

SOME CONTRIBUTIONS TO THE THEORY OF DENUMERABLE MARKOV CHAINS⁽¹⁾

BY
CYRUS DERMAN

Summary. §1 deals with the statistical regularity properties of a denumerable number of particles, all moving about the states of a Markov chain according to the same transition probabilities.

§2 deals with the problem of obtaining a sharper version of a strong limit theorem proved independently by Harris and Lévy.

Introduction. We shall be concerned throughout with a sequence of random variables $\{X_n\}$, $n=0, 1, \dots$, which assume only integer values⁽²⁾ and which have the property that

$$\begin{aligned} \Pr \{X_{m+n} = j \mid X_m = i, X_{m-1}, \dots, X_0\} \\ = \Pr \{X_{m+n} = j \mid X_m = i\} = \Pr \{X_n = j \mid X_0 = i\} \end{aligned}$$

for all integers $m, n \geq 0$ and all states i and j for which the conditional probability is defined. The distribution of X_0 will be fixed but arbitrary. Sequences of this type are known as Markov chains with denumerable states and stationary transition probabilities. The fundamentals of the theory of such chains were laid down by Kolmogorov [14]. Chung [2] and Feller [10] have also given expositions of the main results of the theory. In the remainder of the introduction we shall state, without proof, those results from the general theory which we shall need later on. Also, as much as possible, we shall assign the notation which will be used throughout the remainder of the paper.

The probability given above is called the n th step transition probability which we shall denote by $p_{ij}^{(n)}$. If given a matrix

$$\begin{aligned} P = \{p_{ij}^{(1)} = p_{ij}\}, \quad p_{ij} \geq 0, i, j = 0, 1, \dots \text{ } ^{(3)}; \\ \sum_{j=0}^{\infty} p_{ij} = 1, \quad i = 0, 1, \dots, \end{aligned}$$

Received by the editors November 4, 1954.

⁽¹⁾ This work was sponsored by the United States Air Force through the Office of Scientific Research of the Air Research and Development Command and is part of a dissertation submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy at Columbia University.

⁽²⁾ Each integer denotes a possible state of the Markov chain.

⁽³⁾ We could just as well have included the negative integers in representing the totality of states. It is clearly only a matter of notation.

we can always construct a sequence of random variables $\{X_n\}$, $n=0, 1, \dots$, which is a Markov chain having one-step transition probabilities P . The n th step transition probabilities are then defined recursively by the relation

$$p_{ij}^{(n)} = \sum_{k=0}^{\infty} p_{ik}^{(n-1)} p_{kj}, \quad i, j, = 0, 1, \dots$$

In what follows we shall sometimes refer to the Markov chain as being P and sometimes as $\{X_n\}$.

An important notion in Markov chain theory is that of the number of transitions necessary to reach a certain state for the first time given initially a particular state. We let

$$f_{ij}^{(n)} = \Pr \{ X_{m+n} = j, \quad X_{m+\nu} \neq j \text{ for } 1 \leq \nu < n \mid X_m = i \},$$

$$m, i, j = 0, 1, \dots; n = 1, 2, \dots;$$

$$f_{ij}^* = \sum_{\nu=1}^{\infty} f_{ij}^{(\nu)}; \quad m_{ij} = \sum_{\nu=1}^{\infty} \nu f_{ij}^{(\nu)}; \quad \sigma_{ij}^2 = \sum_{\nu=1}^{\infty} (\nu - m_{ij})^2 f_{ij}^{(\nu)}.$$

We call m_{ij} the mean first passage time from state i to state j . If $i=j$, m_{ii} is called the mean recurrence time of state i . States i and j are said to belong to the same class if $f_{ij}^* f_{ji}^* > 0$, i.e., if there exist integers $n(i, j)$ and $n(j, i)$ such that $p_{ij}^{(n(i,j))} > 0$ and $p_{ji}^{(n(j,i))} > 0$. A state i is called recurrent if $f_{ii}^* = 1$. This implies that $f_{ii}^* = 1$ for all j belonging to the same class as i . If $f_{ii}^* < 1$, i is called transient. If i is recurrent and j belongs to the same class as i , then j is also recurrent. Consequently, all states of a class are either all recurrent or all transient. We can then speak of a class as being recurrent or transient. If all states belong to the same class we refer to the Markov chain as being irreducible recurrent (transient). Even though it might not always be explicitly stated, we shall deal throughout only with irreducible chains. If $f_{ii}^* = 1$ and $m_{ii} < \infty$, i is said to be a positive state. If $m_{ii} = \infty$, i is called null. If i is positive (null) and j belongs to the same class, then j is positive (null). Thus all states of a recurrent class will be positive or null together. A recurrent class is called positive or null accordingly.

If $f_{ii}^* = 1$, $m_{ii} < \infty$, then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\nu=1}^n p_{ji}^{(\nu)} = \frac{1}{m_{ii}}$$

for all states j belonging to the same class as i . If $m_{ii} = \infty$ then $\lim_{n \rightarrow \infty} p_{ji}^{(n)} = 0$. If all states belong to the same positive class, then

$$\sum_{i=0}^{\infty} \frac{1}{m_{ii}} = 1 \quad \text{and} \quad \sum_{i=0}^{\infty} \frac{1}{m_{ii}} p_{ij} = \frac{1}{m_{jj}}, \quad j = 0, 1, \dots$$

In §2 we shall consider the number of times a state is visited in n transitions. This number we shall denote by $N_n(\cdot)$. That is,

$$N_n(i) = \{ \text{the number of } \nu\text{'s such that } X_\nu = i \text{ for } 1 \leq \nu < n \}.$$

1. Doob [8, p. 404] considered the macroscopic regularity of a denumerable number of particles diffusing on the real axis according to some stochastic process. In particular, he gave some sufficient conditions on the nature of the stochastic process such that a certain kind of statistical equilibrium would be maintained. We consider here a denumerable number of particles, independently moving about the states of a Markov chain, all according to the same transition probabilities. We shall be interested in seeing what types of chains maintain statistical equilibrium and in determining the nature of the equilibrium⁽⁴⁾.

