Sankhyā : The Indian Journal of Statistics
1990, Volume **52**, Series A, Pt. 1, pp. 58-83.

A MATRIX LIMIT THEOREM WITH APPLICATIONS TO PROBABILITY THEORY

By PREM S. PURI
Purdue University, West Lafayette
JAMES B. ROBERTSON
University of California, Santa Barbara
and
KALYAN B. SINHA
Indian Statistical Institute

SUMMARY. The classical Poisson limit theorem studies the limit laws of S_n where $S_n = \sum_{j=1}^{n} X_{jn}$ and $X_{1n}, \ldots, X_{nn}$ is a sequence of $\{0, 1\}$ valued, independent, identically distributed random variables. In this paper we will weaken the independence assumption and investigate the possible limit laws for certain types of dependent sequences. This leads us to the study of the limit of $(A_n(s))^n$ where s is a real parameter and $A_n(s)$ is a finite dimensional (the dimension being fixed) matrix of the form $A_n(s) = R(s)+n^{-1}(Q(s)+B_n(s))$ where $\lim_{n \to \infty} B_n(s) = 0$. This problem seem to be of independent interest but does not appear to have been treated in the literature.

1. Introduction

Let $X_{1n}, X_{2n}, \ldots, X_{nn}$ be *independent* Bernoulli random variables with $P(X_{kn} = 1) = 1-P(X_{kn} = 0) = p_n$ for $k = 1, 2, \ldots, n$. Let $S_n = \sum_{j=1}^{n} X_{jn}$. The well known Poisson limit theorem states that $\lim_{n \to \infty} P(S_n = j) = q_j$, where $(q_0, q_1, \ldots)$ is a probability distribution on Z^+ if and only if $\lim_{n \to \infty} np_n = \lambda \geqslant 0$ exists, and that in this case the limit distribution is Poisson with mean λ, i.e. $q_j = \lambda^j \exp(-\lambda)/(j!)$, $j = 0, 1, 2, \ldots$ In this paper we examine the limit behaviour of the distribution of S_n while allowing a certain type of dependence among the X_{jn}'s, which still take values 0 and 1.

A natural way to relax the independence assumption is to assume that $\{X_{1n}, X_{2n}, \ldots, X_{nn}\}$ is stationary. (A sequence $\{X_1, X_2, \ldots, X_n\}$ is said to be

*This research was supported in part by U.S. National Science Foundation Grant No : DMS-850-4319.

AMS (1980) *subject classification :* Primary 60F05 ; Secondary 15A48.

Key words and phrases : Poisson limit theorem, Markov chain, finitary process, Jordan form, matrix limit theorem.

stationary if the joint distribution of $\{X_1, X_2, \ldots, X_{n-1}\}$ and $\{X_2, X_3, \ldots, X_n\}$ are the same. This is equivalent to the condition that there is a stationary process $\{Y_1, Y_2, \ldots\}$ such that $\{X_1, X_2, \ldots, X_n\}$ and $\{Y_1, Y_2, \ldots, Y_n\}$ have the same joint distribution.) If $p_1, \ldots, p_n$ is any probability vector, we may set

$$P(X_1 = x_1, \ldots, X_n = x_n) = p_j \binom{n}{j}^{-1}, \qquad \ldots \quad (1)$$

for every vector $x = (x_1, \ldots, x_n) \in \{0, 1\}^n$ such that $\sum_{k=1}^{n} x_k = j, j = 0, 1, 2, \ldots, n$. This is a stationary sequence since if $x_1 + \ldots + x_{n-1} = j$, then

$$\begin{aligned} P(X_1 &= x_1, \ldots, X_{n-1} = x_{n-1}, X_n = 0) \\ &\quad + P(X_1 = x_1, \ldots, X_{n-1} = x_{n-1}, X_n = 1) \\ &= p_j \binom{n}{j}^{-1} + p_{j+1} \binom{n}{j+1}^{-1} \\ &= P(X_1 = 0, X_2 = x_1, \ldots, X_n = x_{n-1}) \\ &\quad + P(X_1 = 1, X_2 = x_1, \ldots, X_n = x_{n-1}) \qquad \ldots \quad (2) \end{aligned}$$

For this distribution it is clear that $P(S_n = j) = p_j$. It is therefore apparent that one can obtain any limit law if one only assumes stationarity, and that we need to add more restrictions on the sequence $\{X_{kn}\}$ to obtain meaningful results.

In section 4, we shall assume that for some fixed integer $d \geqslant 1$, $\{X_{1n}, \ldots, X_{nn}\}$ is a so called finitary process of dimension d (see section 4). The case with $d = 1$ is just the case where $\{X_{1n}, \ldots, X_{nn}\}$ are independent and identically distributed as above. The finitary processes form a fairly rich class of processes ; in particular they include all functions of finite state Markov chains. To begin with we first consider this latter special case in some detail below.

Consider a d-state $(2 \leqslant d < \infty)$ Markov chain $\{Y_{jn}\}$ with a stationary one step transition matrix P given by

$$P = \begin{bmatrix} P_{11} & 0 \\ P_{21} & P_{22} \end{bmatrix} \qquad \ldots \quad (3)$$

where the $k \times k$ Markov matrix P_{11} corresponds to the states $\{1, 2, \ldots, k\}$, $1 \leqslant k < d$, which is assumed to form an irreducible aperiodic (necessarily positive recurrent) class, while the remaining states $\{k+1, \ldots, d\}$ corresponding to the $(d-k) \times (d-k)$ matrix P_{22} are all assumed to be transient. Thus starting from any of these transient states the process will move to the recurrent class $\{1, \ldots, k\}$ in finite time with probability one. For $i = k+1$, let τ_i be the first passage time to the class $\{1, \ldots, k\}$ given that the process starts with state i at time zero ; with the corresponding probability generating function (p.g.f.) given by

$$G_i(s) = E(s^{\tau_i}),\ i = k+1, \ldots, d,\ |s| \leqslant 1. \qquad \ldots (4)$$

Again let $(\pi_1, \pi_2, \ldots, \pi_k)$, with $\pi_i > 0$, $i = 1, 2, \ldots, k$, and $\sum_{i=1}^{k} \pi_j = 1$, denote the stationary distribution corresponding to the matrix P_{11}. By a well known property of irreducible aperiodic finite state Markov chains we have

$$\lim_{n\to\infty} (P_{11})^n = \begin{bmatrix} \pi_1 & \pi_2 & \ldots & \pi_k \\ & & \ldots & \\ \pi_1 & \pi_2 & \ldots & \pi_k \end{bmatrix} \equiv \Pi^* \text{ (say)}. \qquad \ldots (5)$$

In fact, if we introduce the dummy variable s and define

$$R(s) = \begin{bmatrix} P_{11} & 0 \\ P_{21} & s.\,P_{22} \end{bmatrix},\ |s| \leqslant 1, \qquad \ldots (6)$$

then it can be easily shown that

$$\lim_{n\to\infty} R(s)^n \equiv \lim_{n\to\infty} ((R(s)^n)_{ij})$$

$$= \begin{bmatrix} \pi_1 & \ldots & \pi_k & 0 & \ldots & 0 \\ & \ldots & & & \ldots & \\ \pi_1 & \ldots & \pi_k & 0 & \ldots & 0 \\ \pi_1 G_{k+1}(s) & \ldots & \pi_k G_{k+1}(s) & 0 & \ldots & 0 \\ & \ldots & & & \ldots & \\ \pi_1 G_d(s) & \ldots & \pi_k G_d(s) & 0 & \ldots & 0 \end{bmatrix}$$

$$\equiv \quad \Pi(s) \equiv \begin{bmatrix} \Pi^*_{[k\times k]} & 0_{[k\times(d-k)]} \\ \Pi^*_G(s)_{[(d-k)\times k]} & 0_{[(d-k)\times(d-k)]} \end{bmatrix}, \qquad \ldots (7)$$

where at the end we have conveniently written the limit matrix $\Pi(s)$ as displayed.

Suppose now that the true situation is such that the transition matrix P is perturbed a bit in an n-dependent manner, creating thereby a sequence P_n of transition matrices given by

$$P_n = \begin{bmatrix} P_{11}(n) & P_{12}(n) \\ P_{21} & P_{22} \end{bmatrix}, \qquad \ldots (8)$$

where

$$P_{11}(n) = P_{11} + n^{-1}Q_{11} + o(n^{-1}) \text{ and } P_{12}(n) = n^{-1}Q_{12} + o(n^{-1}). \qquad \ldots (9)$$

Since for each n, P_n is a stochastic matrix, we must have

$$\begin{bmatrix} Q_{11} & Q_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \ldots (10)$$

where $\mathbf{1} = (1, 1, \ldots, 1)^T$. This means at each step we do allow the corresponding Markov chain $\{Y_{in}\}$ to move from the states in the set $\{1, \ldots, k\}$ to those in the set $\{k+1, \ldots, d\}$ but with increasing rarity with increasing n. Now if for $j = 1, 2, \ldots, n$, we define $X_{jn} = 1$ if $Y_{jn} \in \{k+1, \ldots, d\}$, and $X_{jn} = 0$ otherwise, then $S_n = \sum_{j=1}^{n} X_{jn}$ represents the number of visits the Markov chain $\{Y_{jn}\}$ pays to the set $\{k+1, \ldots, d\}$ during the first n steps. We will be interested in the limit behaviour of the distribution of S_n as $n \to \infty$ or equivalently that of the corresponding p.g.f. To this end, once again by introducing the dummy variable s in (8), we define

$$P_n(s) = \begin{bmatrix} P_{11}(n) & s.P_{12}(n) \\ P_{21} & s.P_{22} \end{bmatrix}$$

$$= R(s)+n^{-1}(Q(s)+B_n), \qquad \ldots \quad (11)$$

where $R(s) = \lim_{n \to \infty} P_n(s)$ is as given by (6), B_n is such that $\lim_{n \to \infty} B_n = 0$ and $Q(s) = \lim_{n \to \infty} n(P_n(s)-R(s))$ is given by

$$Q(s) = \begin{bmatrix} Q_{11} & s.Q_{12} \\ 0 & 0 \end{bmatrix}. \qquad \ldots \quad (12)$$

