (13) does not exceed $\mathscr{P}(X_m\in C'(u))$, and if we take into account that $\lim_{m\to\infty}\{X_m\in C'(u)\}=I_u^c$ a.s. implies $\lim_{m\to\infty}\mathscr{P}(X_m\in C'(u))=1-\mathscr{P}(I_u)$ we get that the limit in (12) exists and equals $\mathscr{P}(I_u)$ a.s. and the proof is complete.

4. Invariant Sets for Convergent Sequences

We recall that we have denoted $S_n=\{j:\mu_j^{(n)}>0\}$ for $n\in N$ and that $S_0\supseteq S_1\supseteq S_2\supseteq\ldots$. Since $\sum_{j\in S_n}P(i,j)=1$ for $i\in S_{n-1}$ we can easily deduce that $\lim_{n\to\infty}\{X_n\in S_{n+1}\}$ a.s. exists. But because $\{X_n\in[S_n-S_{n+1}]\}=\{X_n\in S_{n+1}\}^c$, it follows that $\lim_{n\to\infty}\{X_n\in[S_n-S_{n+1}]\}$ a.s. also exists. A Markov chain with the property $\mathscr{P}\{\limsup_{n\to\infty}X_n\in[S_n-S_{n+1}]\}=0$ will be said to be *properly homogeneous* and improperly homogeneous otherwise.

This definition is justified by the fact that if a chain is improperly homogeneous then the temporal homogeneity of its transition probabilities is of little use for the relevant sequence of sets $\{[S_n-S_{n+1}]:n=0,1,\ldots\}$ which consists of mutually disjoint sets and its states belonging to $[S_n-S_{n+1}]$ do not appear in the chain after time n.

In what follows we shall need the following

Lemma 2. *If $\{X_n:n\geqq 0\}$ is a properly homogeneous chain, then any null set in $\mathscr{T}$ is a small set.*

Proof. We have already seen in the introduction that for any null set A, $\mathscr{P}(T^{-n}A)=0$ if $n\in N$. Suppose that $A\in\mathscr{T}$, and $\mathscr{P}(A)>0$. Since $\mathscr{P}(TA\mid X_n=i)=\mathscr{P}(A\mid X_{n+1}=i)=0$ for $i\in S_{n+1}$ we get that

$$\mathscr{P}(TA)=\sum_{i\in[S_n-S_{n+1}]}\mathscr{P}(TA\mid X_n=i)\,\mu_i^{(n)}\leqq\mathscr{P}(X_n\in[S_n-S_{n+1}]).$$

But $\lim_{n\to\infty}\mathscr{P}(X_n\in[S_n-S_{n+1}])=0$ and therefore $\mathscr{P}(TA)=0$. The proof can now be completed by induction.

Let $Y_n = a_n(X_n + b_n)$, $n = 0, 1, \ldots$ where $\{a_n\}$ and $\{b_n\}$ are two sequences of constants and suppose that $\{Y_n : n \geqq 0\}$ converges almost surely to a random variable V. We shall say that V is a *proper random variable* if $\mathscr{P}(-\infty < V < \infty) = 1$. V will be said to be *nondegenerate* if it is not a.s. constant. Denote by $\mathscr{V}$ the σ-field generated by V and $\mathscr{W}$ the class of invariant sets belonging to $\mathscr{V}$. It is easy to see that $\mathscr{W}$ is a σ-field. In what follows we shall prove the following

Theorem 4. *Suppose that $\{X_n : n \geqq 0\}$ is a properly homogeneous chain and $\{Y_n : n \geqq 0\}$ converges a.s. to a proper and non-degenerate random variable V. Then $\lim_{n\to\infty} a_{n+1}/a_n = \alpha$ and $\lim_{n\to\infty} a_{n+1}(b_{n+1} - b_n) = \beta$ exist and are finite and one of the following cases occurs*