An intimately related problem is that of finding positive numbers $\{v_i\}$ satisfying the system of equations

$$(1) \quad \sum_{i=0}^{\infty} v_i p_{ij} = v_j, \quad j = 0, 1, \dots$$

When the Markov chain is irreducible, recurrent, and positive, there exists such a set $\{v_i\}$, having the properties that $0 < v_i < 1$, $i = 0, 1, \dots$, and $\sum_{i=0}^{\infty} v_i = 1$. This set is unique. In fact, $v_i = \lim_{n \rightarrow \infty} (1/n) \sum_{\nu=0}^{\infty} p_{ii}^{(\nu)}$, $i = 0, 1, \dots$. However, for the null and transient cases this is not true. But, if the restriction that the v_i 's be probabilities is removed, the question of existence and uniqueness of such solutions is open. The significance in other connections of the existence of such a set of numbers was pointed out by Harris and Robbins [12]. They were concerned with the application of an ergodic theorem of Hopf [13] to prove some strong laws concerning the ratio of random variables.

Doebelin [8] first proved that

$$(2) \quad \lim_{n \rightarrow \infty} \frac{\sum_{\nu=0}^n p_{ii}^{(\nu)}}{\sum_{\nu=0}^n p_{jj}^{(\nu)}} = \pi_{ij}, \quad i, j = 0, 1, \dots,$$

exists and is both positive and finite for all types of Markov chains. Chung [2] also proved this result and gave some formulas which express (2) in useful terms. Utilizing the results of Doebelin and Chung, the following theorem was proved by Derman [5]:

THEOREM 1. *If all the states of an irreducible Markov chain are recurrent,*

(4) This line of research was suggested by T. E. Harris.

then there exists one and only one set of positive⁽⁵⁾ numbers $\{v_k\}$, $v_0=1$, $k=0, 1, \dots$, satisfying (1) and, in fact, $v_k=\pi_{k0}$, $k=0, 1, \dots$.

Theorem 1 answers the question of existence and uniqueness of numbers $\{v_k\}$ satisfying (1) for the recurrent case. We shall now give some examples to show that the situation is not as simple for the case where the states are transient.

EXAMPLE 1. *A renewal type chain.* Let $p_{i,i+1}=p_i$ and $p_{i,0}=1-p_i$, $i=0, 1, \dots$. If $\{v_k\}$, $v_0=1$, $k=0, 1, \dots$, are to be solutions of (1) we must have

$$(3) \quad v_k = \prod_{\nu=0}^{k-1} p_\nu, \quad k = 1, 2, \dots,$$

$$v_0 = 1 = \sum_{i=0}^{\infty} v_i(1-p_i) = \lim_{n \rightarrow \infty} \left(1 - \prod_{\nu=0}^n p_\nu\right).$$

Thus there is a unique solution to (1) if and only if $\lim_{n \rightarrow \infty} \prod_{\nu=0}^n p_\nu = 0$. However, $\prod_{\nu=0}^n p_\nu$ is the probability of no return to state zero in $(n+1)$ transitions. Hence the above condition is that the probability is one that a return to state zero occur in a finite number of steps, i.e., state zero is recurrent. Since if one state is recurrent the chain is recurrent, we have that a solution to (1) exists if and only if the chain is recurrent.

EXAMPLE 2. *The asymmetrical unrestricted random walk.* Let $p_{i,i+1}=p \neq 1/2$, $p_{i,i-1}=q=1-p$, $i=\dots, -1, 0, 1, \dots$. When $p=q=1/2$ it is known that the Markov chain so defined is recurrent null. When $p \neq q$ it is also known that the chain is transient.

Equations (1) in this case become

$$(4) \quad qv_{k+1} - v_k + v_{k-1}p = 0, \quad k = \dots, -1, 0, 1, \dots$$

The general solution to this difference equation is

$$(5) \quad v_k = c_1(p/q)^k + c_2, \quad k = \dots, -1, 0, 1, \dots; p \neq q$$

where c_1 and c_2 are arbitrary constants. The conditions $v_0=1$ and $v_k > 0$, $k=\dots, -1, 0, 1, \dots$, imply $c_1 \geq 0$, $c_2 \geq 0$, and $c_1+c_2=1$. Thus for any value of c_1 such that $0 \leq c_1 \leq 1$, (5) is a solution to (1); i.e., for the asymmetrical unrestricted random walk, solutions to (1) exist but are not unique.

When $p=q=1/2$, $v_k=c_1k+c_2$ satisfies (4). If $c_1 \neq 0$ then v_k will not always be non-negative. This shows, since the chain is recurrent null, that the condition of non-negativeness is essential for the uniqueness in Theorem 1.

EXAMPLE 3. Let

$$p_{00} = 1/2, p_{01} = 1/2, p_{k,k-1} = 1/2^{k+2} = p_{kk}, p_{k,k+1} = (2^{k+1} - 1)/2^{k+1},$$

$$k = 1, 2, \dots$$

⁽⁵⁾ This positiveness condition can be weakened to non-negativeness. This can be seen by noting that it is impossible for a sequence $\{v_r\}$ to satisfy (1) if one or more but not all v_k 's are zero.

We know that the probability of a direct drift to ∞ starting from 0 is $(1/2) \prod_{k=1}^{\infty} (2^{k+1}-1)/2^{k+1} > 0$. Hence the probability of not returning to 0 in a finite number of steps is positive and, therefore, the chain is transient.

Equations (1) become

$$\begin{aligned} (1/2)v_0 + (1/8)v_1 &= v_0, \\ (6) \quad (1/2)v_0 + (1/8)v_1 + (1/16)v_2 &= v_1, \\ p_{k-2,k-1}v_{k-2} + p_{k-1,k-1}v_{k-1} + p_{k,k-1}v_k &= v_{k-1}, \quad k = 2, 3, \dots \end{aligned}$$

The condition $v_0 = 1$ implies $v_1 = 4$. The v_k 's for $k \geq 2$ are uniquely determined since the equations (6) can be solved one at a time for each successive v_k . It remains to be shown that all v_k 's are positive. To do this we use induction. We have $v_2 > v_1 > 0$. Assume that $v_m > v_{m-1} > 0$ for all integers $m < k$. From (6) we get

$$\begin{aligned} (7) \quad p_{k,k-1}v_k &= v_{k-1}(1 - p_{k-1,k-1}) - v_{k-2}p_{k-2,k-1} \\ &= v_{k-1}(1 - 1/2^{k+1}) - v_{k-2}(1 - 1/2^{k-1}). \end{aligned}$$

Using the induction assumption we have $P_{k,k-1}v_k \geq v_{k-1}(3/2^{k+1})$. Hence $v_k \geq v_{k-1}(3/p_{k,k-1}2^{k+1}) = 6v_{k-1}$; i.e., the relationship holds for all integers. Thus we have an example of a transient chain having a unique solution to (1).

