

Contributions to the Theory of Markov Chains

Kai Lai Chung¹

The fundamentals of the theory of denumerable Markov chains with stationary transition probabilities were laid down by Kolmogorov, and further work was done by Doblin. The theory of recurrent events of Feller is closely related, if not coextensive. Some new results obtained by T. E. Harris turn out to tie up very nicely with some amplifications of Doblin's work. Harris was led to consider the probabilities of hitting one state before another, starting from a third one. This idea of considering three states, one initial, one "taboo", and one final, is more fully developed in the present work. The notion of first passage time to the "union" or "intersection" of two states is also introduced here. The interplay between these notions is illustrated.

The fundamentals of the theory of denumerable Markov chains² with stationary transition probabilities (DMCS) were laid down by Kolmogorov [1]³ and further work was done by Doblin [2]. The theory of recurrent events of Feller [3] is closely related, if not coextensive. Recently some interesting new results were obtained by T. E. Harris [4] and communicated to the author. They turn out to tie up very nicely with some amplifications of Doblin's work the author was engaged in. Although Harris' main purpose lies elsewhere, he was led to consider the probabilities of hitting one state before another, starting from a third one. This idea of considering three (instead of the customary two) states, one initial, one "taboo," and one final, will be more fully developed in the present work. The notion of first passage time to the "union" or "intersection" of two states will also be introduced here. The interplay between these notions will be illustrated.

Recorded results in this paper will be labeled as formulas and theorems, respectively. Relevant remarks as to their origin or significance will be found in the body of the paper. The author is indebted to Dr. Harris for communicating some of his results before publication.

1. The sequence of random variables $\{X_n\}$, $n=0, 1, 2, \dots$ forms a DMCS. The states will be denoted simply by the positive integers. The (one-step) transition probability from the state i to the state j will be denoted by $P_{ij}^{(1)}$. Because of stationarity we have

$$P_{ij}^{(1)} = P(X_{m+1} = j | X_m = i)$$

for all integers $m \geq 0$ for which the conditional probability is defined. With this understanding, we shall permit ourselves to write $m=0$ in the definitions to follow, as if the conditional probabilities were always defined.

NOTATIONS:

n, N, v, r, s , denote positive integers and will be used as time parameters or general numerals;

i, j, k, l, h , denote positive integers and will be used as state labels:

$$P_{ij}^{(n)} = P(X_n = j | X_0 = i); \quad P_{ij}^{(0)} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

$${}_k P_{ij}^{(n)} = P(X_n = j, \quad X_v \neq k, \quad 1 \leq v < n | X_0 = i)$$

$$F_{ij}^{(n)} = P(X_n = j, \quad X_v \neq j, \quad 1 \leq v < n | X_0 = i)$$

$${}_k F_{ij}^{(n)} = P(X_n = j, \quad X_v \neq j, \neq k, \quad 1 \leq v < n | X_0 = i)$$

$$Q^* = \sum_{n=1}^{\infty} Q^{(n)}$$

where Q may stand for any of the symbols ${}_k P_{ij}$, F_{ij} , or ${}_k F_{ij}$.

We offer the following clue to the above notations. The letter P designates "passage"; the letter F , "first passage"; the first right-hand subscript designates the initial state; the second, the final state; the left-hand subscript designates the "taboo state," namely, one to be eschewed during the passage (exclusive of both terminals); the star on a letter with subscripts designates the sum of the corresponding infinite series (finite or $+\infty$) summed from $n=1$ *ad inf.* We admit that this is not the most logical system of notations we could have invented. For instance, we have the superfluity $F_{ij}^{(n)} = {}_j P_{ij}^{(n)}$, and if we had allowed more than one left-hand subscript,⁴ we could have used only one letter P and written ${}_k F_{ij}^{(n)} = {}_{j,k} P_{ij}^{(n)}$. However, we consider our notations to be preferred to the arbitrary use of all sorts of letters from the Latin and Greek alphabets. Also, after painful deliberations we decided not to define ${}_k P_{ij}^{(0)}$, $F_{ij}^{(0)}$, or ${}_k F_{ij}^{(0)}$, while reserving the right to do so later in some cases.

FORMULA I: If $i \neq j$, then

$$F_{ij}^* = {}_i F_{ij}^* (1 + {}_j P_{ii}^*), \tag{I}$$

where on the right side $0 \cdot \infty$ is to be taken as 0 .⁵

⁴ This naturally suggests the consideration of more than one taboo state.
⁵ This follows also from the easily interpreted relations

$$1 + {}_i P_{ii}^* = \sum_{n=0}^{\infty} ({}_i F_{ii}^*)^n = \frac{1}{1 - {}_i F_{ii}^*} = \frac{F_{ij}^*}{{}_i F_{ij}^*}$$

The convention that $0 \cdot \infty$ is to be taken as 0 will be understood in similar circumstances.