For $k = 1, 2, \ldots, d$ let

$$g_{kn}(s) = E(s^{S_n} \mid Y_{0n} = k), \ |s| \leqslant 1 \qquad \ldots \quad (13)$$

be the p.g.f. of S_n given that the process starts at k. Then it can be easily seen that

$$(g_{1n}(s), \ldots, g_{dn}(s))^T = (P_n(s))^n \mathbf{1}. \qquad \ldots \quad (14)$$

Thus in order to study the limit behaviour of (14) as $n \to \infty$, we must study the limit behaviour of $(P_n(s))^n$, that is of

$$(R(s)+n^{-1}(Q(s)+B_n))^n. \qquad \ldots \quad (15)$$

Again if $R(s)$ were the identity matrix I the limit behaviour of (15) is well known (see Kato, 1982, 35-36). On the other hand, if R and Q were commuting with each other, using the result for the identity matrix case and the limit behaviour of $(R(s))^n$ as given by (7), one could easily establish the limit result for (15). Unfortunately in general R and Q do not commute and in order to establish the result in this generality a lot more work needs to be done. This is precisely the content of the next section. A similar matrix power limit question arises for the more general case of finitary processes. This will be taken up in section 4 as an application of the key result to be established in the next section.

2. An operator limit theorem

Here we shall prove the key result (Theorem 2.1) which helps answer the questions raised above and is of independent interest. We shall present the result in the general setting of operators in a complex normed linear space of finite dimensions.

Let X be a d(finite) dimensional complex normed linear space, and let $\mathcal{B}(X)$ denote the normed linear space of all linear operators in X equipped with the operator norm, denoted by $\|.\|$. All convergences of operator sequences in what follows will be in this norm. We shall also denote by $\rho(A)$ the spectral radius of $A \in \mathcal{B}(X)$ so that

$$\rho(A) = \lim_{n \to \infty} \|A^n\|^{n^{-1}} = \inf_n \|A^n\|^{n^{-1}} = \max_j |\lambda_j|, \qquad \ldots \quad (16)$$

where $\{\lambda_j\}$ are the eigenvalues of A.

We state the theorem first and prove it with the help of a series of lemmas.

Theorem 2.1 : *Let R and Q be in $\mathcal{B}(X)$, and let D be a bounded region in $\mathcal{B}(X)$. Suppose furthermore that $B_n : D \to \mathcal{B}(X)$, $n = 1, 2, \ldots$, be such that $\|B_n(Q)\| \to 0$ uniformly in $Q \in D$ as $n \to \infty$.*

(a) *If $\lim_{n \to \infty} R^n = \Pi$ exists, them*

$$\lim_{n \to \infty} [R + n^{-1}(Q + B_n(Q))]^n = \Pi . \exp(\Pi . Q . \Pi), \qquad \ldots \quad (17)$$

the convergence being uniform in $Q \in D$.

(b) *If $\lim_{N \to \infty} N^{-1} \sum_{n=1}^{N} R^n = \Pi$ exists, then*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} [R + n^{-1}(Q + B_n(Q))]^n = \Pi . exp(\pi . Q . \Pi), \qquad \ldots \quad (18)$$

convergence being uniform in $Q \in D$.

We shall 'need the Jordan canonical form for any member of $\mathcal{B}(X)$ (see section 5, ch. 1 of Kato, 1982). Thus any A in $\mathcal{B}(X)$ admits the unique decomposition

$$A = \sum_{i=1}^{r} (\lambda_i P_i + D_i), \qquad \ldots \quad (19)$$

where λ_i is the i-th distinct eigenvalue of A, P_i and D_i are the corresponding eigenprojection and eigennilpotents respectively, and r is the number of distinct eigenvalues of A. It is also known that (see pages 41-43 of Kato, 1982)

$$P_i P_j = \delta_{ij} P_i, \; P_i D_j = D_j P_i = \delta_{ij} D_i$$

$$D_i^{M_i - m_i + 1} = 0, \; D_i D_j = 0 \text{ if } i \neq j, \qquad \ldots \quad (20)$$

where M_i equals the algebraic multiplicity of λ_i, which is equal to the dimension of the range of P_i, and m_i equals the geometric multiplicity of λ_i, which is equal to the dimension of the space of all eigenvectors corresponding to the eigenvalue λ_i. We shall also use the important property that $P_i, D_i, D_i^2, \ldots, D_i^{M_i-m_i}$ are linearly independent in the vector space of all $d \times d$ matrices. Note that $\sum_{i=1}^{r} M_i = d$, $\sum_{i=1}^{r} P_i = I$, and $1 \leqslant m_i \leqslant M_i$. Thus $D_i = 0$ if and only if $m_i = M_i$.

The next lemma studies the structure of an operator R satisfying the hypotheses (a) or (b) of Theorem 2.1.

Lemma 2.2 : *Let* $R \in \mathcal{B}(X)$ *have the Jordan form* (19). *Then*

(a) $R^n \to \Pi$ *as* $n \to \infty$ *if and only if for each* i ($i = 1, 2, \ldots, r$) *either* $\lambda_i = 1$ *with corresponding* $D_i = 0$, *or* $|\lambda_i| < 1$. *Furthermore, in the former case* $P_i \cdot \Pi = \Pi \cdot P_i = P_i$ *while* $P_i \Pi = \Pi \cdot P_i = 0$ *in the latter.*

(b) $N^{-1} \sum_{n=1}^{N} R^n \to \Pi$ *as* $N \to \infty$ *if and only if for each* $i (i = 1, 2, \ldots, r)$ *either* $|\lambda_i| = 1$ *with corresponding* $D_i = 0$ *or* $|\lambda_i| < 1$. *Furthermore,* $P_i \cdot \Pi = \Pi \cdot P_i$ *if* $\lambda_i = 1$, *and otherwise* $P_i \,.\, \Pi = \Pi \cdot P_i = 0$.

Proof : By (19) and (20) we see that for $n \geqslant \max_i (M_i - m_i)$

$$R^n = \sum_{j=1}^{r} (\lambda_j P_j + D_j)^n = \sum_{j=1}^{r} \left(\sum_{k=1}^{M_j - m_j} \binom{n}{k} \lambda_j^{n-k} D_j^k + \lambda_j^n P_j \right), \quad \ldots (21)$$

so that if $m_i < M_i$, then

$$D_i^{M_i-m_i} \cdot R^n = R^n \cdot D_i^{M_i-m_i} = \lambda_i^n D_i^{M_i-m_i}, \quad \ldots (22)$$

and

$$P_i R^n = R^n P_i = \lambda_i^n P_i + \sum_{k=1}^{Mi-m_i} \binom{n}{k} \lambda_i^{n-k} D_i^k. \quad \ldots (23)$$

From (22), (23) and the linear independence of $P_i, D_i, D_i^2, \ldots, D_i^{M_i-m_i}$, for each i, it follows that $R^n \to \Pi$ if and only if for each i, *either* $|\lambda_i| < 1$, in which case $P_i \Pi = \Pi P_i = 0$, or $\lambda_i = 1$, in which case $D_i = 0$ and $P_i \Pi = \Pi P_i = P_i$. This proves part (a).

Since in (21) the terms corresponding to those j for which $|\lambda_j| < 1$ converge to zero as $n \to \infty$, their Cesaro-limit will also be zero. Consequently the proof of the 'if' part of (b) follows by averaging (21) over n and observing that

$$\frac{1}{N} \sum_{n=1}^{N} \lambda_i^n = \frac{\lambda_i (1 - \lambda_i^N)}{N(1 - \lambda_i)} \to 0,$$

as $N \to \infty$, for $|\lambda_i| = 1, \lambda_i \neq 1$.

For the 'only if' part of (b), from (22) we have that $N^{-1} \sum_{n=1}^{N} \lambda_i^n$ must converge as $N \to \infty$, for all λ_i. This necessarily means that $|\lambda_i| \leqslant 1$, for all i. Again from this and (23) it follows that

$$\frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{M_i - m_i} \binom{n}{k} \lambda_i^{n-k} D_i^k$$

must also converge for each i as $N \to \infty$. However we show below that if $|\lambda_i| = 1$, then $N^{-1} \sum_{n=k}^{N} \binom{n}{k} \lambda_i^{n-k}$ does not converge as $N \to \infty$, implying thereby that the corresponding $D_i = 0$. Here we have again used the linear independence of $P_i, D_i, D_i^2, \ldots, D_i^{M_i - m_i}$.