For the last example, we give an example of a recurrent chain having the property that the set $\{v_k\}$ need not be bounded.

EXAMPLE 4. Let $p_{0i} = p_i, i = 0, 1, \dots$; $p_{ii} = f(i), p_{i0} = 1 - f(i), i = 1, 2, \dots$, where $0 < f(i) < 1$ for all i . Clearly the chain is recurrent.

Equations (1) become

$$(8) \quad p_k + v_k f(k) = v_k, \quad k = 1, 2, \dots,$$

hence $v_k = p_k / (1 - f(k)), k = 1, 2, \dots$. Let $g(k)$ be any increasing function. If we choose $f(k) = 1 - p_k / g(k)$, we have $v_k = g(k)$. Thus v_k can be made to increase as fast as desired by choosing $g(k)$ appropriately.

We now consider the denumerable number of particles moving about the states of a chain. More precisely, let $\{X_n(k)\}, k = 1, 2, \dots$, be Markov chains, each having the same matrix of transition probabilities $P = \{p_{ij}\}$ independent of the states of the other chains. We assume P to be irreducible and recurrent. Let $A_i(n)$ be the number of k 's, $k = 1, 2, \dots$, such that $X_n(k) = i, i = 0, 1, \dots$, i.e., $A_i(n)$ denotes the number of particles in state i at time n . The sequence $\{v_k\}$ will denote, throughout, the solutions to (1). We prove the following theorem which establishes the existence of a statistical equilibrium.

THEOREM 2. If $\{A_i(0)\}, i = 0, 1, \dots$, are independently distributed, each having a Poisson distribution with mean $v_i, i = 0, 1, \dots$, then for every $n \geq 0, \{A_i(n)\}, i = 0, 1, \dots$, are independently distributed each having a Poisson distribution with mean $v_i, i = 0, 1, \dots$.

Proof. Consider any subset i_1, \dots, i_r of states and the random variables $A_{i_1}(n), \dots, A_{i_r}(n)$ for any $n \geq 1$. The characteristic function of $A_{i_1}(n), \dots, A_{i_r}(n)$ is

$$\begin{aligned}
 \phi_n(t_1, \dots, t_r) &= \prod_{k=0}^{\infty} \sum_{x=0}^{\infty} (1 + p_{k i_1}^{(n)}(e^{i t_1} - 1) + \dots + p_{k i_r}^{(n)}(e^{i t_r} - 1))^x \frac{v e^{-v} v_k^x}{x!} \\
 &= \prod_{k=0}^{\infty} e^{-v k} \sum_{x=0}^{\infty} \frac{1}{x!} (v_k(1 + p_{k i_1}^{(n)}(e^{i t_1} - 1) + \dots \\
 &\quad + p_{k i_r}^{(n)}(e^{i t_r} - 1))^x \\
 &= \prod_{k=0}^{\infty} e^{-v k + v k} (1 + p_{k i_1}^{(n)}(e^{i t_1} - 1) + \dots + p_{k i_r}^{(n)}(e^{i t_r} - 1)) \\
 (9) \quad &= \prod_{\alpha=1}^r \exp((e^{i t_\alpha} - 1) \sum_{k=0}^{\infty} v_k p_{k i_\alpha}^{(n)}) \\
 &= \prod_{\alpha=1}^r \exp(v_{i_\alpha}(e^{i t_\alpha} - 1)).
 \end{aligned}$$

The last equality in (9) follows by virtue of Theorem 1. Now, the right side of (9) is the characteristic function of r independent Poisson distributions having means v_{i_1}, \dots, v_{i_r} . Since this is true for all subsets i_1, \dots, i_r , the theorem is proved.

The above result is similar to that of Doob. He considered a statistical equilibrium of the following kind: Let I_1, \dots, I_r be any r nonoverlapping intervals. The random variables Z_1, \dots, Z_r denoting the number of particles in the intervals I_1, \dots, I_r respectively, are independently distributed according to Poisson distributions having means proportional to the length of the interval. A particular stochastic process for which this kind of equilibrium exists is the Wiener process.

Let \mathcal{A} be the space of all sequences of non-negative integers $a = \{a_0, a_1, \dots\}$. Then $A_n = \{A_0(n), A_1(n), \dots\}$, $n = 0, 1, \dots$, is a sequence of random variables taking on the values in \mathcal{A} . For any cylinder set $A \in \mathcal{A}$ let $\Phi(A)$ be the probability measure of A defined by independent Poisson distributions having means v_i , $i = 0, 1, \dots$, where v_i is the mean associated with the i th component of A_n . By the Kolmogorov extension theorem Φ can be extended to a probability measure defined over the smallest Borel field \mathcal{A}_F containing all the cylinder sets. Again, for any cylinder set $A \in \mathcal{A}$ the n th-step transition probability

$$\Pr \{A_{k+n} \in A \mid A_k = a\} = P^{(n)}(a, A), \quad n = 1, \dots; k \geq 0,$$

is determined by its characteristic function

$$\phi_{n,a}(t_1, \dots, t_s) = \prod_{k=0}^{\infty} (1 + p_{k i_1}^{(n)}(e^{i t_1} - 1) + \dots + p_{k i_s}^{(n)}(e^{i t_s} - 1))^{a_k}$$

where i_1, \dots, i_s are the coordinates involved in the cylinder set A . Using the Kolmogorov extension theorem again we have that $P^{(n)}(a, A)$, $n=1, 2, \dots$, for any $A \in \mathcal{A}_F$ is uniquely defined. Thus $\{A_n\}$ is a Markov chain having the stationary probability measure Φ and transition probabilities $P^{(1)}(a, A)$. It should be noted that $P^{(n)}(a, A)$ is a genuine probability measure for only those values of a such that $\sum_{i=0}^{\infty} a_i p_{ik}^{(n)} < \infty$ for all k . However, it will be shown in Theorem 3(ii) below that the Φ -measure of such a set is one. Strictly speaking, in defining the above Markov chain, we restrict ourselves to those a which belong to this set. We get the following