¹ National Bureau of Standards and Institute of Statistics, University of North Carolina, Chapel Hill, N. C.
² "Denumerable" means "with a denumerable number of states;" "chain" refers to a process with an integral time parameter.
³ Figures in brackets indicate the literature references at the end of this paper.

PROOF: We start from the formula

$$F_{ij}^{(n)} = \sum_{v=0}^{n-1} j P_{ii}^{(v)} i F_{ij}^{(n-v)}, \quad (1)$$

where we agree that $j P_{ii}^{(0)} = 1$. Equation (1) is proved as follows. Either the state i is not entered at all during the passage from i to j , which contingency contributes the term corresponding to $v=0$ on the right side of (1); or there is a last entry of i , occurring at the v th step, $1 \leq v \leq n-1$, which contingency contributes the general term.

Summing (1) over n , we obtain

$$\begin{aligned} 1 \geq F_{ij}^* &= \sum_{n=1}^{\infty} F_{ij}^{(n)} = \sum_{n=1}^{\infty} \sum_{v=0}^{n-1} j P_{ii}^{(v)} i F_{ij}^{(n-v)} \\ &= \sum_{v=0}^{\infty} j P_{ii}^{(v)} \sum_{n=v+1}^{\infty} i F_{ij}^{(n-v)} = i F_{ij}^* (1 + j P_{ii}^*). \end{aligned}$$

Since the terms of the double series are nonnegative, the inversion is justified and (I) is proved. Moreover, this proves that if $i F_{ij}^* > 0$, then $j P_{ii}^* < \infty$. It follows from (I) that $i F_{ij}^* = 0$ if, and only if, $F_{ij}^* = 0$, namely, $P_{ij}^{(n)} = 0$ for all n .

FORMULA II: If $j \neq k$, then

$${}_k P_{ij}^* = {}_k F_{ij}^* (1 + {}_k P_{ij}^*). \quad (II)$$

(This formula is easily interpreted in terms of mathematical expectations.)

PROOF: We start from the formula

$${}_k P_{ij}^{(n)} = \sum_{v=1}^n {}_k F_{ij}^{(v)} {}_k P_{ij}^{(n-v)}, \quad (2)$$

where as before ${}_k P_{ij}^{(0)} = 1$. If we ignore the left-hand subscripts, (2) reduces to a familiar formula. The proof of the latter extends immediately to (2).

Summing (2) over n we obtain

$$\begin{aligned} {}_k P_{ij}^* &= \sum_{n=1}^{\infty} {}_k P_{ij}^{(n)} = \sum_{n=1}^{\infty} \sum_{v=1}^n {}_k F_{ij}^{(v)} {}_k P_{ij}^{(n-v)} \\ &= \sum_{v=1}^{\infty} {}_k F_{ij}^{(v)} \sum_{n=v}^{\infty} {}_k P_{ij}^{(n-v)} = {}_k F_{ij}^* (1 + {}_k P_{ij}^*). \end{aligned}$$

We note the following corollaries to (I) and (II), to be used later.

FORMULA IIa: If $i \neq j$, then

$$i P_{ij}^* = i F_{ij}^* (1 + i P_{ij}^*). \quad (IIa)$$

FORMULA IIb: If $j \neq k$, then

$${}_j F_{ik}^* {}_k P_{ij}^* = F_{ik}^* {}_k F_{ij}^*. \quad (IIb)$$

FORMULA IIc: If $i \neq j$, then

$$F_{ij}^* (1 + i P_{ij}^*) = i P_{ij}^* (1 + j P_{ii}^*). \quad (IIc)$$

FORMULA III: If $i \neq j$, then

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=0}^N P_{ij}^{(n)}}{\sum_{n=0}^N P_{ii}^{(n)}} = i P_{ij}^* < \infty. \quad (III)$$

PROOF: We start from the formula

$$P_{ij}^{(n)} = \sum_{v=0}^n P_{ii}^{(v)} i P_{ij}^{(n-v)}, \quad (3)$$

where we agree that $i P_{ij}^{(0)} = 0$. The proof of (3) is entirely similar to that of (1).

Summing (3) from $n=0$ to $n=N$, we obtain

$$\sum_{n=0}^N P_{ij}^{(n)} = \sum_{n=0}^N \sum_{v=0}^n P_{ii}^{(v)} i P_{ij}^{(n-v)} = \sum_{v=0}^N P_{ii}^{(v)} \sum_{n=0}^{N-v} i P_{ij}^{(n)}. \quad (4)$$

We need an elementary lemma which is frequently useful in such connections.

