ON LAST EXIT TIMES

BY

K. L. CHUNG[1]

1. The terminology and notation of this paper follow that of my book [1], where not explicitly explained. Results cited without amplification can also be found in the book.

Let $\{x_t, t \geqq 0\}$ be a well-separable, measurable Markov chain with the discrete state space I, the initial distribution $\{p_i\}$ and stationary, standard transition matrix $((p_{ij}))$, $i, j \in I$. Let

$$(1) \quad {}_ip_{kj}(t) = P\{x(t_0 + t, w) = j, x(t_0 + s, w) \neq i, 0 < s < t \mid x(t_0, w) = k\};$$

for every $t_0 \geqq 0$ for which the conditional probability is defined; thus ${}_ip_{kj}(t) \equiv 0$ if $k = i$ or $i = j$, by stochastic continuity. We note that if k is a stable state and $k = i$, the definition (1) differs from the one adopted in [1].

Writing as usual

$$S_i(w) = \{t : x(t, w) = i\}, \qquad \overline{S_i(w)} = \text{closure of } S_i(w),$$

we define

$$(2) \qquad \gamma_i(t, w) = \sup \{\overline{S_i(w)} \cap [0, t]\}$$

and call it *the last exit time from i before time t.* The separability and measurability of the process ensure that the corresponding w-function $\gamma_i(t)$ is a random variable. Under the hypothesis that $x(0, w) = i$, the stochastic continuity of the process implies that $\gamma_i(t)$ has a distribution function $\Gamma_i(\cdot, t)$ vanishing at zero, continuous in $(0, t)$, and making a jump of magnitude $p_{ii}(t)$ at t to reach the value one. We have clearly, if $0 \leqq s \leqq t$,

$$(3) \qquad \Gamma_i(s, t) = \textstyle\sum_{k \neq i} p_{ik}(s)[1 - F_{ki}(t - s)],$$

where F_{ki} is the *first entrance time distribution from k to i.* We define similarly

$$(4) \qquad \begin{aligned} \Gamma_{ij}(s, t) &= P\{\gamma_i(t, w) \leqq s; x(t, w) = j \mid x(0, w) = i\} \\ &= \textstyle\sum_k p_{ik}(s)\, {}_ip_{kj}(t - s), \end{aligned}$$

noting that the term corresponding to $k = i$ vanishes. Thus we have

$$\Gamma_i(s, t) = \textstyle\sum_{j \neq i} \Gamma_{ij}(s, t).$$

2. The set of sample functions with $x(0, w) = i$ and $x(t, w) = j$ can be decomposed into subsets according to the location of $\gamma_i(t, w)$ in a dyadic partition of $[0, t]$. Since the terminating dyadics $\{v2^{-n}\}$ form a separability set,

Received November 16, 1959.

[1] This research was supported by the United States Air Force through the Air Force Office of Scientific Research and Development Command.

we have by a familiar argument

$$(5)\qquad p_{ij}(t) = \lim_{n\to\infty} \sum_{v=1}^{[2^n t]} p_{ii}\left(\frac{v-1}{2^n}\right) \sum_k p_{ik}\left(\frac{1}{2^n}\right) {}_ip_{kj}\left(t - \frac{v}{2^n}\right).$$

Let us write this more suggestively as

$$(6)\qquad p_{ij}(t) = \lim_{n\to\infty} \int_0^t \phi_{ij}^{(n)}(t-s)\, d\pi_i^{(n)}(s),$$

where

$$\pi_i^{(n)}(s) = \sum_{v=1}^{[2^n s]} p_{ii}\left(\frac{v-1}{2^n}\right)\frac{1}{2^n}, \qquad \phi_{ij}^{(n)}(s) = 2^n \sum_k p_{ik}\left(\frac{1}{2^n}\right) {}_ip_{kj}(s).$$

This motivates the investigation of $\phi_{ij}^{(n)}(s)$ as $n \to \infty$, which we now proceed with. For $i \neq j$, $\delta > 0$, and $s \geqq 0$, we set

$$(7)\qquad \phi_{ij}(\delta; s) = (1/\delta)\textstyle\sum_k p_{ik}(\delta)\, {}_ip_{kj}(s).$$

1°. *Given $\varepsilon > 0$, there exists $\eta(\varepsilon) > 0$ such that for any positive integer m satisfying $m\delta \leqq \eta(\varepsilon)$ we have*

$$(8)\qquad \phi_{ij}(m\delta; s) \geqq (1-\varepsilon)\phi_{ij}(\delta; s)$$

for all $s \geqq 0$.