THEOREM 3. (i) *If $A = \{a \mid \sum_{i=0}^{\infty} a_i < \infty\}$ then $\Phi(A) = 0$ or 1 according as $\{p_{ij}\}$ is null or positive.*

(ii) *If $A = \{a \mid \sum_{i=0}^{\infty} a_i p_{ij}^{(n)} < \infty$ for all n and $j\}$ then $\Phi(A) = 1$.*

Proof. (i) For each n , $\sum_{i=0}^{\infty} A_i(n)$ is a series whose terms are independent Poisson random variables having as means (variances) v_0, v_1, \dots . If $\{p_{ij}\}$ is positive, $\sum_{i=0}^{\infty} v_i < \infty$. It follows then that $\sum_{i=0}^{\infty} A_i(n) < \infty$ with probability (Φ -measure) one (see [8, p. 108]). If $\{p_{ij}\}$ is null, $\sum_{i=0}^{\infty} v_i = \infty$. The probability that $A_i(n) = 0$ for $i \geq k$ is $\exp(-\sum_{i=k}^{\infty} v_i) = 0$ for $k=0, 1, \dots$; $n=0, 1, \dots$. But

$$\Pr \left\{ \sum_{i=0}^{\infty} A_i(n) < \infty \right\} \leq \sum_{k=0}^{\infty} \Pr \{ A_i(n) = 0 \text{ for } i \geq k \} = 0.$$

(ii) On taking the expectation with respect to the Φ -probability measure we have since the a_i 's are non-negative

$$(10) \quad E \left(\sum_{i=0}^{\infty} a_i p_{ik}^{(n)} \right) = \sum_{i=0}^{\infty} v_i p_{ik}^{(n)} = v_k < \infty.$$

The assertion is an immediate consequence.

We may note that as a consequence of the stationarity of the Markov chain $\{A_n\}$, the ergodic theorem of Birkhoff is applicable in proving certain strong laws concerning functions of the sequence of random variables $\{A_n\}$. However, the following theorem is also implied:

THEOREM 4. (i) *Except for a set of Φ -measure zero, and for all $\mathcal{E} \in \mathcal{A}_F$, there exists a function $H(a, \mathcal{E})$ such that*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\nu=1}^n P^{(\nu)}(a, \mathcal{E}) = H(a, \mathcal{E}).$$

(ii) *Except for a set of Φ -measure zero, there exists a function $K_j(a)$ for every j such that*

$$\lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} a_k \frac{1}{n} \sum_{\nu=1}^n p_{kj}^{(\nu)} = K_j(a).$$

(iii) If $H(a, \mathcal{E})$ is independent of a except for a set of Φ -measure zero, then $H(a, \mathcal{E}) = \Phi(\mathcal{E})$.

Proof. (i) and (ii) follow immediately from a theorem of Doob [7, p. 400]. The crucial assumption in applying Doob's theorem is the existence of a stationary probability measure Φ .

(iii) We have from the stationarity property

$$(11) \quad \frac{1}{n} \int_{\mathcal{A}} \sum_{\nu=1}^n P^{(\nu)}(a, \mathcal{E}) \Phi(da) = \Phi(\mathcal{E}) \quad \text{for all } n.$$

But since $(1/n) \sum_{\nu=1}^n P^{(\nu)}(a, \mathcal{E})$ is bounded and $\Phi(\mathcal{A}) = 1$

$$(12) \quad \begin{aligned} \int_{\mathcal{A}} H(a, \mathcal{E}) \Phi(da) &= H(\mathcal{E}) \int_{\mathcal{A}} \Phi(da) = H(\mathcal{E}) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \int_{\mathcal{A}} \sum_{\nu=1}^n P^{(\nu)}(a, \mathcal{E}) \Phi(da) = \Phi(\mathcal{E}). \end{aligned}$$

We now consider the limit in (ii) of Theorem 4 for the case where $\{p_{ij}\}$ is null. Let

$$(13) \quad f_j(k, N) = \sup_{n > N} \frac{1}{n} \sum_{\nu=1}^n p_{kj}^{(\nu)}, \quad j, k, N = 0, 1, \dots$$

THEOREM 5. Let $\{p_{ij}\}$ be null and C_j any set of numbers.

(i) If $A = \{a \mid K_j(a) = C_j\}$ then $\Phi(A) = 0$ or 1.

(ii)⁽⁶⁾ If $C_j = v_j$, $j = 0, 1, \dots$, and if for some N , $\sum_{k=0}^{\infty} v_k^{1/2} f_j(k, N) < \infty$, then $\Phi(A) = 1$.

Proof. (i) For any integer m , let a be any point of \mathcal{A} differing from a' only at the first m coordinates where a' is any point in A . We then have

$$(14) \quad \sum_{k=0}^{\infty} a_k \frac{1}{n} \sum_{\nu=1}^n p_{kj}^{(\nu)} = \sum_{k=0}^m (a_k - a'_k) \frac{1}{n} \sum_{\nu=1}^n p_{kj}^{(\nu)} + \sum_{k=0}^{\infty} a'_k \frac{1}{n} \sum_{\nu=1}^n p_{kj}^{(\nu)}.$$

But since $\{p_{ij}\}$ is null, $\lim_{n \rightarrow \infty} (1/n) \sum_{\nu=1}^n p_{kj}^{(\nu)} = 0$ for all k . Thus on letting $n \rightarrow \infty$, (14) tends to C_j and hence $a \in A$. This shows that the conditions determining the set A are independent of the first m coordinates for $m = 0, 1, \dots$. Therefore, by the zero-one-law, $\Phi(A) = 0$ or 1.