LEMMA. Let $0 \leq a_v \leq 1$, $b_v \geq 0$; $\sum_{v=0}^{\infty} a_v > 0$, $\lim_{v \rightarrow \infty} b_v = B \leq +\infty$. Then

$$\lim_{N \rightarrow \infty} \frac{\sum_{v=0}^N a_v b_{N-v}}{\sum_{v=0}^N a_v} = B.$$

Applying the lemma to (4) we obtain (III). That $i P_{ij}^* < \infty$ is clear from (IIa), and the remarks at the end of the proof of (I).

THEOREM 1. The limit

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=0}^N P_{ii}^{(n)}}{\sum_{n=0}^N P_{ij}^{(n)}} \quad (5)$$

exists, and is equal to any of the following three expressions:

$$\frac{1 + j P_{ii}^*}{1 + i P_{ii}^*}, \quad \frac{F_{ij}^*}{i P_{ii}^*}, \quad \frac{F_{ij}^* j F_{ii}^*}{F_{ii}^* i F_{ij}^*}, \quad (IVa, b, c)$$

the first always, the second if $i \neq j$, the third if $F_{ii}^* F_{ii}^* > 0$.

PROOF. Doblin [2] has shown, trivially, that

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=0}^N P_{ij}^{(n)}}{\sum_{n=0}^N P_{ii}^{(n)}} = F_{ij}^*. \quad (6)$$

Comparing (III) and (6), we obtain (IVb) if $i \neq j$. (IVa) now follows from (IIc) and obviously holds for $i=j$. If $F_{ii}^* F_{ii}^* > 0$, then the denominator of (IVc) is not zero, and this is then equal to (IVb) by (IIb), with $k=i$.

That the limit (5) exists, and is finite and not zero, was proved by Doblin [2]; that it is equal to (IVa) was previously proved by the author [5]. The present approach seems to be the simplest.

COROLLARY. If i, j, k, l , are distinct states of one class,⁶ then

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N P_{ij}^{(n)}}{\sum_{n=1}^N P_{kl}^{(n)}} = \frac{F_{ij}^*(1 + lP_{ii}^*)}{F_{kl}^*(1 + jP_{ll}^*)}$$

Naturally there are other expressions for it, and we omit the tedious considerations when some of the states are identical.

2. We now consider two states i and j belonging to the same recurrent class, namely:

$$F_{ii}^* = F_{jj}^* = F_{ii}^* = 1.$$

A fundamental idea in the theory of DMCS, already found in Kolmogorov's work, is that whatever transpires between successive entries at a recurrent state forms a sequence of independent events. Using this idea, Harris [4] and Lévy [10], independently of each other, discovered theorem 2. Our proof is somewhat different from theirs.

Let $i \neq j$ and define

Y_n = the number of v , $1 \leq v \leq n$, such that $X_v = i$;

Z_n = the number of v , $1 \leq v \leq n$, such that $X_v = j$.

In words, Y_n (or Z_n) is the number of entries at the state i (or j) in the first n steps. Using the average ergodic theorem (see (11) below) it is easy to show that if j is a positive state, and $P(X_0 \in C) = 1$, where C is the class containing j , then we have

$$P\left(\lim_{n \rightarrow \infty} \frac{Z_n}{\sum_{v=0}^n P_{ij}^{(v)}} = 1\right) = 1.$$

The following theorem covers both positive and null classes.

THEOREM 2 (Harris-Lévy). If i and j are two states in a recurrent class C and $P(X_0 \in C) = 1$, then

$$P\left[\lim_{n \rightarrow \infty} \frac{Z_n}{Y_n} = \lim_{n \rightarrow \infty} \frac{\sum_{v=0}^n P_{ij}^{(v)}}{\sum_{v=0}^n P_{ii}^{(v)}} = 1\right] = 1. \quad (7)$$

PROOF. Since i is recurrent, we have $P(\lim_{n \rightarrow \infty} Y_n = +\infty) = 1$. Let $v_1 < v_2 < \dots$ be the successive indices v , such that $X_v = i$. Let W_s = the number of v , $v_s < v < v_{s+1}$ such that $X_v = j$. Then as remarked above, the W_s 's are independent, identically dis-

tributed random variables. Evidently we have

$$E(W_s) = \sum_{n=1}^{\infty} {}_iP_{ij}^{(n)} = {}_iP_{ij}^* < \infty.$$

Now by definition we have $v_{Y_n} \leq n < v_{Y_n+1}$ and

$$\sum_{s=1}^{Y_n-1} W_s \leq Z_n \leq v_1 + \sum_{s=1}^{Y_n} W_s.$$

Consequently,

$$\frac{\sum_{s=1}^{Y_n-1} W_s}{Y_n} \leq \frac{Z_n}{Y_n} \leq \frac{v_1}{Y_n} + \frac{\sum_{s=1}^{Y_n} W_s}{Y_n}. \quad (8)$$