(i) $\alpha = 1$, $\beta = 0$ *and* $\mathscr{V} = \mathscr{W}$ *a.s.*

(ii) *At least one of the inequalities* $\alpha \neq 1$ *and* $\beta \neq 0$ *holds,* $\mathscr{W} \subset \mathscr{V}$ *a.s., and* $\mathscr{W}$ *is generated by the family of invariant sets*

$$\mathscr{C} = \Big\{ V \in \bigcup_{n=0}^{\infty} [a\alpha^n + \beta(\alpha^n - 1)(\alpha - 1)^{-1}, b\alpha^n + \beta(\alpha^n - 1)(\alpha - 1)^{-1}]^* \cup \bigcup_{n=1}^{\infty} [(a-\beta)\alpha^{-n} - \beta\alpha^{-1}(\alpha^{-n+1} - 1)(\alpha^{-1} - 1)^{-1}, (b-\beta)\alpha^{-n} - \beta\alpha^{-1}(\alpha^{-n+1} - 1)(\alpha^{-1} - 1)^{-1}]^* ; a, b \in R \Big\} \qquad \text{if } \alpha \neq 1$$

and

$$\mathscr{C} = \Big\{ V \in \bigcup_{n=-\infty}^{\infty} (a + n\beta, b + n\beta) \Big\} \qquad \text{if } \alpha = 1$$

where $[x_1, x_2]^*$ stands for the closed interval $[x_1, x_2]$ if $x_1 < x_2$ and for $[x_2, x_1]$ otherwise.

Proof. First, we notice that according to Theorem 4 of [8] the limits $\lim_{n\to\infty} a_{n+1}/a_n = \alpha$ and $\lim_{n\to\infty} a_{n+1}(b_{n+1} - b_n) = \beta$ exist, $-\infty < \alpha, \beta < \infty$ and $\alpha \neq 0$. We shall further show that if x_0 is a continuity point of F, where F is the distribution function of V, then $\alpha x_0 + \beta$ is also a continuity point of F. Suppose, for definiteness that $\alpha > 0$. Then for any $\varepsilon > 0$

$$\begin{aligned} &F(\alpha(x_0 + \varepsilon) + \beta) - F(\alpha(x_0 - \varepsilon) + \beta) \\ &\quad = \lim_{n\to\infty} \sum_{i \in S} \mathscr{P}(Y_n \in (\alpha(x_0 - \varepsilon) + \beta, \alpha(x_0 + \varepsilon) + \beta) \mid X_1 = i)\, \mu_i^{(1)} \\ &\quad = \lim_{n\to\infty} \sum_{i \in S} \mathscr{P}_i(a_n(X_{n-1} + b_n) \in (\alpha(x_0 - \varepsilon) + \beta, \alpha(x_0 + \varepsilon) + \beta)\, \mu_i^{(1)} \\ &\quad = \sum_{i \in S} \mathscr{P}_i(V \in (x_0 - \varepsilon, x_0 + \varepsilon))\, \mu_i^{(1)}. \end{aligned} \tag{14}$$

But

$$F(x_0 + \varepsilon) - F(x_0 - \varepsilon) = \sum_{i \in S} \mathscr{P}_i(V \in (x_0 - \varepsilon, x_0 + \varepsilon))\, \mu_i^{(0)} \tag{15}$$

and since $\mu_i^{(0)} > 0$ for all $i \in S$ the continuity of F at x_0 and (15) imply $\lim_{\varepsilon\to 0} \mathscr{P}_i(V \in (x_0 - \varepsilon, x_0 + \varepsilon)) = 0$ for all $i \in S$. Using this in (14) we get that F is continuous at $\alpha x_0 + \beta$. We can similarly show that the continuity of F at x_0 also implies its continuity at $\alpha^{-1}(x_0-\beta)$ (We start with $F(x_0+\varepsilon)-F(x_0-\varepsilon)$ and proceed as in (14), etc. ...). It follows, by induction, that if x_0 is a continuity point of F and if $\alpha \neq 1$, then

$$\{\alpha^n x_0+\beta(\alpha^n-1)(\alpha-1)^{-1}, (x_0-\beta)\,\alpha^{-n} \\ -\beta\alpha^{-1}(\alpha^{-n+1}-1)(\alpha^{-1}-1)^{-1};\, n=1,2,\dots\}$$

are also continuity points of F.