(ii) If for some N , $\sum_{k=0}^{\infty} |a_k - v_k| f_j(k, N) < \infty$ then

⁽⁶⁾ It has been pointed out by the referee that it is sufficient to assume $\sum_k v_k (f_j(k, N))^2 < \infty$. Under this condition the left side of (14) converges in mean-square to v_j . Using this fact, together with Theorem 4 (ii), his assertion follows.

$$\lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} (a_k - v_k) (1/n) \sum_{\nu=1}^n p_{kj}^{(\nu)} = 0.$$

Let B be the set of all points $a \in A$ such that $\sum_{k=0}^{\infty} a_k p_{kj}^{(n)} < \infty$ for all n and such that the above series converges. For any point $a \in B$

$$(15) \quad \sum_{k=0}^{\infty} a_k \frac{1}{n} \sum_{\nu=1}^n p_{kj}^{(\nu)} = \sum_{k=0}^{\infty} (a_k - v_k) \frac{1}{n} \sum_{\nu=1}^n p_{kj}^{(\nu)} + \sum_{k=0}^{\infty} v_k \frac{1}{n} \sum_{\nu=1}^n p_{kj}^{(\nu)}.$$

On letting $n \rightarrow \infty$, (15) tends to v_j . Now since $A_k(0)$ has variance v_k , $k=0, 1, \dots$, $E|A_k(0) - v_k| \leq v_k^{1/2}$, $k=0, 1, \dots$. Thus by a theorem (see [8, p. 108]) on the convergence of a series of random variables if $\sum_{k=0}^{\infty} v_k^{1/2} f_j(k, N) < \infty$ and $\sum_{k=0}^{\infty} v_k f_j^2(k, N) < \infty$ (but here the convergence of the first series implies the convergence of the second series) then with probability one $\sum_{k=0}^{\infty} |A_k(0) - v_k| f_j(k, N)$ converges. Since the set $\mathcal{E} = \{a \mid \sum_{i=0}^{\infty} a_i p_{ij}^{(n)} = \infty \text{ for some } n\}$ has measure zero by Theorem 2, it follows that $\Phi(B) = 1$.

If $\{p_{ij}\}$ is positive, using (i) of Theorem 3 it is easy to see that if $A = \{a \mid \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} a_k \frac{1}{n} \sum_{\nu=1}^n p_{kj}^{(\nu)} = p_j, \sum_{k=0}^{\infty} a_k < \infty\}$ then $\Phi(A) = 1$ where $p_j = \lim_{n \rightarrow \infty} (1/n) \sum_{\nu=1}^n p_{kj}^{(\nu)}$, $k=0, 1, \dots$.

We now investigate the behavior of $\{A_i(n)\}$, $i=0, 1, \dots$, for large n given $A_0 = a$. Analogous to (13) we define

$$(16) \quad g_j(k, N) = \sup_{n > N} p_{kj}^{(n)}, \quad k, j, N = 0, 1, \dots$$

Also let

$$(17) \quad h_j(n) = \sup_{k \geq 0} p_{kj}^{(n)}, \quad j = 0, 1, \dots; n = 1, 2, \dots$$

We shall refer to the following conditions:

$$(18) \quad \sum_{k=0}^{\infty} |a_k - v_k| g_j(k, N) < \infty \quad \text{for some } N,$$

$$(19) \quad \lim_{n \rightarrow \infty} h_j(n) = 0.$$

We have now

THEOREM 6. *If (19) holds for a set of states $j=i_1, \dots, i_r$, and A is a set of points in \mathcal{A} such that (18) holds for $j=i_1, \dots, i_r$, then*

$$\lim_{n \rightarrow \infty} \Pr \{A_{i_1}(n) = x_1, \dots, A_{i_r}(n) = x_r \mid A_0 = a\} = \prod_{\alpha=1}^r \frac{e^{-v_{i_\alpha}} v_{i_\alpha}^{x_\alpha}}{x_\alpha!}$$

for all $a \in A$, where x_1, \dots, x_r is any set of non-negative integers.

Proof. The characteristic function of the distribution of $A_{i_1}^{(n)}, \dots, A_{i_r}^{(n)}$ given $A_0 = a$ is

$$(20) \quad \phi_{n,a}(t_1, \dots, t_r) = \prod_{k=0}^{\infty} (1 + p_{ki_1}^{(n)}(e^{it_1} - 1) + \dots + p_{ki_r}^{(n)}(e^{it_r} - 1))^{a_k}.$$

On taking logarithms we get

$$(21) \quad \begin{aligned} \log \phi_{n,a} &= \sum_{k=0}^{\infty} a_k \log (1 + \psi_{n,k}(t_1, \dots, t_r)) \\ &= \sum_{k=0}^{\infty} a_k (\psi_{n,k}(t_1, \dots, t_r) + \theta |\psi_{n,k}|^2) \end{aligned}$$

where $|\theta| < 1$ and $\psi_{n,k} = p_{ki_1}^{(n)}(e^{it_1} - 1) + \dots + p_{ki_r}^{(n)}(e^{it_r} - 1)$. Since $\{p_{ij}\}$ is null we can assume n large enough, i.e., $p_{ki_\alpha}^{(n)}$, $\alpha = 1, \dots, r$, small enough, so that the expansion of $\log(1 + \psi_{n,k})$ in (21) is valid. By Theorem 1 and assumption (18) for $j = i_1, \dots, i_r$ and the fact that $\psi_{n,k} \rightarrow 0$ uniformly as $n \rightarrow \infty$, we have

$$(22) \quad \begin{aligned} \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} a_k \psi_{n,k} &= \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} (a_k - v_k) \psi_{n,k} + \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} v_k \psi_{n,k} \\ &= v_{i_1}(e^{it_1} - 1) + \dots + v_{i_r}(e^{it_r} - 1). \end{aligned}$$

Also, by assumption (18) and (19)

$$(23) \quad \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} a_k |\psi_{n,k}|^2 = \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} (a_k - v_k) |\psi_{n,k}|^2 + \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} v_k |\psi_{n,k}|^2 = 0.$$

Therefore, we have

$$(24) \quad \lim_{n \rightarrow \infty} \log \phi_{n,a}(t_1, \dots, t_r) = v_{i_1}(e^{it_1} - 1) + \dots + v_{i_r}(e^{it_r} - 1)$$

and hence $\lim_{n \rightarrow \infty} \phi_{n,a}(t_1, \dots, t_r) = \prod_{\alpha=1}^r \exp v_{i_\alpha}(e^{it_\alpha} - 1)$.