Applying Khintchine-Kolmogorov's strong law of large numbers (see, e. g., [9] p. 208) to the sequence $\{W_s\}$ we obtain:

$$P\left(\lim_{Y_n \rightarrow \infty} \frac{\sum_{s=1}^{Y_n} W_s}{Y_n} = {}_iP_{ij}^*\right) = 1. \quad (9)$$

Moreover, $P(v_1 < +\infty) = 1$. It follows from (8) and (9) that

$$P\left(\lim_{n \rightarrow \infty} \frac{Z_n}{Y_n} = {}_iP_{ij}^*\right) = 1. \quad (10)$$

Now $F_{ij}^* = 1$. Hence theorem 2 follows from (10) and theorem 1, using (IVb) there.

This theorem includes as special case a previous result by Erdős and the author [7]. Consider independent, identically distributed random variables $\{U_n\}$ which assume only integer values with mean zero. They form a DMCS with all integers as the states. Since the mean is zero, all possible states are recurrent by a theorem of Fuchs and the author [8].⁷ Without loss of generality, we may suppose that every integer is a possible, therefore recurrent, state.

Writing $S_n = \sum_{v=1}^n U_v$, we see that

$$P_{ij}^{(n)} = P(S_n = j - i).$$

Hence, $P_{ii}^{(n)} = P_{jj}^{(n)} = P_{00}^{(n)}$, and (7) becomes

$$P\left(\lim_{n \rightarrow \infty} \frac{Z_n}{Y_n} = 1\right) = 1,$$

which is theorem 8 in [7]. Needless to say, as far as this statement is concerned, Harris' approach is incomparably better. However, we note that there we actually proved a sharper result, i. e.

$$P\left(\left|\frac{Z_n - Y_n}{Y_n}\right| > M_n^{-\frac{1+\epsilon}{4}} \text{ i.o.}\right) = 0$$

⁶ Slightly generalizing Kolmogorov, we define a class to be a set of states such that for any two states i and j belonging to it, we have $F_{ij}^* F_{ji}^* > 0$. See [6].

⁷ This important step cannot be circumvented by the present, more general, approach.

for every $\epsilon > 0$, where $M_n = \sum_{v=1}^n P(S_v = i)$. See also theorem 7 in [7]. It would be of interest to investigate corresponding strong relations for the general Markov-chain case, using perhaps a more precise form of the strong law of large numbers.

3. We now consider a positive recurrent class C . According to Kolmogorov, in C all mean recurrence and first passage times are finite, namely, for all $i, j \in C$ we have

$$m_{ij} = \sum_{n=1}^{\infty} n F_{ij}^{(n)} < \infty.$$

We introduce the notions of first passage to $j\mathbf{U}k$ and to $j\mathbf{\cap}k$, as follows. Let $j \neq k$.

$T(i, j\mathbf{U}k)$ = the smallest integer $n \geq 1$ such that $X_n = j$ or $X_n = k$, whichever happens first, given that $X_0 = i$;

$T(i, j\mathbf{\cap}k)$ = the smallest integer $n \geq 1$ such that there exist two integers n_1 and n_2 such that $n_1 \leq n_2$, $1 \leq n_1 \leq n$, $1 \leq n_2 \leq n$, and $X_{n_1} = j$, $X_{n_2} = k$, given that $X_0 = i$;

$$m(i, j\mathbf{U}k) = E\{T(i, j\mathbf{U}k)\};$$

$$m(i, j\mathbf{\cap}k) = E\{T(i, j\mathbf{\cap}k)\}.$$

Let w denote the sample point. Put

$$e_n^i = \{w: X_0(w) = i, X_v(w) \neq j, \neq k \\ \text{for } 1 \leq v \leq n; X_n(w) = j\};$$

$$e^j = \bigcup_{n=1}^{\infty} e_n^j.$$

Thus e^j is the event that $X_0 = i$ and the state j is reached before the state k . Since i, j , and k belong to one recurrent class, we have

$$e^j \mathbf{U} e^k = \{w: X_0(w) = i\}.$$

Let $P(X_0 = i) = c > 0$. We have the following relations, immediate consequences of the definitions.