By Lemma 2 we get that whenever $\lim_{n\to\infty}\{X_n\in A_n\}=A$ a.s., $\{A_n\}\in\mathfrak{S}$ and $A\Delta \lim_{n\to\infty}\{X_n\in A_n\}$ a.s. is a small set and in such a case Theorem B implies that $TA = \lim_{n\to\infty}\{X_n\in A_{n+1}\}$ a.s.

Consider now the event $\{a\leqq V\leqq b\}$ where a and b are continuity points of F. Then we can easily check that

$$T\{a\leqq V\leqq b\}=\lim_{n\to\infty}\{a_n^{-1}\,a-b_n\leqq X_{n+1}\leqq a_n^{-1}\,b-b_n\} \\ =\{a\alpha+\beta\leqq V\leqq b\alpha+\beta\} \quad \text{a.s.}$$

Further if c is a jump point for F i.e. if $P(V=c)>0$ we can choose a decreasing sequence of numbers $\{c_n\}$ such that $\lim_{n\to\infty} c_n=0$ and $\{c-c_n, c+c_n\colon n=1,2,\dots\}$ are continuity points of F. Proposition 1(ii) implies that $T\{V=c\}=\bigcap_{n=1}^{\infty} T\{c-c_n\leqq V\leqq c+c_n\}=\{V=\alpha c+\beta\}$ a.s. An upshot of these considerations is $T\{a\leqq V\leqq b\}=\{a\alpha+\beta\leqq V\leqq b\alpha+\beta\}$ a.s. for any a, $b\in R$. We can similarly show that $T^{-1}\{a\leqq V\leqq b\}=\{(a-\beta)\alpha^{-1}\leqq V\leqq(b-\beta)\alpha^{-1}\}$ a.s. for a, $b\in R$.

Suppose now that $\lim_{n\to\infty} a_{n+1}/a_n=1$ and $\lim a_{n+1}(b_{n+1}-b_n)=0$. Then the above equality yields $T\{a\leqq V\leqq b\}=\{a\leqq V\leqq b\}$ a.s. and according to Theorem 6 of [1], $\{a\leqq V\leqq b\}$ is either an invariant set or differs from an invariant set by a small set. Because the family of sets $\{\{a\leqq V\leqq b\};a,b\in R\}$ generates $\mathscr{V}$, (i) follows.

Suppose now that at least one of the inequalities $\alpha\neq 1$, $\beta\neq 0$ holds. Then as we have seen before $T\{a\leqq V\leqq b\}=\{a\alpha+\beta\leqq V\leqq b\alpha+\beta\}$ a.s. If we choose a and b such that $\mathscr{P}(\{a\leqq V\leqq b\})>0$ and $a\alpha+\beta>b$ then $\{a\leqq V\leqq b\}$ and $T\{a\leqq V\leqq b\}$ are a.s. disjoint and as a consequence we get $\mathscr{V}\supset\mathscr{W}$ a.s.

We can inductively show that

$$T^n\{a\leqq V\leqq b\}=\{a\alpha^n+\beta(\alpha^n-1)(\alpha-1)^{-1}\leqq V \\ \leqq b\alpha^n+\beta(\alpha^n-1)(\alpha-1)^{-1}\} \quad \text{a.s.}$$

for $n=1,2,\dots$

and

$$T^{-n}\{a\leqq V\leqq b\}=\{(a-\beta)\,\alpha^{-n}-\beta\alpha^{-1}(\alpha^{-n+1}-1)(\alpha^{-1}-1)^{-1} \\ \leqq V\leqq(b-\beta)\,\alpha^{-n}-\beta\alpha^{-1}(\alpha^{-n+1}-1)(\alpha^{-1}-1)^{-1}\}$$

for $n=1,2,\dots$.