Since $N^{-1} \sum_{n=k}^{N} \binom{n}{k} \to \infty$, as $N \to \infty$, for $k \geqslant 1$, we may assume that $|\lambda_i| = 1$ and $\lambda_i \neq 1$. Furthermore a simple calculation shows that

$$\frac{1}{N} \sum_{n=k}^{N} \binom{n}{k} \lambda_i^{n-k} = \frac{1}{N(k!)} \frac{d^k}{d\lambda_i^k} \left(\sum_{n=k}^{N} \lambda_i^n \right)$$

$$= \frac{1}{N(k!)} \frac{d^k}{d\lambda_i^k} \left(\frac{\lambda_i^{N+1} - \lambda_i^k}{\lambda_i - 1} \right)$$

$$\equiv \sum_{j=-1}^{k-1} a_j(\lambda_i, N, k)\, N^j, \qquad \ldots (24)$$

where for $-1 \leqslant j \leqslant k-1$, the $a_j(\lambda_i, N, k)$ are absolutely bounded with respect to N. In particular

$$a_{k-1}(\lambda_i, N, k) = \frac{\lambda_i^{N+1-k}}{k\,!(\lambda_i - 1)}; \; 1 \leqslant k \leqslant M_i - m_i. \qquad \ldots (25)$$

Because in (24), the power of N varies from term to term, it is enough to note that for $k \geqslant 2$, $a_{k-1} \neq 0$, and a_0 oscillates as $N \to \infty$. Thus the left hand side of (24) does not converge for every $k = 1, 2, \ldots, M_i - m_i$; $i = 1, 2, \ldots, r$.□

Remark 2.1 : It is clear from Lemma 2.2 that if $\Pi \neq 0$, then there is one i with $\lambda_i = 1$ and this i we set equal to 1 by convention, so that $\lambda_1 = 1$. With this convention we can write $\Pi = P_1$, the projection corresponding to the eigenvalue 1. Thus in the case (a) of Lemma 2.2, we can write

$$R = P_1 + \Lambda, \text{ with } P_1 \cdot \Lambda = \Lambda \cdot P_1 = 0, \text{ and } \rho(\Lambda) < 1. \qquad \ldots (26)$$

Similarly in case (b) of Lemma 2.2, we can write

$$R = R_0 + \Lambda \equiv \sum_{i=1}^{s} \lambda_i P_i + \Lambda, \quad \ldots (27)$$

where $|\lambda_i| = 1$ for $i = 1, 2, \ldots, s$ $(s \leqslant r)$, $R_0 \cdot \Lambda = \Lambda \cdot R_0 = 0$. It follows from Lemma 2.2 that

$$\|R_0\| = 0 \text{ or } 1, \text{ and } \rho(\Lambda) < 1. \quad \ldots (28)$$

We note that in this case that $R_0 = R \cdot P_0$ where $P_0 = \sum_{i=1}^{s} P_i$. Also observe that Λ is not necessarily diagonalizable in either case, and from (16) it follows that there exists a positive integer n_0 such that

$$\|\Lambda^n\| < 1 \text{ and } \|R^n\| \leqslant 1 \ \forall\, n \geqslant n_0. \quad \ldots (29)$$

On the other hand, $\Pi = 0$ in case (a) if and only if $|\lambda_i| \leqslant 1$ for all i without any λ_i being equal to 1.

Lemma 2.3 : *Let* $A \in \mathcal{B}(X)$, R *be as in Lemma* 2.2, *and* $A(n) = n^{-1} \sum_{j=0}^{n-1} R^j \cdot A \cdot R^{n-1-j}$, *where n is a positive integer greater than or equal to one. Then*

(a) *in both cases* (a) *and* (b) *of Lemma* 2.2, $\Pi \cdot A(n) \cdot \Pi = \Pi \cdot A \cdot \Pi$,

(b) *In case* (b) *of Lemma* 2.2, *given any* $\epsilon > 0$ *there exists a positive integer* $n = n(\epsilon)$ *such that*

$$\left\| \frac{1}{n} \sum_{j=0}^{n-1} R_0^j - \Pi \right\| < \epsilon, \quad \ldots (30)$$

where R_0 *is as in* (27), *and*

$$\|\Lambda^n\| \leqslant 1. \quad \ldots (31)$$

Furthermore

$$\|P_0 \cdot A(n) \cdot P_0\| \leqslant \|A\| \quad \ldots (32)$$

and

$$\|(P_0 . A(n) . P_0 - \Pi . A . \Pi) . \Pi\| \leqslant \|A\| \epsilon. \quad \ldots (33)$$

Proof of (a) : By Remark 2.1, $\Pi = P_1$. Hence $\Pi . R = R . \Pi = \Pi$ and the result follows.

Proof of (b) : Since $R_0 = RP_0 = P_0 R$, we consider R_0 to be an operator acting on the space $P_0 X$ and set $R_0^0 = P_0$. Then

$$\frac{1}{n} \sum_{j=0}^{n-1} R_0^j - \Pi = \sum_{k=2}^{s} \left(\frac{1}{n} \sum_{j=0}^{n-1} \lambda_k^j \right) P_k,$$

and thus

$$\left\| \frac{1}{n} \sum_{j=0}^{n-1} R_0^j - \Pi \right\| = \frac{1}{n} \sup \left\{ \left| \sum_{j=0}^{n-1} \lambda_2^j \right|, \ldots, \left| \sum_{j=0}^{n-1} \lambda_s^j \right| \right\}.$$

But

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} \lambda_k^j \right| = \left| \frac{(1-\lambda_k^n)}{n(1-\lambda_k)} \right| \leqslant 2 \left(\inf_{2 \leqslant k \leqslant s} |1-\lambda_k| \right)^{-1} n^{-1} \to 0$$

as $n \to \infty$ uniformly in k, and we have (30). Inequality (31) is the same as inequality (29). That $\|P_0 . A(n) . P_0\| \leqslant \|A\|$ follows from the relation

$$P_0 . A(n) . P_0 = \frac{1}{n} \sum_{j=0}^{n-1} R_0^j . A . R_0^{n-1-j}$$

and the fact that $\|R_0\| = 1$. Finally,

$$P_0 . A(n) . P_0 - \Pi . A . \Pi) . \Pi = \left(\frac{1}{n} \sum_{j=0}^{n-1} R_0^j . A . R_0^{n-1-j} - \Pi . A . \Pi \right) . \Pi$$

$$= \left(\frac{1}{n} \sum_{j=0}^{n-1} R_0^j - \Pi \right) . A . \Pi$$

and (33) follows from (30). □

Remark 2.2 : It is clear from the proof of Lemma 2.3 that even when d is countably infinite, the lemma will be valid if 1 is not an accumulation point of the eigenvalues of R_0. On the other hand if 1 is an accumulation point of the eigenvalues of R_0 in a separable Hilbert space X, then a simple application of the dominated convergence theorem shows that the strong limit of $n^{-1} \sum_{j=0}^{n-1} R_0^j$ as $n \to \infty$ is Π.

Lemma 2.4 : *Let* $A_n(Q)$, $C_n(Q)$, $\in \mathcal{B}(X)$, $n = 1, 2, \ldots$, $Q \in D \subseteq \mathcal{B}(X)$ *be such that* $\|A_n(Q)\|^k \leqslant M$ *for all* n, $Q \in D$, $k \leqslant n$, *and* $\|C_n(Q)\| \to 0$ *as* $n \to \infty$ *uniformly in* $Q \in D$. *Then* $\|(A_n(Q)+n^{-1} C_n(Q))^n - (A_n(Q))^n\| \to 0$ *uniformly in* $Q \in D$.

Proof : The proof follows from the inequality

$$\|(A_n(Q)+n^{-1} C_n(Q))^n - (A_n(Q))^n\| \leqslant \sum_{j=1}^{n} \binom{n}{j} n^{-j} \|A_n(Q)\|^{n-j} \|C_n(Q)\|^j$$

$$\leqslant M . \|C_n(Q)\| . \exp(\|C_n(Q)\|). \ \square$$

It is clear from the above lemma that it suffices to prove Theorem 2.1 in both cases with $B_n(Q) = 0$. Lemmas 2.5 and 2.7 are approximation results preparing the stage for the proof of the main theorem while Lemma 2.6 is an auxiliary result used in the proof of Lemma 2.7.

Lemma 2.5 : *Let* $A \in \mathcal{B}(X)$ *such that* $A = A_0+A_1$ *with* $A_0.A_1 = A_1.A_0 = 0$, $\|A_0\| \leqslant 1$ *and* $\|A_1\| < 1$. *Also let* P_0 *be a projection in* X *such that* $A_0 = A.P_0 = P_0.A$. *Then* $\|(A+n^{-1}B(Q))^n-(A_0+n^{-1}\ P_0.B(Q).P_0)^n\| \to 0$ *as* $n \to \infty$, *uniformly in* $Q \in D$, *where* $B(Q) \in \mathcal{B}(X)$ *is such that* $\|B(.)\|$ *is bounded on* D *and* $(I-P_0).B(Q).P_0 = 0$.

Proof: For brevity, set $B_{00} = P_0.B(Q).P_0$, $B_{01} = P_0.B(Q).(I-P_0)$, $B_{10} = (I-P_0).B(Q).P_0 = 0$ and $B_{11} = (I-P_0).B(Q).(I-P_0)$. Then we write

$$A+n^{-1}B(Q) = (A_0+n^{-1}B_{00})+(A_1+n^{-1}B_{11})+n^{-1}B_{01}. \qquad \ldots \quad (34)$$

Set

$$A_n \equiv (A_0+n^{-1}B_{00})+(A_1+n^{-1}B_{11}). \qquad \ldots \quad (35)$$

It is easy to see that $(A_n^i.\ B_{01}.\ A_n^j).\ (A_n^k.\ B_{01}.\ A_n^l) = 0$ for all $i, j, k, l \geqslant 0$. Therefore

$$\begin{aligned}
&(A+n^{-1}\ B(Q))^n-(A_0+n^{-1}\ B_{00})^n \\
&= (A_n+n^{-1}B_{01})^n-A_n^n+(A_1+n^{-1}B_{11})^n \\
&= \sum_{j=0}^{n-1} A_n^{n-1-j}.(n^{-1}B_{01}).\ A_n^j+(A_1+n^{-1}B_{11})^n \\
&= \frac{1}{n}\sum_{j=0}^{n-1} (A_0+n^{-1}B_{00})^{n-1-j}.\ B_{01}.\ (A_1+n^{-1}\ B_{11})^j \\
&\quad +(A_1+n^{-1}\ B_{11})^n. \qquad \ldots \quad (36)
\end{aligned}$$

Since $\|A_1\| \leqslant 1$, it follows that $\|A_1+n^{-1}\ B_{11}\| < 1$ for all $Q \in D$ and sufficiently large n so that $\|(A_1+n^{-1}\ B_{11})^n\| \to 0$ as $n \to \infty$. The norm of the first term in the right hand side of (36) is bounded by $n^{-1}\|B_{01}\| \exp(\|B_{00}\|) \cdot \sum_{j=0}^{\infty} (\|A_1\|+n^{-1}\|B_{11}\|)^j$ showing that this term also converges to 0 in norm as $n \to \infty$. □

Note that Lemma 2.5 is essentially a special case of Theorem 2.1(a) where $B(Q) = Q$ and $(I-P_0)B(Q)P_0 = 0$, $\forall\ Q \in D$. This latter condition will be removed in Lemma 2.7.