$$cm(i, j\mathbf{U}k) = \sum_{n=1}^{\infty} \left\{ \int_{e_n^i} nP(dw) + \int_{e_n^k} nP(dw) \right\}$$

$$cm_{ij} = \sum_{n=1}^{\infty} \left\{ \int_{e_n^i} nP(dw) + \int_{e_n^k} (n + m_{kj})dP(w) \right\}$$

$$= cm(i, j\mathbf{U}k) + P(e^k)m_{kj}$$

$$cm(i, j\mathbf{\cap}k) = \sum_{n=1}^{\infty} \left\{ \int_{e_n^i} (n + m_{jk})P(dw) \right. \\ \left. + \int_{e_n^k} (n + m_{kj})P(dw) \right\}$$

$$= cm(i, j\mathbf{U}k) + P(e^j)m_{jk} + P(e^k)m_{kj}.$$

Now by definition we have

$$\frac{P(e^j)}{c} = {}_kF_{ij}^*, \quad \frac{P(e^k)}{c} = {}_jF_{ik}^*.$$

Hence we obtain from the above:

FORMULA VI. If $j \neq k$, then

$$m(i, j\mathbf{U}k) = m_{ij} - {}_jF_{ik}^* m_{kj} = m_{ik} - {}_kF_{ij}^* m_{jk}. \quad (\text{VI})$$

FORMULA VII. If $j \neq k$, then

$$m(i, j\mathbf{\cap}k) = m_{ij} + {}_kF_{ij}^* m_{jk} = m_{ik} + {}_jF_{ik}^* m_{kj}. \quad (\text{VII})$$

Since

$${}_kF_{ij}^* + {}_jF_{ik}^* = 1$$

we deduce from (VII):

FORMULA VIII. If $j \neq k$, then

$$m_{ik} + m_{kj} - m_{ij} = {}_kF_{ij}^* (m_{jk} + m_{kj}). \quad (\text{VIII})$$

We note the following special case ($i = k$) of (VIII):

$$m_{kk} = {}_kF_{ki}^* (m_{jk} + m_{kj}). \quad (\text{VIIIa})$$

This last formula is due to Harris [4], who also derived from it the following relation:

$$\frac{m_{jj}}{m_{kk}} = \frac{{}_jF_{jk}^*}{{}_kF_{ki}^*}. \quad (\text{VIIIb})$$

Now in a positive class the ergodic theorem of Kolmogorov holds:⁸

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{v=1}^n P_{ii}^{(v)} = \frac{1}{m_{jj}}. \quad (11)$$

Thus (VIIIb) turns out to be a special case of theorem 1, using (IVc) and noting that $F_{jk}^* = F_{ki}^* = 1$.

Dividing (VIII) by the product of (VIIIa) and (VIIIb) we obtain

$$\frac{m_{ik} + m_{kj} - m_{ij}}{m_{jj}} = \frac{{}_kF_{ij}^*}{{}_jF_{ik}^*}. \quad (12)$$

By formula (IIb), the right side is ${}_kP_{ij}^*$, since $F_{ik}^* = 1$ in the present case. Thus we obtain

FORMULA (IX). If $j \neq k$ then

$${}_kP_{ij}^* = \frac{m_{ik} + m_{kj} - m_{ij}}{m_{jj}}. \quad (\text{IX})$$

As an application consider, as in the Central Limit Theorem for Markov chains, random variables $\{Y_n\}$ attached to the Markov chain $\{X_n\}$ in the following way: $Y_n = x_i$ if $X_n = i$ where the x_i 's are arbitrary real numbers.

THEOREM 3. Let i be a positive state. Given $X_0 = i$, let v_0 denote the smallest $n \geq 1$ such that $X_n = i$. Then if the series on the right-hand side converges absolutely,

⁸ This theorem actually establishes the limit of $P_{ij}^{(n)}$ as $n \rightarrow \infty$. The average form (11) is an easy consequence of a Hardy-Littlewood Tauberian theorem.

$$m(i, j\mathbf{U}k) + m(i, j\mathbf{\cap}k) = m_{ij} + m_{ik}.$$

This relation is in striking resemblance to a familiar formula in the elementary calculus of probabilities, according to which if A and B are any two events then

$$P(A\cup B) + P(A\cap B) = P(A) + P(B).$$

The generalizations of the last relation to any finite number of events is known as Poincaré's formula (see, for example, [9], p. 61); and we immediately suspect that the same may be true for the mean first passage times. This is indeed so. We define $m(i, j_1\mathbf{U} \dots \mathbf{U}j_s)$ and $m(i, j_1\mathbf{\cap} \dots \mathbf{\cap}j_s)$ as the obvious extensions from the case $s=2$. We shall also write $j_1\mathbf{\cap} \dots \mathbf{\cap}j_s \div j$ to denote $j_1\mathbf{\cap} \dots \mathbf{\cap}j_{r-1}\mathbf{\cap}j_{r+1}\mathbf{\cap} \dots \mathbf{\cap}j_s$ if $j=j_r$ ($1 \leq r \leq s$) and $j_1\mathbf{\cap} \dots \mathbf{\cap}j_s$ if j is not one of the j_r 's.