Using the countable additivity of T and Proposition 1(iii) we get that if $\Lambda\in\mathscr{V}$, $\psi(\Lambda)=\bigcup_{n=-\infty}^{\infty} T^n\Lambda\in\mathscr{W}$ and that $\psi(\Lambda)$ is the smallest invariant set containing Λ. It follows that $\mathscr{C}$ is the class of smallest invariant sets containing $\mathscr{C}_1=\{\{a\leqq V\leqq b\}; a, b\in R\}$. It is not difficult to see that ψ commutes with complementation, countable unions and intersections and since $\mathscr{C}_1$ generates $\mathscr{V}$ it follows that $\mathscr{C}$ generates $\mathscr{W}$ and the proof for the case $\alpha>0$ is complete. The other cases can be treated in the same way.

Corollary. *Suppose that* $\{X_n: n\geqq 0\}$ *is a properly homogeneous chain,* $\{Y_n: n\geqq 0\}$ *converges a.s. to a proper random variable* V *and* $\mathscr{I}$ *is trivial. Then one of the following two cases occurs*

(i) V *is constant with probability* 1,

(ii) V *assumes the countable set of values*

$$\{\gamma\,\alpha^n+\beta(\alpha^n-1)(\alpha-1)^{-1};\ n=0,1,\ldots\}\cup\{(\gamma-\beta)\,\alpha^n-\beta\,\alpha^{-1}(\alpha^{-n+1}-1)(\alpha^{-1}-1)^{-1}; n=1,2,\ldots\}\qquad \text{if } \alpha\neq 1$$

and

$$\{\gamma+n\beta;\quad n=\ldots-1,0,1,\ldots\}\qquad \text{if } \alpha=1$$

for some $\gamma\in R$.

Proof. Suppose that F is strictly increasing on a certain interval (a,b) and that $\alpha>0$. Then we can choose two numbers $a', b'\in(a,b)$ such that $F(b')-F(a')>0$ and either $a'\alpha+\beta>b'$ or $b'\alpha+\beta<a'$. Assume for definiteness that $a'\alpha+\beta>b'$; we get that $\Lambda=\bigcup_{n=-\infty}^{\infty} T^n\{a\leqq V\leqq b\}$ is an invariant set and $0<\mathscr{P}(\Lambda)<1-(F(a'\alpha+\beta)-F(b'))<1$ which contradicts the triviality of $\mathscr{I}$. Therefore V is a discrete random variable. Suppose now that γ is a number with the property $\mathscr{P}(V=\gamma)>0$. Then Proposition 1(iii) in conjunction with Lemma 2 implies that $\mathscr{P}(T^n\{V=\gamma\})>0$ for all $n\in Z$. Since $\bigcup_{n=-\infty}^{\infty} T^n\{V=\gamma\}$ is an invariant set, its probability must be 1 and the proof is complete.

We shall further give a result that parallels Theorem 4(ii) in the case when the sequence $\{Y_n: n\geqq 0\}$ converges a.s. to an improper random variable and $\lim_{n\to\infty} a_{n+1}/a_n=\infty$. Such a situation occurs for some branching processes with infinite mean (see [17]) where $\lim_{n\to\infty} X_n/a_n=V$ a.s. with $\mathscr{P}(V=\infty)>0$ and $\mathscr{P}(V\neq\{0,\infty\})>0$. Using a reasoning similar to that employed in the proof of Theorem 4 we can show that $\lim_{n\to\infty}\{a\leqq a_n(X_{n+1}+b_n)\leqq b\}=T\{a\leqq V\leqq b\}$ a.s. Further $T\{a\leqq V\leqq b\}$ assumes positive probability provided that $\mathscr{P}(\{a\leqq V\leqq b\})>0$. Under these conditions Theorem 1.1 of [16] implies that $V_1=\lim_{n\to\infty} a_n(X_{n+1}+b_n)$ exists a.s. It is not hard to see that $\lim_{n\to\infty} a_{n+1}/a_n=\infty$ entails

$$\mathscr{P}(\{-\infty<V<\infty, V\neq 0\}\cap\{-\infty<V_1<\infty, V_1\neq 0\})=0$$

and hence, unlike the situation described by Theorem 4(ii), V_1 is not expressible as a function of V.