Lemma 2.6 : *Let* $\theta, a > 0$ *and* l, N *positive integers. Then* $\prod_{j=1}^{l} (1+\theta a^{i_j})$ *attains its maximum (subject to the conditions :* i_j *is a positive integer for* $j = 1, 2, \ldots, l$ *and* $\sum_{j=1}^{l} i_j = N$*) when all the* i_j*'s except one are equal to unity.*

Proof: Without loss of generality we may assume that $a \neq 1$. For $l = 2$ it is equivalent to maximizing $a^i + a^{N-i}$ over $1 \leqslant i \leqslant N-1$ which can be easily seen to occur at $i = 1$ or $i = N-1$. Continuing by induction on l, we have

$$\max\left\{\prod_{j=1}^{l+1}(1+\theta a^{i_j}) : \sum_{j=1}^{l+1} i_j = N,\ i_1, \ldots, i_{l+1} \geqslant 1\right\}$$

$$= \max_{1 \leqslant k \leqslant N-l}\left\{(1+\theta a^k)\max\left\{\prod_{j=1}^{l}(1+\theta a^{i_j}) : \sum_{j=1}^{l} i_j = N-k,\ i_1, \ldots, i_l \geqslant 1\right\}\right.$$

$$= (1+\theta a)^{l-1}\max_{1 \leqslant k \leqslant N-l}(1+\theta a^k)(1+\theta a^{N-k-l+1})$$

$$= (1+\theta a)^l(1+\theta a^{N-l})$$

as desired. □

Lemma 2.7 : *Let A, P_0, A_0, A_1 be as in Lemma 2.5 and $B(Q) \in \mathcal{B}(X)$ be such that $\|B(.)\|$ is bounded on D. Then $\|(A+n^{-1}B(Q))^n - (A_0+n^{-1} P_0 . B(Q). P_0)^n\| \to 0$ as $n \to \infty$ uniformly in $Q \in D$.*

Proof: We define B_{11}, B_{01}, B_{10} as in the proof of Lemma 2.5 and note that now $B_{10} = (I-P_0).\ B.\ P_0$ is not assumed to be 0. Then $A+n^{-1} B(Q) = \tilde{A}_n + n^{-1}B_{10}$, where $\tilde{A}_n = (A_0+n^{-1}B_{00}) + (A_1+n^{-1}B_{11}) + n^{-1}B_{01}$. By Lemma 2.5 $\|\tilde{A}_n^n - (A_0+n^{-1}B_{00})^n\| \to 0$, uniformly in $Q \in D$ as $n \to \infty$.

Expanding $(\tilde{A}_n + n^{-1}B_{10})$ and observing that $B_{10}^2 = 0$ we have

$$(A+n^{-1}B(Q))^n = \tilde{A}_n^n + \sum_{k=2}^{[(n+3)/2]} \sum_{j=1}^{4} I_j(k), \qquad \ldots (37)$$

where $[a]$ is the greatest integral part of the positive real number a and

$$I_1(k) = \Sigma_{(1)}\, \tilde{A}_n^{i_1}. (n^{-1} B_{10}). \tilde{A}_n^{i_2}. (n^{-1} B_{10}) \ldots \tilde{A}_n^{i_{k-1}}. (n^{-1} B_{10}),$$

$$I_2(k) = \Sigma_{(2)}\, (n^{-1} B_{10}). \tilde{A}_n^{i_2}. (n^{-1} B_{10}) \ldots \tilde{A}_n^{i_{k-1}}. (n^{-1} B_{10}) . \tilde{A}_n^{i_k},$$

$$I_3(k) = \Sigma_{(3)}\, (n^{-1} B_{10}) . \tilde{A}_n^{i_2} . (n^{-1} B_{10}) \ldots \tilde{A}_n^{i_{k-1}} . (n^{-1} B_{10}),$$

$$I_4(k) = \Sigma_{(4)}\, \tilde{A}_n^{i_1}. (n^{-1} B_{10}) . \tilde{A}_n^{i_2} . (n^{-1} B_{10}) \ldots \tilde{A}_n^{i_{k-1}} . (n^{-1} B_{10}). \tilde{A}_n^{i_k} \ldots. \quad (38)$$

In (37) $k-1$ is the number of times B_{10} appears in the expansion, and each sum in (38) runs over the corresponding i's taking values greater than or equal to one and adding up to $n-k+1$ for each k. By convention $I_3(2) = 0$. It is easy to see that the number of terms in $\Sigma_{(1)}$ and $\Sigma_{(2)}$ are $\binom{n-k}{k-2}$ while $\Sigma_{(3)}$ and $\Sigma_{(4)}$ have $\binom{n-k}{k-3}$ and $\binom{n-k}{k-1}$ terms respectively. The desired result shall follow if we show that the sum in (37) converges to 0 uniformly in $Q \epsilon D$ as $n \to \infty$.

A simple calculation shows that

$$\tilde{A}_n^m = (A_0+n^{-1} B_{00})^m+(A_1+n^{-1} B_{11})^m + \frac{1}{n} \sum_{i=0}^{m-1} (A_0+n^{-1} B_{00})^i . B_{01} . (A_1+n^{-1} B_{11})^{m-1-i}, \quad \ldots \quad (39)$$

$$\tilde{A}_n^m . B_{10} = (A_1+n^{-1} B_{11})^m . B_{10} + \frac{1}{n} \sum_{i=0}^{m-1} (A_0+n^{-1} B_{00})^i . B_{01} . (A_1+n^{-1}B_{11})^{m-1-i} B_{10}. \quad \ldots \quad (40)$$

Thus as in the proof of Lemma 2.5, we have for all $m \leqslant n$ and $Q \epsilon D$

$$\|\tilde{A}_n^m\| \leqslant M_1 \text{ and } \|\tilde{A}_n^m . B_{10}\| \leqslant \frac{M_2}{n} (1+n\alpha^m), \quad \ldots \quad (41)$$

where M_1, M_2 are two absolute constants and $\alpha = \sup_{Q \epsilon D} (\|A_1\|+n^{-1}\|B_{11}\|) < 1$ for sufficiently large n.

By (38) and (41) we have

$$\|I_1(k)\| \leqslant \binom{n-k}{k-2} (n^{-1}\|B_{10}\|)^{k-1} M_1^{k-1}$$

$$\leqslant \frac{M_3^{k-1}}{n(k-2)!}$$

and similarly

$$\|I_2(k)\| \leqslant \frac{M_3^{k-1}}{n(k-2)!}$$

$$\|I_3(k)\| \leqslant \frac{M_4^{k-2}}{n^2(k-3)!}$$

where M_3 and M_4 are absolute constants. It is now easily seen that $\sum_{k=2}^{[(n+3/2]} \|I_j(k)\| \to 0$ as $n \to \infty$ uniformly in $Q \epsilon D$ for $j = 1, 2, 3$.

Finally from (38) and (41) it follows that

$$\|I_4(k)\| \leqslant M_1(n^{-1})^{k-1}\Sigma_{(4)} \prod_{j=1}^{k-1} \|\bar{A}_n^{i_j} \,.\, B_{10}\|$$

$$\leqslant M_1 M_2^{k-1} n^{-(2k-2)} \sum_{i=1}^{n-2k+2} \Sigma_{(5)} \prod_{j=1}^{k-1} (1+n\alpha^{i_j}), \quad \ldots \quad (42)$$

where $\Sigma_{(5)}$ for every fixed i (the value of i_k) runs over $\{(i_1, \ldots, i_{k-1}) : \sum_{j=1}^{k-1} i_j = n-k-i+1, i_1, \ldots, i_{k-1} \geqslant 1\}$. Also note that $\Sigma_{(5)}$ has $\binom{n-k-i}{k-2}$ terms. By Lemma 2.6 and (42) we conclude that

$$\|I_4(k)\| \leqslant \frac{M_1 M_2^{k-1}}{n^{2k-2}} \sum_{i=1}^{n-2k+2} \binom{n-k-i}{k-2} (1+n\alpha)^{k-2}(1+n\alpha^{n-2k-i+3})$$

$$\leqslant \frac{M_5\, M_6^{k-1}}{(k-2)!} \sum_{i=1}^{n-2k+2} \left[\frac{1}{n^2}+\frac{\alpha^{n-2k-i+3}}{n}\right],$$

where M_5 and M_6 are suitable constants. Since $\alpha < 1$, the above inequality leads to the result that $\lim_{n\to\infty} \sum_{k=2}^{[(n+3)/2]} \|I_4(k)\| = 0$. □