FORMULA XI. *If j_1, \dots, j_s are distinct states in a positive class to which i also belongs, then*

$$\begin{aligned} m(i, j_1\mathbf{U} \dots \mathbf{U}j_s) &= \sum_{r=1}^s m(i, j_r) - \sum_{1 \leq r_1 < r_2 \leq s} m(i, j_{r_1}\mathbf{\cap}j_{r_2}) \\ &+ \sum_{1 \leq r_1 < r_2 < r_3 \leq s} m(i, j_{r_1}\mathbf{\cap}j_{r_2}\mathbf{\cap}j_{r_3}) - \dots + \\ &(-1)^{s-1} m(i, j_1\mathbf{\cap} \dots \mathbf{\cap}j_s). \end{aligned} \quad (\text{XI})$$

PROOF. Put

$$e_n^r = \{w: X_0(w) = i, X_v(w) \neq j_1, \dots, j_s \text{ for } 1 \leq v < n, X_n(w) = j_r\}$$

$$e^r = \bigcup_{n=1}^{\infty} e_n^r.$$

We have, as at the beginning of section 3,

$$\begin{aligned} &m(i, j_{r_1}\mathbf{\cap} \dots \mathbf{\cap}j_{r_l}) \\ &= c^{-1} \sum_{r=1}^s \sum_{n=1}^{\infty} \int_{e_n^r} \{n + m(j_r, j_{r_1}\mathbf{\cap} \dots \mathbf{\cap}j_{r_l} \div j_r)\} P(dw) \\ &= m(i, j_1\mathbf{U} \dots \mathbf{U}j_s) \\ &+ \sum_{r=1}^s c^{-1} P(e^r) m(j_r, j_{r_1}\mathbf{\cap} \dots \mathbf{\cap}j_{r_l} \div j_r). \end{aligned} \quad (13)$$

Substitute (13) into the right side of (XI) and consider the typical term $c^{-1} P(e^r) m(j_r, j_{r_1}\mathbf{\cap} \dots \mathbf{\cap}j_{r_l})$, where r is distinct from r_1, \dots, r_l . It appears on the right side of (XI) once in $m(i, j_{r_1}\mathbf{\cap} \dots \mathbf{\cap}j_{r_l})$ and once in $m(i, j_r \mathbf{\cap} j_{r_1}\mathbf{\cap} \dots \mathbf{\cap}j_{r_l})$, with opposite signs; and does not appear in any other terms. Hence its net coefficient on the right side of (XI) is zero. It remains only to consider the term $m(i, j_1\mathbf{U} \dots \mathbf{U}j_s)$. This appears once in every term and hence its net coefficient is

$$\binom{s}{1} - \binom{s}{2} + \binom{s}{3} - \dots + (-1)^{s-1} = 1.$$

Therefore (XI) is established.

we have

$$E\{Y_1 + \dots + Y_{v_0} | X_0 = i\} = m_{ii} \sum_{j=1}^{\infty} \frac{x_j}{m_{jj}}, \quad (\text{Xa})$$

$$E\{(Y_1 + \dots + Y_{v_0})^2 | X_0 = i\}$$

$$= m_{ii} \sum_{j=1}^{\infty} \frac{x_j^2}{m_{jj}} + 2m_{ii} \sum_{\substack{j=1 \\ j \neq i}}^{\infty} \frac{x_j}{m_{jj}} \sum_{k=1}^{\infty} \frac{m_{ji} + m_{ik} - m_{jk}}{m_{kk}} x_k. \quad (\text{Xb})$$

PROOF. It is more convenient to consider new variables Z_n defined as follows:

$$Z_n = \begin{cases} 0 & \text{if } X_v = i \text{ for some } v, 1 \leq v < n; \\ x_j & \text{if } X_n = j \text{ and } X_v \neq i, 1 \leq v < n, \end{cases}$$

where j may be i in the last-written line. Evidently we have then

$$\begin{aligned} E\left\{\sum_{n=1}^{v_0} Y_n \mid X_0 = i\right\} &= E\left\{\sum_{n=1}^{\infty} Z_n \mid X_0 = i\right\} \\ &= \sum_{n=1}^{\infty} E\left\{Z_n \mid X_0 = i\right\} = \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} i P_{ii}^{(n)} x_j \\ &= \sum_{j=1}^{\infty} i P_{ij}^* x_j = \sum_{j=1}^{\infty} \frac{m_{ii}}{m_{jj}} x_j, \end{aligned}$$

by (IX) with $i=k$.