We get the following

COROLLARY 1⁽⁷⁾. *If for some N , $\sum_{k=0}^{\infty} v_k^{1/2} g_j(k, N) < \infty$, $j = i_1, \dots, i_r$, then the set A of Theorem 6 has Φ -measure 1.*

Proof. The proof follows using the same type of argument as used to prove (ii) of Theorem 5.

Suppose, instead of A_0 having the probability measure Φ , it has probability measure Φ^* generated, also, by a sequence of independent random variables. Then we get

(7) It seems that the referee's strengthening of Theorem 5 (ii) should carry over here. However, in this case, there is no analogue to the probability one limit of Theorem 4(ii) which he used in his proof.

COROLLARY 2. *If $E|A_k(0) - v_k| = \alpha_k$, $E|A_k(0) - v_k|^2 - \alpha_k^2 = \sigma_k^2$ and if for some N , $\sum_{k=0}^{\infty} \alpha_k g_j(k, N) < \infty$, $j = i_1, \dots, i_r$, and $\sum_{k=0}^{\infty} \sigma_k^2 g_j^2(k, N) < \infty$, $j = i_1, \dots, i_r$, then the set A of Theorem 6 has Φ^* -measure 1.*

Proof. The same remark used for the proof of Corollary 1 holds again.

We remark that the conditions of Corollary 1 for $j=0, 1, \dots$ imply metric transitivity of the Markov chain $\{A_n\}$ (see [7, p. 401]).

As an application of Theorem 6 and Corollary 2 let us consider the symmetrical unrestricted random walk. Here we have $v_k \equiv 1$, $k = \dots, -1, 0, 1, \dots$ and

$$p_{kj}^{(n)} = \binom{n}{(n - |k-j|)/2} \frac{1}{2^n}$$

if $n \geq |k-j|$ and $n + |k-j|$ is even, $p_{kj}^{(n)} = 0$ otherwise. For all k, j and $n \geq |k-j|$ we have

$$(25) \quad p_{kj}^{(n)} \leq \binom{n}{(n - \Delta_1)/2} \frac{1}{2^n} \leq \binom{|k-j|}{(|k-j| - \Delta_2)/2} \frac{1}{2^{|k-j|}}$$

where $\Delta_1 = 0$ if n is even, $\Delta_1 = 1$ if n is odd and $\Delta_2 = 0$ if $|k-j|$ is even, $\Delta_2 = 1$ if $|k-j|$ is odd. Hence we have

$$g_j(k, 1) \leq \binom{|k-j|}{(|k-j| - \Delta_2)/2} \frac{1}{2^{|k-j|}} \quad \text{and} \quad h_j(n) \leq \binom{n}{(n - \Delta_2)/2} \frac{1}{2^n}.$$

It is clear that $h_j(n) \rightarrow 0$ as $n \rightarrow \infty$. Also, it is known that $g_j(k, 1) \leq O(|k-j|^{-1/2})$. Thus by Theorem 6, for any set i_1, \dots, i_r and for all a such that $\sum_{k=-\infty}^{\infty} |a_k - 1| (1/|k|)^{1/2}$ we have

$$\lim_{n \rightarrow \infty} \Pr \{A_{i_1}(n) = x_1, \dots, A_{i_r}(n) = x_r \mid A_0 = a\} = \prod_{\alpha=1}^r \frac{e^{-1}}{x_\alpha!}.$$

By Corollary 2 to Theorem 6, if the probability measure Φ^* associated with A_0 is such that $\sum_{k=-\infty}^{\infty} E|A_{|k|} - 1| |k|^{-1/2} < \infty$ and $\sum_{k=-\infty}^{\infty} E(A_{|k|} - 1)^2 |k|^{-1} < \infty$, then

$$\lim_{n \rightarrow \infty} \Pr \{A_{i_1}(n) = x_1, \dots, A_{i_r}(n) = x_r\} = \prod_{\alpha=1}^r \frac{e^{-1}}{x_\alpha!}.$$

We note, in closing this section, that most of the theorems proved here rely heavily only on the existence of solutions to (1). Thus, transient chains, for which solutions to (1) exist, have the stationarity property of Theorem 2. The interesting feature, here, is that in such cases the drift is of such a nature that although more than one distribution could serve as the stationary distribution (for those cases where the solutions to (1) are not unique) the initial distribution is maintained.

2. In [11] and [15] Harris and Lévy each prove the following theorem:

THEOREM 1. *If i and j are any two states belonging to the same recurrent class, then with probability one*

$$\lim_{n \rightarrow \infty} \frac{N_n(j)}{N_n(i)} = \lim_{n \rightarrow \infty} \frac{\sum_{v=0}^n p_{jj}^{(v)}}{\sum_{v=0}^n p_{ii}^{(v)}} = \pi_{ji}.$$

Chung [2] remarked that it would be of interest to prove some sharper versions of the theorem. For example, if the Markov chain is that formed by considering sums $\{S_n\}$, $n=1, 2, \dots$, of independent and identically distributed random variables $\{U_n\}$, $n=1, 2, \dots$, which assume only integer values and have means zero, then Chung and Erdős [4] proved that

$$\Pr \left\{ \frac{|N_n(i) - N_n(j)|}{N_n(j)} > M_n^{(-1+\epsilon)/4} \text{ infinitely often} \right\} = 0$$

for every $\epsilon > 0$ where $M_n = \sum_{v=1}^n \Pr \{S_v = j\}$. In this section we shall prove a sharper version of Theorem 1.

First we define the following stochastic process $\{Y_r\}$. Let i_1, \dots, i_r be any finite subset of states and let $Y_k = i_\alpha$ if $X_{v_k} = i_\alpha$, $\alpha = 1, \dots, r$, where v_k is the k th v such that X_v equals any one of the states i_1, \dots, i_r . Lévy, in [15], considered such a "subordinate" chain.