Proof of Theorem 2.1 : In case (a) choose a positiveinteger ν so that (29) is satisfied. In case (b) given any arbitrary $\epsilon > 0$ choose $\nu = \nu(\epsilon)$ as in Lemma 2.3(b). Having chosen this ν we hold it fixed, and write $n = l\nu+\mu$ with $l = 0, 1, 2, \ldots$ and $\mu = 0, 1, 2, \ldots, \nu-1$. Since μ is bounded, we note that $n \to \infty$ if and only if $l \to \infty$. Thus as in Lemma 2.3, defining $Q(\nu) = \nu^{-1} \sum_{j=0}^{\nu-1} R^j Q R^{\nu-1-j}$, we find that

$$(R+n^{-1}Q)^n-(R^\nu+l^{-1}Q(\nu))^l.\ R^\mu$$
$$=[(R+(l\nu+\mu)^{-1}Q)^{l\nu}-(R^\nu+l^{-1}Q(\nu))^l].\ R^\mu$$
$$+(R+(l\nu+\mu)^{-1}Q)^{l\nu}.\ (R+(l\nu+\mu)^{-1}Q)^\mu-R^\mu]. \quad \ldots \quad (43)$$

Note that

$$(R+(l\nu+\mu)^{-1}Q)^\nu=R^\nu+(l\nu+\mu)^{-1} \sum_{j=0}^{\nu-l} R^j Q R^{\nu-1-j}+O(l^{-2}),$$

where $O(\cdot)$ is uniform in $Q \epsilon D$ as $l \to \infty$. Furthermore since $(l\nu+\mu)^{-1} = (l\nu)^{-1}-\mu(l\nu)^{-1}(l\nu+\mu)^{-1}$, from the above we obtain

$$(R+(l\nu+\mu)^{-1}Q)^\nu=R^\nu+l^{-1}Q(\nu)+l^{-1}C_l(\nu, Q), \quad \ldots \quad (44)$$

where $C_l(\nu,Q) \to 0$ uniformly in $Q \in D$ as $l \to \infty$. Using (29), we have $\|R^\nu + l^{-1}Q(\nu)\|^k \leqslant (1+l^{-1}\|Q\|\,\|R\|^{\nu-1})^k \leqslant \exp\,(\|Q\|\,\|R\|^{\nu-1}) \leqslant$ constant (depending only on ν) for all $k \leqslant l$ and $Q \in D$. Using this and (44) we can apply Lemma 2.4 to the first term on the right hand side of (43) to conclude that its norm converges to 0 as $l \to \infty$, uniformly in $Q \in D$. The fact that $(R^\nu + l^{-1}Q(\nu))^l$ is uniformly bounded also tells us that $\|(R+(l\nu+\mu)^{-1}Q)^{l\nu}\|$ is uniformly bounded in l and $Q \in D$. Since μ is bounded, we see that $\|(R+(l\nu+\mu)^{-1}Q)^\mu - R^\mu\| \leqslant \sum_{k=1}^{\mu} \binom{\mu}{k} \|R\|^{\mu-k}\, ((l\nu+\mu)^{-1}\,\|Q\|)^k \to 0$ as $l \to \infty$ uniformly in $Q \in D$. Thus we arrive at the convergence of $(R+n^{-1}Q)^n - (R^\nu + l^{-1}Q(\nu))^l .\, R^\nu$ to 0 as $l \to \infty$ uniformly in Q. By the choice of ν, $R^\nu = R_0^\nu + \Lambda^\nu$ with $\|\Lambda^\nu\| < 1$ and an application of Lemma 2.7 leads us to

$$\|(R+n^{-1}Q)^n - (R_0^\nu + l^{-1}P_0 \,.\, Q(\nu) \,.\, P_0)l.\, R_0^\mu\| \to 0 \qquad \dots \quad (45)$$

as $l \to \infty$ uniformly in $Q \in D$.

Case (a) : In this case $R_0 = P_0 = P_1 = \Pi$ and we get the desired result for $B_n(Q) = 0$ (and hence also for $B_n(Q) \neq 0$ by the remark after Lemma 2.4) by applying Lemma 2.3(a) and the result that $\lim_{l\to\infty} (I+l^{-1}A)^l = \exp(A)$ uniformly for all A in a bounded set of $\mathcal{B}(X)$ (see Kato 1982, 35—36).

Case (b) : Write $N = L\nu + J$ with $L = 0, 1, 2, \dots,\ J = 0, 1, 2, \dots, \nu-1$. Since $\|R_0^\nu + L^{-1}P_0 \,.\, Q(\nu) \,.\, P_0\|^L \leqslant (1+L^{-1}\|Q\|)^L \leqslant \exp\,(\|Q\|)$, we get that

$$\lim_{N\to\infty} N^{-1}(R_0^\nu + L^{-1}P_0 \,.\, Q(\nu) \,.\, P_0)^L \sum_{\mu=J+1}^{\nu-1} R_0^\mu = 0 \qquad \dots \quad (46)$$

uniformly in $Q \in D$. Thus by (45) one has

$$\lim_{N\to\infty} \frac{1}{N} \left\| \sum_{n=1}^{N} (R+n^{-1}Q)^n - \sum_{l=0}^{L} (R_0^\nu + l^{-1}P_0 \,.\, Q(\nu)P_0)^l \sum_{\mu=0}^{\nu-1} R_0^\mu \right\| = 0 \qquad \dots \quad (47)$$

also uniformly in $Q \in D$.

Given an arbitrary $\epsilon > 0$ we have fixed ν so that (30) is satisfied and hence by Lemma 2.3(b)

$$\left\| \frac{\nu}{N} \sum_{l=0}^{l} (R_0^\nu + l^{-1}P_0 \,.\, Q(\nu) \,.\, P_0)^l \,.\, \left\{ \frac{1}{\nu} \sum_{\mu=0}^{\nu-1} R_0^\mu - P_1 \right\} \right\| \leqslant \epsilon \exp\,(\|Q\|). \quad \dots \quad (48)$$

Since $P_1 = \Pi$ (see Remark 2.1), we have by (32) and (33)

$$\|(R_0^\nu + l^{-1}P_0 \, . \, Q(\nu) \, . \, P_0)^l \, . \, \Pi - (\Pi + l^{-1}\Pi \, . \, Q \, . \, \Pi)^l \, . \, \Pi\|$$

$$= \|[(R_0^\nu + l^{-1}P_0 \, . \, Q(\nu) \, . \, P_0)^l - (R_0^\nu + l^{-1}\Pi \, . \, Q \, . \, \Pi)^l] \, . \, \Pi \,\|$$

$$\leqslant \frac{1}{l} \sum_{j=0}^{l-1} \|(R_0^\nu + l^{-1}P_0 \, . \, Q(\nu) \, . \, P_0)\|^j \|(P_0 \, . \, Q(\nu) \, . \, P_0 - \Pi \, . \, Q \, . \, \Pi) \, . \, \Pi \,\|$$

$$. \, \|I + l^{-1}\Pi \, . \, Q \, . \, \Pi \,\|^{l-1-j}$$

$$\leqslant \|Q\| \exp(2\|Q\|)\epsilon. \qquad \ldots \quad (49)$$

Combining (47), (48), (49) and the standard result referred to earlier in case (a) we arrive at

$$\lim_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} (R + n^{-1}Q)^n - \Pi, \exp(\Pi \, . \, Q \, . \, \Pi) \right\| < C\epsilon,$$

where C is a constant independent of Q and ϵ. Since ϵ is arbitrary, we have the desired result for $B_n(Q) = 0$ and by the remark after Lemma 2.6 we arrive at the general result. □

3. Markov chains

Returning to the Markov chain example discussed in Section 1, in view of (10), using (7) with $s = 1$, it easily follows that

$$Q_{11} \, . \, \Pi^* + Q_{12} \, . \, \Pi_G(1)^* = 0. \qquad \ldots \quad (50)$$

In view of (7), applying Theorem 2.1(a) to (15) and using (12), we have

$$\lim_{n\to\infty} (P_n(s))^n = \Pi(s) \, . \, \exp(\Pi(s) \, . \, Q(s) \, . \, \Pi(s)) = \Pi(s) \, . \, \exp(Q(s) . \Pi(s))$$

$$= \Pi(s) \, . \begin{bmatrix} \exp(Q_{11} \, . \, \Pi^* + sQ_{1\mathrm{v}} \, . \, \Pi_G^*) & 0 \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} \Pi^* \, . \, \exp(Q_{11} \, . \, \Pi^* + sQ_{12} \, . \, \Pi_G^*(s)) & 0 \\ \Pi_G^*(s) \, . \, \exp(Q_{11} \, . \, \Pi^* + sQ_{12} \, . \, \Pi_G^*(s)) & 0 \end{bmatrix} . \qquad \ldots \quad (51)$$

Finally using this in (14) we have the desired limit result for the distribution of S_n in terms of its p.g.f. given by

$$[g_1(s) \ldots g_d(s)]^T \equiv \lim_{n\to\infty} [g_{1n}(s) \ldots g_{dn}(s)]^T$$

$$= \begin{bmatrix} \Pi^* \\ \Pi_G^*(s) \end{bmatrix}_{d\times k} . \, \exp(Q_{11} \, . \, \Pi^* + sQ_{12} \, . \, \Pi_G^*(s)) \, . \, \mathbf{1}_{k\times 1}$$

$$= \begin{bmatrix} \Pi^* \\ \Pi_G^*(s) \end{bmatrix} . \, \exp(-Q_{12} \, . \, (\Pi_G^*(1) - s\Pi_G^*(s))) \, . \, \mathbf{1} \qquad \ldots \quad (52)$$

where at the end we have used (50) and where $\Pi_G^*(1)-s\Pi_G^*(s)$ is a $(d-k)\times k$ matrix given by

$$\Pi_G^*(1)-s\Pi_G^*(s) = \begin{bmatrix} \Pi_1(1-sG_{K+1}(s)) & \dots & \Pi_k(1-sG_{k+1}(s)) \\ \dots & \dots & \dots \\ \Pi_1(1-sG_d(s)) & \dots & \Pi_k(1-sG_d(s)) \end{bmatrix}. \quad \dots \quad (53)$$

The above limit $[g_1(s)\dots g_d(s)]$ gives us a matrix analog of the p.g.f. of a compound Poisson distribution. Again instead of asking for the limit behaviour of S_n, the time spent in the set $\{k+1, \dots, d\}$ during the first n steps, we might consider studying the limit behaviour of the joint distribution of the vector $[S_n(k+1)\dots S_n(d)]$ where $S_n(j)$ is the time spent in the state j during the first n steps, with $j = k+1, \dots, d$ and $S_n = \sum_{j=k+1}^{d} S_n(j)$. The above analysis can be easily adjusted by introducing $d-k$ dummy variables $(s_{k+1}, \dots, s_d)$ instead of a single dummy variable s. An application of Theorem 2.1 as before, would this time lead to a matrix analog of the p.g.f. of a multivariate compound Poisson distribution.