Furthermore, we have

$$\begin{aligned} E\left\{\left(\sum_{n=1}^{v_0} Y_n\right)^2 \mid X_0 = i\right\} &= E\left\{\left(\sum_{n=1}^{\infty} Z_n\right)^2 \mid X_0 = i\right\} \\ &= E\left\{\sum_{n=1}^{\infty} \left(Z_n^2 + 2 \sum_{1 \leq r < s \leq n} Z_r Z_s\right) \mid X_0 = i\right\}. \end{aligned}$$

As before, we obtain readily

$$E\left\{\sum_{n=1}^{\infty} Z_n^2 \mid X_0 = i\right\} = \sum_{j=1}^{\infty} \frac{m_{ii}}{m_{jj}} x_j^2.$$

Next we have

$$\begin{aligned} &\sum_{r=1}^{\infty} \sum_{s=r+1}^{\infty} E(Z_r Z_s | X_0 = i) = \\ &\sum_{r=1}^{\infty} \sum_{s=r+1}^{\infty} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} i P_{ij}^{(r)} x_j i P_{ik}^{(s-r)} x_k = \sum_{\substack{j=1 \\ j \neq i}}^{\infty} i P_{ij}^* x_j \sum_{k=1}^{\infty} i P_{ik}^* x_k. \end{aligned}$$

By (IX) this reduces to (Xb).

The two expressions on the left sides of (Xa) and (Xb) play an important role in Doblin's Central Limit Theorem for DMCS. We refer the reader to [2] for details. They are here evaluated in what seems to us more tangible terms.

4. From formulas (VI) and (VII) it follows that

⁹ Doob pointed out to me that this is a case of Wald's equation for Markov chains.

We remark that trivial as this proof is, it does not exactly correspond with the familiar proof of Poincaré's formula, and we do not know if there is any closer relation between the two apparent twins. We also leave possible extensions suggested by the known extensions of Poincaré's formula to the interested reader.

5. We now give another method of computing ${}_k P_{ij}^*$. This method requires the ergodic theorem (11). An interesting byproduct is the following:

THEOREM 5. *If i, j and k (not necessarily distinct) belong to a positive class, then*

$$\sum_{n=1}^{\infty} \{P_{ik}^{(n)} - P_{jk}^{(n)}\} = \frac{m_{jk} - m_{ik}}{m_{kk}}$$

PROOF. Using the familiar formula (cf. the remark after (2))

$$P_{ik}^{(n)} = \sum_{v=1}^n F_{ik}^{(v)} P_{kk}^{(n-v)},$$

we have

$$\sum_{n=1}^N \{P_{ik}^{(n)} - P_{jk}^{(n)}\} = \sum_{v=1}^N \{F_{ik}^{(v)} - F_{jk}^{(v)}\} \sum_{n=0}^{N-v} P_{kk}^{(n)}. \quad (14)$$

Substituting from (11), we see that the right side of (14) is, as $N \rightarrow \infty$, asymptotically equal to

$$\sum_{v=1}^N \{F_{ik}^{(v)} - F_{jk}^{(v)}\} \frac{N-v}{m_{kk}}. \quad (15)$$

Now since

$$F_{ik}^* = F_{jk}^* = 1, \quad \sum_{v=1}^{\infty} v F_{ik}^{(v)} = m_{ik} < \infty, \quad \sum_{v=1}^{\infty} v F_{jk}^{(v)} = m_{jk} < \infty,$$

we have, as $N \rightarrow \infty$

$$N \sum_{v=1}^N \{F_{ik}^{(v)} - F_{jk}^{(v)}\} = N \sum_{v=N+1}^{\infty} \{F_{ik}^{(v)} - F_{jk}^{(v)}\} \rightarrow 0.$$

Using this in (15), we see that its limit as $N \rightarrow \infty$ is

$$\lim_{N \rightarrow \infty} \frac{1}{m_{kk}} \sum_{v=1}^N (-v) \{F_{ik}^{(v)} - F_{jk}^{(v)}\} = \frac{m_{jk} - m_{ik}}{m_{kk}}.$$