Theorem 5. *Suppose that $\{X_n : n \geqq 0\}$ is a properly homogeneous Markov chain, $\{Y_n : n \geqq 0\}$ converges a.s. to an improper and nondegenerate random variable V and that $\lim_{n\to\infty} a_{n+1}/a_n = \infty$. Then $\lim_{n\to\infty} a_n(X_{n+k} + b_n) = V_k$ exists a.s. and is nondegenerate for all $k \in Z$; $\mathscr{W} \subset \mathscr{V}$ a.s. and $\mathscr{W}$ is generated by the family of invariants sets $\mathscr{D} = \left\{ \bigcup_{n=-\infty}^{\infty} \{a \leqq V_n \leqq b\}, a, b \in R \right\}$.*

The finite mean supercritical branching processes (see [2]) and the irregular branching processes with infinite mean (see [17]) provide examples of Markov chains to which Theorems 4 and 5 respectively apply. Let us notice that any non-degenerate supercritical branching process $\{Z_n : n \geqq 0\}$ is a properly homogeneous chain. Indeed, suppose that x and y are two states such that $P(1,x)P(1,y) > 0$ and let t be the greatest common divisor of the numbers $\{i : i \geqq 1, P(1,i) > 0\}$. According to a result by Dubuc (Proposition 2, [11]) there exists a number d such that any $j \in \{i : x^n + d \leqq i \leqq y^n - d\}$ with $j = x^n (\bmod t)$ is accessible at time n from 1 (i.e. $P^{(n)}(1,j) > 0$) for $n = 1, 2, \ldots$. If we choose $x = \min\{i : i \geqq 1, P^{(n)}(1,i) > 0\}$ then we can prove that for n sufficiently large $\mu_j^{(n)} > 0$ for all $j = x^n (\bmod t)$ with $j > x^n + d$. Toward this aim let us notice that given $\mu_2^{(0)} > 0$, if we choose n such that $y^n - d > 2x^n + 2d$ then any $j \in \{i : 2x^n + 2d \leqq i \leqq 2y^n - 2d\}$ with $j = 2x^n (\bmod t)$ is accessible at time n from 2. But $\{i : x^n + d \leqq i \leqq y^n - d\}$ and $\{i : 2x^n + 2d \leqq i \leqq 2y^n - 2d\}$ overlap and therefore $\mu_j^{(n)} > 0$ for all $j \in \{i : x^n + d \leqq i \leqq 2y^n - 2d\}$ with $j = x^n (\bmod t)$. Further $\mu_4^{(0)} > 0$ and $y^n - d > 2x^n + 2d$ also implies that $\{i : 2x^n + 2d \leqq i \leqq 2y^n - 2d\}$ and $\{i : 4x^n + 4d \leqq i \leqq 4y^n - 4d\}$ overlap, etc. We conclude that for n sufficiently large $\mu_j^{(n)} > 0$ for all $j = x^n (\bmod t)$ with $j > x^n + d$. Therefore $[S_n - S_{n+1}] \subset \{1, 2, \ldots, x^n + d\}$.