4. Finitary processes

4.1 *Generalities.* In section 3 we started with a Markov chain $\{Y_{1n}, \dots, Y_{nn}\}$ with state space $\mathcal{S} = \{1, \dots, d\}$ and then studied the *function of this process* given by $X_{in} = f(Y_{in})$ were f was the indicator function of the set $\{k+1, \dots, d\}$. We shall now describe the concept of finitary processes (see Robertson, 1973 for more details) which generalizes the notion of function of Markov chains and at the same time allows one to apply the matrix theory methods of Markov chains.

Let S be a finite set (we will take $S = \{0, 1\}$). Let d be a positive integer (d will be fixed throughout this section) and η and $\xi \in \boldsymbol{R}^d$. For seach $i \in S$, let A_i be a $d\times d$ matrix. We assume the following axioms are satisfied.

Axiom 4.1 : $\eta^T . \xi = 1$,

Axiom 4.2 : $A . \xi = \xi$, *where* $A = \sum_{i \in S} A_i$,

Axiom 4.3 : *For every positive integer m and for all $i_1, \dots, i_m \in S$ we have*

$$\eta^T . A_{i_1} \dots A_{i_m} . \xi \geqslant 0.$$

A system $(\boldsymbol{R}^d ; \eta ; \xi ; A_i, i \in S)$ satisfying Axioms 4.1, 4.2 and 4.3 is called a *finitary system.* This system is said to be *stationary* if

$$\eta^T . A = \eta^T. \quad \dots \quad (54)$$

For every finitary system ($\boldsymbol{R}^d$; η ; ξ ; A_i, $i \in S$) there exists a stochastic process $\{X_1, X_2, \ldots\}$ such that for every positive integer m and for all $i_1, \ldots, i_m \in S$ we have

$$P[X_1 = i_1, \ldots, X_m = i_m] = \eta^T . A_{i_1} \ldots A_{i_m} . \xi. \quad \ldots \quad (55)$$

Such a stochastic process $\{X_1, X_2, \ldots\}$ is called a *finitary process.* The following summarizes several facts whose proofs can be found in Robertson (1973).

Proposition 4.1 : (a) *If* ($\boldsymbol{R}^d$; η ; ξ ; A_i, $i \in S$) *is a finitary sy tem, then there exists another finitary system* ($\boldsymbol{R}^{d'}$; η' ; ξ' ; A'_i, $i \in S$) *(called a reduced system) with* $d' \leqslant d$, *such that*

(1) *For every positive integer m and for all $i_1, \ldots, i_m \in S$,* $\eta'^T . A'_{i_1} \ldots A'_{i_m} . \xi'$ $= \eta^T . A_{i_1} \ldots A_{i_m} . \xi$.

(2) $\{A'_{i_1} \ldots A'_{i_m} . \xi' : m \geqslant 0, i_1, \ldots, i_m \in S\}$ *spans* $\boldsymbol{R}^{d'}$.

(3) $\{\eta'^T . A'_{i_1} \ldots A'_{i_m} : m \geqslant 0, i_1, \ldots, i_m \in S\}$ *spans* $\boldsymbol{R}^{d'}$.

(b) *Every function of a finite Markov chain is a finitary process, and a finitary process is a function of a Markov chain if and only if it has a finitary system* (R^d ; η ; ξ ; A_i, $i \in S$) *such that all the entries of the vectors η and ξ and all the entries of the matrices A_i, $i \in S$ are nonnegative.*

(c) *A finitary process is stationary if and only if its reduced finitary system is stationary.*

From the above proposition it is clear that without loss of generality we may assume that the system is reduced. However, the system with nonnegative entries referred to in Proposition 4.1(b) will in general not be reduced.

We now take $S = \{0, 1\}$ and suppose for each positive integer n that $\{X_{1n}, \ldots, X_{nn}\}$ is a finitary process given by the reduced finitary system ($\boldsymbol{R}^d$; η_n ; ξ_n ; $A_i(n)$, $i \in S$). We want to find the possible limit laws of $S_n = \sum_{i=1}^{n} X_{in}$, as $n \to \infty$. For this we look at the p.g.f., $g_n(s)$, of S_n, which is easily seen, using (55), to be given by

$$g_n(s) = \eta_n^T . (A_n(s))^n . \xi_n, \quad \ldots \quad (56)$$

where $0 \leqslant s \leqslant 1$ and $A_n(s) = A_0(n) + sA_1(n)$.

In the sequel we assume that η_n and ξ_n both converge as $n \to \infty$. Consequently it is sufficient to study the possible nonzero limits of $(A_n(s))$ as $n \to \infty$. This we do in the next subsection for the case with $d = 2$.

4.2 *Finitary processes with* $d = 2$. In Theorem 2.1 we studied the limit of $(A_n)^n$ with

$$A_n = R + \frac{1}{n}(Q + B_n(Q)) \qquad \ldots \quad (57)$$

given that R^n converges and $\|B_n(Q)\| \to 0$, as $n \to \infty$ uniformly in $Q \in D$. We now study the converse question (with $d = 2$) namely that given $(A_n)^n$ converges as $n \to \infty$ for the same form of A_n as above, we prove that R^n does converge. This along with Theorem 2.1 will allow us to compute the form of the limit of $(A_n(s))^n$ as $n \to \infty$ for a finitary process with $d = 2$.

Theorem 4.2 : *Let* $A_n = R + n^{-1}(Q + B_n(O))$ *with* R, Q *and* $B_n(Q)$ *as defined in Theorem* 2.1. *Assume furthermore that* $(A_n)^n$ *converges to a nonzero limit, say* C. *Then* R^n *converges to a nonzero operator, say* Π.

Remark 4.1 : The proof of this theorem is given in the appendix. Note that our Theorem 4.2 is limited to A_n's specifically of the form (57). In fact if we only assume that $A_n \to R$, $(A_n)^n \to C (\neq 0)$ and $R^n \to \Pi (\neq 0)$ hold, then (57) is not necessarily satisfied. For instance take

$$A_n = \left(1 + \frac{\alpha}{n}\right) P_1 + \frac{1}{\sqrt{n}} P_2, \qquad \ldots \quad (58)$$

with α real and P_1 and P_2 are mutually orthogonal projections with their sum equal to the identity.

Finitary systems with $d = 2$ can be given fairly explicitly. Consider the closed convex cone $\mathcal{C}$ consisting of all limits of all linear combinations with nonnegative coefficients of vectors of the form $A_{i_1} \ldots A_{i_m} . \xi$ where $m \geqslant 0$ and $i_1, \ldots, i_m \in \{0, 1\}$. This cone is clearly invariant under the operators A_i for all $i \in S$. By Axiom 4.3, $\mathcal{C} \neq \boldsymbol{R}^2$ and the case were $\mathcal{C}$ is one dimensisional is trivial. Thus there are two linear independent vectors, α and β such that $\mathcal{C} = \{a\alpha + b\beta : a, \ b \geqslant 0\}$. Since $\xi \in \mathcal{C}$, we may choose α and β such that $\xi = \alpha + \beta$. We first express the matrices ξ, A_0 and A_1 in terms of basis (α, β). The fact that $\mathcal{C}$ is invariant under A_0 and A_1 implies that the entries of A_0 and A_1 are nonnegative. Next Axiom 4.2 states that $A_0 + A_1$ is a stochastic matrix. From this it also follows that

$$\xi^T = [1, 1], \ \eta^T = [p, 1-p], \text{ for } 0 \leqslant p \leqslant 1. \qquad \ldots \quad (59)$$

If the process is stationary then η will be the invariant probability vector for A_0+A_1. We thus have

$$A(s) = A_0+sA_1 = \begin{bmatrix} 1-a-c-(1-s)b & a+sc \\ d+se & 1-d-e-(1-s)f \end{bmatrix} \quad \dots \quad (60)$$

where all the parameters are nonnegative, $a+b+c \leqslant 1$, and $d+e+f \leqslant 1$. The eigenvalues of $A(s)$ are given by

$$\lambda_{\pm} = \frac{1}{2}\,[\alpha(s)+\beta(s)\pm\sqrt{(\alpha(s)-\beta(s))^2+4(a+sc)(d+se)}] \quad \dots \quad (61)$$

where $\alpha(s) = 1-a-(1-s)b-c$ and $\beta(s) = 1-d-e-(1-s)f$.

Suppose now that the parameters a, b, c, ... depend on n with the matrix in (60) denoted by $A_n(s)$. Since we are interested only in the nonzero limits of $(A_n(s))$ with $A_n(s)$ of the form (57), in view of Theorem 4.2 and the Remark 2.1, the maximal eigenvalue of $A_n(s)$, λ_{+n} in (61), must converge to one as $n \to \infty$. The next lemma gives the necessary and sufficient conditions for this to occur. The proof of the lemma, being straightforward, is omitted.

Lemma 4.3 : (a) $0 \leqslant |\lambda_{-n}| \leqslant \lambda_{+n} \leqslant 1$.

(b) $\lambda_{+n} \to 1$ *if and only if at least one of the following three conditions holds*

(i) a_n, b_n *and* $c_n \to 0$.

(ii) d_n, e_n *and* $f_n \to 0$.