We note that theorem 5 gives a convenient determination of the mean first passage times in terms of the transition probabilities; in particular

$$m_{jk} = \left(1 + \sum_{n=1}^{\infty} \{P_{kk}^{(n)} - P_{jk}^{(n)}\}\right) \lim_{n \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N P_{kk}^{(n)}.$$

We do not know what the situation is in a null class. All we can infer from theorem 1, (IVc), is that if i and j belong to one recurrent class and ${}_j F_{ii}^* \neq {}_i F_{ii}^*$, then

$$\sum_{n=1}^{\infty} \{P_{ii}^{(n)} - P_{jj}^{(n)}\} = \pm \infty.$$

To return to ${}_k P_{ij}^*$. If $j \neq k$, we have evidently

$$\sum_{h \neq k} {}_k P_{ih}^* P_{hj}^{(1)} = {}_k P_{ij}^{*(n+1)}, \quad (16)$$

where, as later, an unspecified summation runs from 1 to ∞ . Summing (16) over n , we obtain

$$\sum_{h \neq k} {}_k P_{ih}^* P_{hj}^{(1)} = {}_k P_{ij}^* - {}_k P_{ij}^{(1)},$$

or

$$\sum_h {}_k P_{ih}^* P_{hj}^{(1)} = {}_k P_{ij}^* + P_{ki}^{(1)} - P_{ij}^{(1)} \quad (17)$$

since

$${}_k P_{ik}^* = 1, \quad {}_k P_{ij}^{(1)} = P_{ij}^{(1)}.$$

We assert that in general

$$\sum_h {}_k P_{ih}^* P_{hj}^{(n)} = {}_k P_{ij}^* + \sum_{v=1}^n \{P_{ki}^{(v)} - P_{ij}^{(v)}\}.$$

This is readily shown by induction on n , starting with (17). Now sum from $n=1$ to $n=N$, divide by N , and let $N \rightarrow \infty$. By (11) and theorem 5 we obtain

$$\left(\sum_h {}_k P_{ih}^*\right) \frac{1}{m_{jj}} = {}_k P_{ij}^* + \sum_{v=1}^{\infty} \{P_{ki}^{(v)} - P_{ij}^{(v)}\} = {}_k P_{ij}^* + \frac{m_{ij} - m_{kj}}{m_{jj}}. \quad (18)$$

Now,

$$\begin{aligned} \sum_h {}_k P_{ih}^* &= \sum_{h=1}^{\infty} \sum_{n=1}^{\infty} {}_k P_{ih}^{(n)} \\ &= \sum_{n=1}^{\infty} P(X_v \neq k, 1 \leq v < n | X_0 = i) = \sum_{n=1}^{\infty} \sum_{v=n}^{\infty} F_{ik}^{(v)} = m_{ik}. \end{aligned} \quad (19)$$

(18) and (19) give (IX).

References

- [1] A. N. Kolmogorov, Anfangsgründe der Theorie der Markoffschen Ketten mit unendlich vielen möglichen Zuständen, *Matematicheskii Sbornik*, N. S., vol. **1**, 607 to 610 (1936); *Bul. Univ. Etat Moscow* **1**, (1937) (in Russian).
- [2] W. Doblin, Sur deux problèmes de M. Kolmogoroff concernant les chaînes dénombrables, *Bul. Soc. Math. France* **66**, 210 to 220 (1938).
- [3] W. Feller, Fluctuation theory of recurrent events, *Trans. Am. Math. Soc.* **67**, 98 to 119 (1949). See also [10], chap. 12.
- [4] T. E. Harris, First passage and recurrence distributions, *Trans. Am. Math. Soc.* **73**, 471 to 486 (1952).
- [5] K. L. Chung, An ergodic theorem for stationary Markov chains with a countable number of states, *Proc. International Congress of Math.* **II**, 568 (1950).
- [6] K. L. Chung, Lecture notes on Markov chains, *Graduate Mathematical Statistics Society*, Columbia University, New York, N. Y., (1951).
- [7] K. L. Chung and P. Erdős, Probability limit theorems assuming only the first moment I, *Mem. Am. Math. Soc.* **6** (1951).
- [8] K. L. Chung and W. Fuchs, On the distribution of values of sums of random variables, *Mem. Am. Math. Soc.* **6** (1951).
- [9] W. Feller, An introduction to probability theory and its applications (John Wiley & Sons, New York, N. Y., 1950).
- [10] P. Lévy, Systèmes Markoviens et stationnaires. Cas dénombrables. *Ann. Sci. École Norm. Sup.* (3), **68**, 327 to 381 (1952).

WASHINGTON, January 14, 1952.