LEMMA 1. *$\{Y_r\}$ is a Markov chain with stationary transition probabilities.*

Proof. Since i_1, \dots, i_r are recurrent states, $\{Y_k\}$, $k=1, 2, \dots$, are defined except for a set of probability measure zero. Let $\Omega_n = \{\omega | v_k = n\}$. Then for any set y_1, \dots, y_{k+1} where y_v ($v=1, \dots, k+1$) is equal to one of the values i_1, \dots, i_r we have

$$\begin{aligned} & \Pr \{Y_{k+1} = y_{k+1} | Y_k = y_k, \dots, Y_1 = y_1\} \\ &= \sum_{n=k}^{\infty} \Pr \{\Omega_n\} \Pr \{X_{n+v} = y_{k+1} \text{ for some } v \geq 1, X_{n+v'} \neq i_1, \dots, i_r \\ & \quad \text{for } 1 \leq v' < v | X_n = y_k\} \\ (1) \quad &= \sum_{n=k}^{\infty} \Pr \{\Omega_n\} \Pr \{X_v = y_{k+1} \text{ for some } v \geq 1, X_{v'} \neq i_1, \dots, i_r \\ & \quad \text{for } 1 \leq v' < v | X_0 = y_k\} \\ &= \Pr \{X_v = y_{k+1} \text{ for some } v \geq 1, X_{v'} \neq i_1, \dots, i_r \\ & \quad \text{for } 1 \leq v' < v | X_0 = y_k\}. \end{aligned}$$

Since this is true for any k and independent of y_1, \dots, y_{k-1} the lemma follows.

We shall denote the transition probability given in (1) by $\bar{p}_{i\alpha i\beta}$.

Let us consider the special case $r=2$, namely $i_1=i, i_2=j$. Here $\bar{p}_{ij} = {}_{ij}\bar{p}_{ij}^*$; $\bar{p}_{ii} = {}_{ij}\bar{p}_{ii}^*$; $\bar{p}_{ji} = {}_{ij}\bar{p}_{ji}^*$; $\bar{p}_{jj} = {}_{ij}\bar{p}_{jj}^*$ where ${}_{kl}\bar{p}_{ij}^* = \sum_{n=1}^{\infty} \Pr \{X_n=j, X_v \neq k, l \text{ for } 1 \leq v < n | X_0=i\}$. Since the original chain $\{X_n\}$ was recurrent and irreducible, all four transition probabilities are positive. Some computations, see [10], show that $\lim_{n \rightarrow \infty} \bar{p}_{ii}^{(n)} = \lim_{n \rightarrow \infty} \bar{p}_{ji}^{(n)} = {}_{ij}\bar{p}_{ji}^* / ({}_{ij}\bar{p}_{ij}^* + {}_{ij}\bar{p}_{ji}^*)$ and $\lim_{n \rightarrow \infty} \bar{p}_{jj}^{(n)} = \lim_{n \rightarrow \infty} \bar{p}_{ij}^{(n)} = {}_{ij}\bar{p}_{ij}^* / ({}_{ij}\bar{p}_{ij}^* + {}_{ij}\bar{p}_{ji}^*)$. We have, also, that

$$\bar{f}_{jj}^{(1)} = \bar{p}_{jj}; \quad \bar{f}_{jj}^{(\nu)} = \bar{p}_{ji}\bar{p}_{ij}(\bar{p}_{ii})^{\nu-2}, \quad \nu = 2, 3, \dots,$$

where $\bar{f}^{(\nu)}$ are the recurrence time probabilities of state j for the chain $\{Y_k\}$. Some further computations yield

$$(2) \quad \bar{\sigma}_{jj}^2 = \sum_{\nu=1}^{\infty} (\nu - \bar{m}_{jj})^2 \bar{f}_{jj}^{(\nu)} = \frac{{}_{ij}\bar{p}_{ji}^*}{{}_{ij}\bar{p}_{ij}^{*2}} (2 - {}_{ij}\bar{p}_{ij}^* - {}_{ij}\bar{p}_{ji}^*)$$

where $\bar{m}_{jj} = \sum_{\nu=1}^{\infty} \nu \bar{f}_{jj}^{(\nu)}$. We now prove

THEOREM 2. *If i and j are any two states belonging to the same recurrent class and if $\phi(n)$ is any nondecreasing, unbounded function, then with probability 1*

$$\left| \frac{N_n(j) - N_n(i)\pi_{ji}}{(N_n(i) + N_n(j))^{1/2}\phi(N_n(i) + N_n(j))} \right| > c_{ij}$$

will be satisfied for infinitely many or at most finitely many n according as $\sum_{n=1}^{\infty} (\phi(n)/n)e^{-(\phi^2(n))/2}$ diverges or converges where

$$c_{ij} = \bar{\sigma}_{jj} \frac{{}_{ij}\bar{p}_{ij}^{*3/2}}{{}_{ij}\bar{p}_{ji}^*} ({}_{ij}\bar{p}_{ij}^* + {}_{ij}\bar{p}_{ji}^*)^{-1/2}.$$

Proof. Let $\bar{N}_k(j) = \{\text{the number of } v\text{'s such that } y_v=j, 1 \leq v \leq k\}$. It follows from a theorem proved by Feller [9, p. 114] that with probability 1

$$(3) \quad \left| \frac{\bar{N}_k(j) - k\bar{p}_j}{k^{1/2}\bar{\sigma}_{jj}\bar{p}_j^{3/2}\phi(k)} \right| > 1$$

will be satisfied for infinitely many or at most finitely many k according as $\sum_{k=1}^{\infty} (\phi(k)/k)e^{-(\phi^2(k))/2}$ diverges or converges, where $\bar{p}_j = \lim_{n \rightarrow \infty} \bar{p}_{jj}^{(n)}$. This follows from the fact that the Markov chain under consideration has only two states and must therefore satisfy the conditions of Feller's theorem. The correspondence of the $\{Y_n\}$ chain to the $\{X_n\}$ chain defines a transformation from the space Ω of all sequences ω of non-negative integers to the space Ω' of all sequences ω' of integers i and j . Thus the inverse image of the set of all

sequences ω' having probability 1 by virtue of Feller's theorem applied to the sequence (3) must also have probability 1. Now for any ω and n , k is defined by $k = N_n(i, \omega) + N_n(j, \omega)$ and $\bar{N}_k(j) = N_n(j, \omega)$ where $N_n(\cdot, \omega)$ is the value of $N_n(\cdot)$ for a given ω . Then