(iii) b_n, c_n, e_n and $f_n \to 0$.

We will consider the three cases of Lemma 4.3 one by one, but first we consider a "case 0." Suppose that a_n, b_n, c_n, d_n, e_n and $f_n \to 0$. In accord with (57) we assume that $na_n \to a$, $nb_n \to b$, $nc_n \to c$, $nd_n \to d$, $ne_n \to e$ and $nf_n \to f$. Then

$$R(s) = \lim_{n\to\infty} A_n(s) = I,\ \Pi = \lim_{n\to\infty} R^n = I \quad \dots \quad (62)$$

$$Q = \lim_{n\to\infty} n(A_n(s)-I) = \begin{bmatrix} -a-(1-s)b-c & a+sc \\ d+se & -d-e-(1-s)f \end{bmatrix} \dots \quad (63)$$

We can easily calculate the exponential of a 2×2 matrix B with trace τ, determinant δ, and eigenvalues $\lambda_{\pm}$. The matrix satisfies its characteristic polynomial $B^2 = \tau B - \delta I$. This can be used to obtain a difference equation for B^n. If the eigenvalues are distinct, the result is

$$B^n = \frac{\lambda_+^n - \lambda_-^n}{\lambda_+ - \lambda_-} B - \lambda_+\lambda_- \frac{\lambda_+^{n-1} - \lambda_-^{n-1}}{\lambda_+ - \lambda_-} I.$$

If λ is the only eigenvalue, then

$$B^n = n\lambda^{n-1}B - (n-1)\lambda^n I.$$

From this we may calculate respectively

$$\exp(B) = \sum_{n=0}^{\infty} \frac{B^n}{n!} = \frac{e^{\lambda_+} - e^{\lambda_-}}{\lambda_+ - \lambda_-} B - \frac{\lambda - e^{\lambda_+} - \lambda_+ e^{\lambda_-}}{\lambda_+ - \lambda_-} I,$$

$$\exp(B) = e^{\lambda}B + (1-\lambda)e^{\lambda}I.$$

Thus from (63) we obtain

$$\exp(Q) = \frac{(e^{\lambda_+} - e^{\lambda_-})Q + (\hat{\lambda}_+ e^{\lambda_-} - \hat{\lambda}_- e^{\lambda_+})I}{\hat{\lambda}_+ - \hat{\lambda}_-} \qquad \ldots \quad (64)$$

where

$$\hat{\lambda}_\pm = \hat{\lambda}_\pm(s) = \frac{1}{2}\left(\hat{\alpha} + \hat{\beta} \pm \sqrt{(\hat{\alpha} - \hat{\beta})^2 + 4(a+sc)(d+se)}\right),$$

$$\hat{\alpha} = \hat{\alpha}(s) = -a - (1-s)b - c, \; \hat{\beta} = \hat{\beta}(s) = -d - e - (1-s)f, \qquad \ldots \quad (65)$$

$$a, b, c, d, e, f \geqslant 0.$$

The conditions that the two eigenvalues of Q would be equal for all s implies that $Q = -\hat{\lambda}(1-s)I$, which implies that the limiting distribution is Poisson. Thus we may assume that (64) is the form of $\exp(Q)$ except for the two values of s that are the roots of the quadratic equation $(\hat{\alpha} - \hat{\beta})^2 + 4(a+sc)(d+se) = 0$. For $\exp(Q)$ as in (64) the limiting p.g.f. of S_n is given by

$$\lim_{n \to \infty} \eta^T [A_n(s)]^n \xi = [p, 1-p].\exp(Q).\xi$$

$$= \frac{\theta(1-s)(e^{\hat{\lambda}_-} - e^{\hat{\lambda}_+}) + \hat{\lambda}_+ e^{\hat{\lambda}_-} - \hat{\lambda}_- e^{\hat{\lambda}_+}}{\hat{\lambda}_+ - \hat{\lambda}_-} \qquad \ldots \quad (66)$$

where $\theta = p(b+c) + (1-p)(e+f) \geqslant 0$ and $\hat{\lambda}_\pm = \hat{\lambda}_\pm(s)$ are as in (65). In contrast to the cases considered below we are not able to describe the distribution of this p.g.f. in terms of well known distributions. In what follows we treat cases (1), (2), and (3), but assume that we are not in "case 0."

Next consider the case 1 of Lemma 4.3b where $a_n \to 0$, $b_n \to 0$ and $c_n \to 0$. We suppose that $na_n \to a$, $nb_n \to b$, $nc_n \to c$, $d_n \to d$, $e_n \to e$ and $f_n \to f$ where $d+e+f > 0$. Also we suppose that $n(d_n - d) \to x$, $n(e_n - e) \to y$ and $n(f_n - f) \to z$.

One may then calculate the following matrices as before. For $0 \leqslant s \leqslant 1$, we have

$$R(s) = \begin{bmatrix} 1 & 0 \\ d+se & 1-d-e-(1-s)f \end{bmatrix} \quad \ldots \quad (67)$$

$$Q = \begin{bmatrix} -a-(1-s)b-c & a+sc \\ x+sy & -x-y-(1-s)z \end{bmatrix} \quad \ldots \quad (68)$$

$$\Pi = \begin{bmatrix} 1 & 0 \\ \dfrac{d+se}{d+e+(1-s)f} & 0 \end{bmatrix} \quad \ldots \quad (69)$$

$$\Pi \,.\, \exp(\Pi \,.\, Q \,.\, \Pi) = \begin{bmatrix} e^{-\gamma(s)} & 0 \\ \delta(s)e^{-\gamma(s)} & 0 \end{bmatrix} \quad \ldots \quad (70)$$

where

$$\gamma(s) = \frac{(1-s)\,(ae+af+bd+be+cd+cf+ce+bf+s(ce-bf))}{d+e+(1-s)f}$$

$$\delta(s) = \frac{d+se}{d+e+(1-s)f}.$$

Finally

$$\lim_{n \to \infty} \eta^T[A_n(s)]^n \xi = [p,\, 1-p].\Pi.\exp\,(\Pi.Q.\Pi).\xi = (p+(1-p)\delta(s))e^{-\gamma(s)}. \quad \ldots \quad (71)$$

This distribution has a fairly nice interpretation. We will transform the parameters to make this more apparent. Set

$$\lambda = \gamma(0) = \frac{a(e+f)}{d+e+f}+b+c \geqslant 0$$

$$q = \frac{f}{d+e+f}$$

$$1-r = \frac{af(e+f)+c(e+f)\,(d+e+f)}{(d+e+f)\,[a(e+f)+(b+c)\,(d+e+f)]}$$

$$1-r' = \delta(0) = \frac{d}{d+e+f}.$$

Note that q, r, r' all belong to $[0, 1)$ and $\gamma(s) = \lambda\left(1-s\left(1-r+\frac{r(1-q)s}{1-qs}\right)\right)$. Then $e^{-\gamma(s)}$ is the p.g.f. of

$$Y = \begin{cases} 0 & \text{if } N = 0 \\ \sum_{n=1}^{N} X_n & \text{if } N \geqslant 1 \end{cases}$$

where $\{N, X_1, X_2, \ldots\}$ are independent random variables with N being Poisson with mean λ and $P[X_n = 1] = 1-r$ and $P[X_n = m] = r(1-q)\, q^{m-2}$ for $m \geqslant 2$ and all n. The factor $p+(1-p)\delta(s)$ in (71) is the p.g.f of a random variable Z for which

$$P[Z = 0] = p+(1-p)\,(1-r')$$

$$P[Z = n] = (1-p)\, r'(1-q)q^{n-1},\; n \geqslant 1.$$

Thus the p.g.f. in (71) is the p.g.f. of the sum of the two independent random variables Y and Z respectively.

The second case of Lemma 4.3(b) where $d_n \to 0$, $e_n \to 0$ and $f_n \to 0$ can be dealt with as in the previous case by interchanging the states and thus does not yield any new limit laws.

Consider now the final case of Lemma 4.3(b) where $b_n, c_n, e_n, f_n \to 0$ as $n \to \infty$. We suppose that $a_n \to a$ and $d_n \to d$ with $a \geqslant 0$, $d \geqslant 0$ and $a+d > 0$. We also suppose that $nb_n \to b$, $nc_n \to c$, $ne_n \to e$, $nf_n \to f$, $n(a_n - a) \to x$ and $n(d_n - d) \to y$. Then as before we obtain from (57)

$$R = \begin{bmatrix} 1-a & a \\ d & 1-d \end{bmatrix}, \quad \Pi = \frac{1}{a+d}\begin{bmatrix} d & a \\ d & a \end{bmatrix}$$

$$Q = \begin{bmatrix} -x-c-(1-s)b & x+sc \\ y+se & -y-e-(1-s)f \end{bmatrix}$$

$$\Pi\,.\,Q\,.\,\Pi = -(1-s)\lambda\Pi$$

where $\lambda = (ae+af+bd+cd)/(a+d)$. Finally with η and ξ as in (71) we have

$$\eta^T\,.\,\Pi\,.\,\exp(\Pi\,.\,Q\,.\,\Pi)\,.\,\xi = e^{-\lambda(1-s)}, \qquad \ldots \quad (72)$$

which is the p.g.f. of a Poisson random variable with mean λ.

This completes the discussion of the two dimensional finitary case.

Appendix

Here we give the proof of Theorem 4.2, for which we need a result about countinuity of eigenvalues, eigenprojections and eigenilpotents under continuous perturbation and some lemmas.