Suppose now that $m = E(Z_1) < \infty$. Then the Seneta-Heyde theorem ([2], p. 30) asserts that there exists a continuously distributed random variable W and some norming constants $\{c_n\}$ with $\lim_{n\to\infty} c_{n+1}/c_n = m$ such that $\lim Z_n/c_n = W$ a.s. Now $c_n = c_n/c_{n-1}\, c_{n-1}/c_{n-2} \cdots c_1$ yields that there exists a number A and an integer k such that $c_n \geqq A(m-\varepsilon)^{n-k}$ for a preassigned $\varepsilon > 0$ and a sufficiently large n. If we choose ε such that $m - \varepsilon > x$ and take into account that $\mathscr{P}(W > 0) = \mathscr{P}(\lim_{n\to\infty} Z_n = \infty \text{ a.s.})$ we get that $\lim_{n\to\infty} \mathscr{P}(Z_n \in [S_n - S_{n+1}]) = 0$. Hence $\{Z_n : n \geqq 0\}$ is properly homogeneous.

In the infinite mean case $\{Z_n : n \geqq 0\}$ grows quicker to infinity and the general theory of such processes given by Schuh and Barbour [17] can be easily seen to imply that these processes are properly homogeneous (in both the regular and irregular cases).

We mention that Athreya and Ney [2] Chapter 2, p. 96 have identified the invariant sets $\left\{ \bigcup_{n=-\infty}^{\infty} \{m^n < W \leqq m^{x+n}\}, x \in R \right\}$ for a finite mean supercritical branching process, under the additional assumption that $E(Z_1 \log Z_1^+) < \infty$. This result is a particular case of Theorem 4(ii). Athreya and Ney derived two different proofs of their result on p. 96–97 of [2] but neither of them seems extendable to the general case considered here.

References

1. Abrahamse, A.F.: The tail σ-field of a Markov chain. Ann. Math. Statist. **40**, 127–136 (1969)
2. Athreya, K.B., Ney, P.E.: Branching Processes. New York: Springer 1972
3. Blackwell, D.: On transient Markov processes with a countable number of states and stationary transition probabilities. Ann. Math. Statist. **26**, 654–658 (1955)
4. Breiman, L.: On transient Markov chains with application to the uniqueness problem for Markov processes. Ann. Math. Statist. **28**, 499–503 (1957)
5. Chung, K.L.: Markov chains with stationary transition probabilities. 2nd Edition. New York: Springer 1967
6. Cohn, H.: On the tail events of a Markov chain. Z. Wahrscheinlichkeitstheorie verw. Gebiete **29**, 65–72 (1974)
7. Cohn, H.: On the tail σ-field of the countable Markov chains. Rev. Roumaine Math. Pures Appl. **4**, 850–858 (1976)
8. Cohn, H.: On the norming constants occurring in convergent Markov chains. Bull. Austral. Math. Soc. **17**, 193–205 (1977)
9. Derriennic, Y.: Lois "zero ou deux" pour les processus de Markov. Applications aux marches aléatoires. Ann. Inst. H. Poincaré. Section B. XII, **2**, 111–129 (1976)
10. Doob, J.L.: A Markov chain theorem. Probability and Statistics. (H. Cramér memorial volume) ed. U. Grenander, 50–57. Stockholm and New York: Almqvist and Wiksel, 1959
11. Dubuc, S.: Etats accessibles dans un processus de Galton-Watson. Canad. Math. Bull. **17**, 111–113 (1974)
12. Griffeath, D.: Partial coupling and loss of memory for Markov chains. Ann. Probability **4**, 850–858 (1976)
13. Küchler, U.: Über die σ-Algebra der asymptotischen Ereignisse bei diskreten Geburts- und Todesprozessen. Math. Nachr. **65**, 321–329 (1975)
14. Loève, M.: Probability Theory. Third Edition. Princeton: Van Nostrand, 1963
15. Ornstein, D. and Sucheston, L.: An operator theorem on L_1 convergence to zero with applications to Markov kernels. Ann. Math. Statist. **41**, 1631–1639 (1970)
16. Padmanabhan, A.R.: Convergence in probability and allied results. Math. Jap. **15**, 111–117 (1970)
17. Schuh, H.-J., Barbour, A.: On the asymptotic behaviour of the branching processes with infinite mean. Advances Appl. Probability **9**, 681–723 (1977)

Received September 10, 1977; in revised form January 2, 1979