$$(4) \quad \frac{N_k(j) - k\bar{p}_j}{k^{1/2}\bar{\sigma}_{jj}\bar{p}_j^{3/2}\phi(k)} = \frac{N_n(j, \omega) - \bar{p}_j(N_n(i, \omega) + N_n(j, \omega))}{(N_n(i, \omega) + N_n(j, \omega))^{1/2}\bar{\sigma}_{jj}\bar{p}_j^{3/2}\phi(N_n(i, \omega) + N_n(j, \omega))}$$

$$= c'_{ij} \frac{N_n(j, \omega) - (i_j\bar{p}_{ij}^*/i_j\bar{p}_{ji}^*)N_n(i, \omega)}{(N_n(i, \omega) + N_n(j, \omega))^{1/2}\bar{\sigma}_{jj}\phi(N_n(i, \omega) + N_n(j, \omega))}$$

where $c'_{ij} = (1 - \bar{p}_j)/\bar{p}_j^{3/2} = (i_j\bar{p}_{ji}^*/i_j\bar{p}_{ij}^{*3/2})(i_j\bar{p}_{ij}^* + i_j\bar{p}_{ji}^*)^{1/2}$. Let $c_{ij} = \bar{\sigma}_{jj}/c'_{ij}$. From a result due to Chung [2] it follows that $i_j\bar{p}_{ij}^*/i_j\bar{p}_{ji}^* = \pi_{ji}$. Since $\{X_n\}$ is recurrent we have that for all ω except for a set of probability measure zero infinitely many k implies infinitely many n and finitely many k implies finitely many n . Hence, from (3) and (4) we get the desired result.

We close this section by proving a result, intuitively expected, concerning the derived Markov chain $\{Y_n\}$.

THEOREM 3. *If $\{Y_n\}$ is the Markov chain derived for any subset i_1, \dots, i_r from an irreducible recurrent Markov chain $\{X_n\}$, and if $\{\bar{p}_{ij}\}$ is its matrix of transition probabilities, then*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{v=1}^n \bar{p}_{i_\alpha i_\alpha}^{(v)} = \frac{v_{i_\alpha}}{v_{i_1} + \dots + v_{i_r}}, \quad \alpha = 1, \dots, r.$$

Proof. From Theorem 1 we have with probability 1 that

$$\lim_{k \rightarrow \infty} \frac{\bar{N}_k(i_\alpha)}{\bar{N}_k(i_1)} = \lim_{n \rightarrow \infty} \frac{N_n(i_\alpha)}{N_n(i_1)} = \pi_{i_\alpha i_1} \quad \text{for } \alpha = 1, \dots, r.$$

But from Theorem 1, §1, we know that $\pi_{i_\alpha i_1}$, $\alpha = 1, \dots, r$, must be unique, up to a multiplicative factor, positive solutions to the equations

$$(5) \quad \sum_{\alpha=1}^r \bar{v}_\alpha \bar{p}_{i_\alpha i_\beta} = \bar{v}_\beta, \quad \beta = 1, \dots, r.$$

It follows therefore that

$$\frac{\pi_{i_\alpha i_1}}{\pi_{i_1} + \dots + \pi_{i_r i_1}} = \frac{v_{i_\alpha}}{v_{i_1} + \dots + v_{i_r}} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{v=1}^n \bar{p}_{i_\alpha i_\alpha}^{(v)}, \quad \alpha = 1, \dots, r.$$

Acknowledgments. I am indebted to Professor K. L. Chung for having originally stimulated my interest in the subject and for many helpful letters and conversations. I am also deeply grateful to Dr. T. E. Harris without whose suggestions and encouragement this paper would not have been written.

REFERENCES

1. K. L. Chung, *Lecture notes on Markov chains*, Grad. Math. Stat. Soc., Columbia University, New York, 1951.
2. ———, *Contributions to the theory of Markov chains I*, Journal of Research of the National Bureau of Standards vol. 50 (1953) pp. 203–208.
3. ———, *Limit theorems in probability*, Lecture Notes, Grad. Math. Stat. Soc., Columbia University, New York, 1950.
4. K. L. Chung and P. Erdős, *Probability limit theorems assuming only the first moment I*, Memoirs of the American Mathematical Society, no. 6, 1951.
5. C. Derman, *A solution to a set of fundamental equations in Markov chain*, Proc. Amer. Math. Soc. vol. 5 (1954) pp. 332–334.
6. W. Doeblin, *Sur deux problèmes de M. Kolmogoroff concernant les chaînes dénombrables*, Bull. Soc. Math. France vol. 66 (1938) pp. 210–220.
7. J. L. Doob, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. vol. 63 (1948) pp. 393–421.
8. ———, *Stochastic processes*, New York, Wiley, 1953.
9. W. Feller, *Fluctuation theory of recurrent events*, Trans. Amer. Math. Soc. vol. 67 (1949) pp. 98–119.
10. ———, *An introduction to probability theory and its applications*, New York, Wiley, 1950.
11. T. E. Harris, *First passage and recurrence distributions*, Trans. Amer. Math. Soc. vol. 73 (1952) pp. 471–486.
12. T. E. Harris and H. Robbins, *Ergodic theory of Markov chains admitting an infinite invariant measure*, Proc. Nat. Acad. Sci. U.S.A. vol. 39 (1953) pp. 860–864.
13. E. Hopf, *Ergodentheorie*, Chelsea, 1948.
14. A. N. Kolmogorov, *Anfangsgründe der Theorie der Markoffschen Ketten mit unendlich vielen möglichen Zuständen*, Mat. Sbornik N. S. vol. 1 (1936) pp. 607–610; Bul. Univ. État Moscou vol. 1 (1937).
15. P. Lévy, *Systèmes Markoviens et stationnaires. Cas dénombrable*, Ann. École Norm. (3) vol. 68 (1951) pp. 327–381.

COLUMBIA UNIVERSITY,
 NEW YORK, N. Y.
 SYRACUSE UNIVERSITY,
 SYRACUSE, N. Y.