Let $G_n (n = 1, 2, \ldots)$ and $G \epsilon \mathcal{B}(X)$, $\dim(X) = d < \infty$, and let their Jordan forms be given as :

$$\left.\begin{aligned} G_n &= \sum_{i=1}^{\gamma_n} [\lambda_i(n)P_i(n)+D_i(n)], \\ G &= \sum_{j=1}^{\gamma} [\lambda_j P_j+D_j]. \end{aligned}\right\} \quad \ldots (73)$$

For each $\lambda_j (j = 1, \ldots, \gamma)$ define $S(\lambda_j \equiv \{i \mid 1 \leqslant i \leqslant d, \lim_{n\to\infty} \lambda_i(n) = \lambda_j\}$. Then we have the following version of a theorem in Kato, 1982, (Theorem 5.1, pages 107-108).

Theorem A.1 : *Let G_n converge to G as $n \to \infty$. Then for $j = 1.2, \ldots, \gamma$,*

(a) *the sets $S(\lambda_j)$ are non-empty and mutually disjoint,*

(b) $\sum_{i \epsilon s(\lambda_j)} P_i (n)$ *and* $\sum_{i \epsilon s(\lambda_j)} D_i(n)$ *converge to P_j and D_j respectively as $n \to \infty$.*

The following lemma is given without proof and will be needed in the sequel.

Lemma A.2 : *Let a_n be a sequence of complex numbers converging to a with $|a| = 1$. Furrthermore, let $(a_n)^n$ converge to 0 as $n \to \infty$. Then $n \mid a_n - a \mid \to \infty$ as $n \to \infty$.*

For $d = 2$, we write the Jordan forms for A_n, R, C of Theorem 4.2 as follows :

$$\left.\begin{aligned} A_n &= \lambda_1(n)P_1(n) + \lambda_2(n) P_2(n)+D(n), \\ R &= \lambda_1 P_1+\lambda_2 P_2+D, \\ C &= \rho_1 \widetilde{P}_1+\rho_2 \widetilde{P}_2+\widetilde{D} \end{aligned}\right\} \quad \ldots (74)$$

In (74) it is understood that $D(n)$, D and $\widetilde{D}$ are zero respectively if the corresponding two eigenvalues are distinct.

Lemma A.3 : *Let* A_n, R *and* C *be as in* (74) *satisfying*

(a) $A_n \to R$, *and*

(b) $(A_n)^n \to C$

as $n \to \infty$. *Then*

(i) $\lambda_1(n) \to \lambda_1$, $\lambda_2(n) \to \lambda_2$, $P_1(n) \to P_1$, $P_2(n) \to P_2$, $D(n) \to D$,

(ii) $(\lambda_1(n))^n \to \rho_1$, $(\lambda_2(n))^n \to \rho_2$, $\widetilde{P}_1 = P_1$, $\widetilde{P}_2 = P_2$,

(iii) $n[(\lambda_1(n))^{n-1} + (\lambda_2(n))^{n-1}]\, D(n) \to \widetilde{D}$.

The proof of this lemma is an easy consequence of Theorem A.1. Note that if $|\lambda_1| < 1$ and $|\lambda_2| < 1$ then $C = 0$. The next lemma rules out one of the various possibilities for the eigenvalues of C.

Lemma A.4 : *Assume* (a), (b) *of Lemma* A.3, *equation* (57) *and also* $\rho_1 = \rho_2 = 0$. *Then* $C = 0$.

Proof: As mentioned above if $|\lambda_i| < 1$, $i = 1, 2$, then $C = 0$. On the other hand when $\lambda_1(n) \neq \lambda_2(n)$ for sufficiently large n (which must happen if $\lambda_1 \neq \lambda_2$ and may happen even when $\lambda_1 = \lambda_2$), $D(n) = 0$. This means $\tilde{D} = C = 0$.

The only remaining case is when $|\lambda_i| = 1$, $i = 1, 2$ and $\lambda(n) \equiv \lambda_1(n) = \lambda_2(n)$ for sufficiently large n onwards. Writing $\lambda(n) = \lambda + \alpha_n$, using (a), (b) of Lemma A.3 and the result of Lemma A.2, we find that $R = \lambda + D$, $\alpha_n \to 0$, $n\,|\alpha_n| \to \infty$. Next by (57) we have

$$A_n = \lambda + \alpha_n + D(n) = \lambda + D + n^{-1}Q + o(n^{-1}) \qquad \ldots \quad (75)$$

or

$$(\alpha_n + D(n))^2 = \alpha_n^2 + 2\alpha_n D(n) = n^{-1}(DQ + QD) + o(n^{-1})$$

or

$$\alpha_n + 2D(n) = (n\alpha_n)^{-1}(DQ + QD) + o((n\,|\alpha_n|)^{-1}) \to 0$$

as $n \to \infty$ and we conclude $D = 0$. Substituting this in (75) above, we get $\alpha_n + D(n) = n^{-1}Q + o(n^{-1})$, and on multipllying both sides by $n(\lambda(n))^{n-1}$ we find

$$n\alpha_n(\lambda(n))^{n-1} + n(\lambda(n))^{n-1}D(n) = Q(\lambda(n))^{n-1} + o(1). \qquad \ldots \quad (76)$$

Note that $Q(\lambda(n))^{n-1} \to 0$ and $n(\lambda(n))^{n-1}D(n) \to \frac{1}{2}\tilde{D}$ by Lemma A.3(iii) as $n \to \infty$ proving the convergence of $n\alpha_n(\lambda(n))^{n-1}$ to β, say. Thus $\widetilde{D} + 2\beta = 0$ which implies $\tilde{D} = 0$. Hence $C = \tilde{D} = 0$. □

Proof of Theorem 4.2 : In view of Lemma A.4, we need to consider only two cases, case I : $\rho_i \neq 0$, $i = 1, 2$, and Case II : $\rho_1 \neq 0, \rho_2 = 0$.

Case I : Consider the subcase $\rho_1 \neq \rho_2$. Then it is essy to see that $\lambda_1(n) = 1+n^{-1}\alpha_1+o(n^{-1})$ and $\lambda_2(n) = 1+n^{-1}\alpha_2+o(n^{-1})$ where $\alpha_i = \log\ \rho_i$, $i = 1, 2$. Since $\lambda_1(n) \neq \lambda_2(n)$. we. have that $D(n) = 0 = D$ and $R = I$, and $C = \rho_1 P_1 + \rho_2 P_2$. Now if instead $\rho_1 = \rho_2 \equiv \rho \neq 0$, we see that if also $\lambda_1(n) = \lambda_2(n) \equiv \lambda(n)$, then $\lambda(n) = 1+n^{-1}\alpha+o(n^{-1})$ with $\alpha = \log\ \rho$. On the other hand from Lemma A.3(iii), $2n(\lambda(n))^{n-1}D(n) \to \tilde{D}$ leading to $nD(n) \to (2\rho)^{-1}\ \tilde{D}$ or $D = 0$ which implies that $R = I$. If $\lambda_1(n) \neq \lambda_2(n)$ for all n, then $D(n) = D = 0$ so that $R = I$ again,.

Case II : In this case $\lambda_1(n) = 1+n^{-1}\alpha_1+o(n^{-1})$ and $\lambda_2(n) = \lambda_2+\alpha_n$ with $\alpha_1 = \log\ \rho_1$, $|\lambda_2| \leqslant 1$. Since by Lemma A.2, $n|\alpha_n| \to \infty$, $\lambda_1(n) \neq \lambda_2(n)$ for sufficiently large n onwards, and hence $D(n) = 0$. Thus $D = \tilde{D} = 0$ and $R = P_1+\lambda_2 P_2$. If moreover $|\lambda_2| < 1$, then $R^n \to P_1$. On the other hand if $\lambda_2 = 1$, then $R = I$.

Finally let $|\lambda_2| = 1, \lambda_2 \neq 1$ and $P_2 \neq 0$. In such a case we show that (57) is violated. From (57) we have that

$$(\lambda_2-1)\,[P_2(n)-P_2]+\alpha_n P_2(n) = n^{-1}\,[Q-\alpha_1 P_1(n)]+o(n^{-1}). \tag{77}$$

Using an orthonormal basis $\{e_i\}_{i=1,2}$ of $\boldsymbol{R}^2$ such that e_i is an eigenvector of P_i with eigenvalue 1, we find that

$$(\lambda_2-1+\alpha_n)\,(e_1, P_2(n)e_2) = n^{-1}(e_1, [Q-\alpha_1 P_1(n)]e_2)+o(n^{-1}),$$

leading to $(e_1, P_2(n)e_2 = O(n^{-1})$. Next the simple identity :

$$(e_2, [P_2-P_2(n)]e_2) = \frac{|(e_1, P_2(n)e_2)|^2}{(e_2, P_2(n)e_2)} \quad \ldots \quad (78)$$

tells us that

$$(e_2, [P_2-P_2(n)]e_2) = O(n^{-2}), \quad \ldots \quad (79)$$

since $(e_2, P_2(n)e_2) \to 1$. On the other hand from (77) one has

$$(\lambda_2-1)n(e_2,\ [P_2(n)-P_2]e_2$$
$$=\{(e_2, Qe_2)-\alpha_1(e_2,\ P_1(n)e_2)\}-n\alpha_n(e_2,\ P_2(n)e_2)+o(1). \quad \ldots \quad (80)$$

Since $n|\alpha_n| \to \infty$, clearly (79) contradicts (80). Thus $P_2 = 0$ which implies $R = P_1 = I$. $\square$

Acknowledgement. This research work was carried out jointly at the Indian Statistical Institute, New Delhi, India, where the first two authors were visiting (Prem S. Puri for the years 1986-88 and James B. Robertson for 1986-87). These authors are grateful to the Institute for providing facilities and a stimulating atmosphere for the conduct of their research.

References

Kato, T. (1982) : *A Short Introduction to Perturbation Theory for Linear Operators*, New York : Springer Verlag.

Robertson, J. B. (1973) : A spectral representation of the states of a measure preserving transformation. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, **27**, 185-194.

Paper received : *November*, 1988.

Revised : *February*, 1989.