Proof. Consider the *discrete skeleton* $\{x_{n\delta},\ n \geqq 0\}$ and write

$$ {}_ip_{ik}^{(m-v)} = P\{x((m-v)\delta, w) = k,\ x(n\delta, w) \neq i,\ 1 \leqq n \leqq m - v - 1 \mid x(0,w) = i\}.$$

We have

$$p_{ik}(m\delta) = \sum_{v=0}^{m-1} p_{ii}(v\delta)\, {}_ip_{ik}^{(m-v)}(\delta).$$

This is the *last entrance formula* for the discrete skeleton. Furthermore it follows from the definitions that

$$ {}_ip_{ik}^{(m-v)}(\delta) = \textstyle\sum_{l\neq i} p_{il}(\delta)\, {}_ip_{lk}^{(m-v-1)}(\delta) \geqq \sum_l p_{il}(\delta)\, {}_ip_{lk}((m-v-1)\delta).$$

Hence we have, using the semigroup property of $(({}_ip_{lj}))$ for a fixed i,

$$\begin{aligned}
\phi_{ij}(m\delta; s) &\geqq \frac{1}{m\delta}\sum_{v=0}^{m-1} p_{ii}(v\delta) \sum_k \sum_l p_{il}(\delta)\, {}_ip_{lk}((m-v-1)\delta)\, {}_ip_{kj}(s) \\
&= \frac{1}{m\delta}\sum_{v=0}^{m-1} p_{ii}(v\delta) \sum_l p_{il}(\delta)\, {}_ip_{lj}((m-v-1)\delta + s) \\
&\geqq \frac{1}{m\delta}\sum_{v=0}^{m-1} p_{ii}(v\delta) \sum_l p_{il}(\delta)\, {}_ip_{lj}(s)\, {}_ip_{jj}((m-v-1)\delta).
\end{aligned}$$

We choose $\eta(\varepsilon)$ so that if $m\delta \leqq \eta(\varepsilon)$, then

$$\min\nolimits_{0\leqq v<m} p_{ii}(v\delta) \cdot \min\nolimits_{0\leqq v<m} {}_ip_{jj}((m-v-1)\delta) \geqq 1 - \varepsilon.$$

This is possible since $\lim_{t\downarrow 0} {}_ip_{jj}(t) = 1$. Then we have

$$\phi_{ij}(m\delta; s) \geqq \frac{1}{m\delta} m(1 - \varepsilon) \sum_l p_{il}(\delta) \, {}_ip_{lj}(s) = (1 - \varepsilon)\phi_{ij}(\delta; s).$$

2°. *For every $i \neq j$ and every $s \geqq 0$,*

$$\lim_{\delta\downarrow 0} \phi_{ij}(\delta; s)$$

exists and is a bounded function of s in any finite interval.

Proof. Let

$$\liminf_{\delta\downarrow 0} \phi_{ij}(\delta; s) = g_{ij}(s).$$

Since the series in (7) is dominated by the series $\sum_k p_{ik}(\delta)$ which converges uniformly with respect to δ in any finite interval, we see that $\phi_{ij}(\delta; s)$ is continuous in δ for each s. Furthermore, it follows from (8) that g_{ij} is bounded in any finite interval since $\phi_{ij}(\eta, \cdot)$ is for every η. Now for all sufficiently small δ we may choose $m\delta$ so that $\phi_{ij}(m\delta; s)$ is near $g_{ij}(s)$; hence the existence of the limit asserted in 2° follows from the inequality (8) and the definition of $g_{ij}(s)$. The above argument is similar to one by Kolmogorov [4] (cf. Theorem 2.5 of [1]); indeed Kolmogorov's theorem corresponds to the case $s = 0$ here.

If we could show that the convergence in 2° is uniform with respect to s, or equivalently (see below in 3°) that the limit function g_{ij} is continuous, then we could pass to the limit in (6) and obtain the desired result. Unfortunately we have to do this in a rather devious way. Let us denote by $P_{ij}(t)$, ${}_iP_{kj}(t)$, $G_{ij}(t)$, and $\Phi_{ij}^{(n)}(t)$ the integrals of p_{ij}, ${}_ip_{kj}$, g_{ij}, and $\phi_{ij}^{(n)}$ from 0 to t.

3°. *We have*

$$P_{ij}(t) = \int_0^t G_{ij}(t - s) p_{ii}(s) \, ds. \tag{9}$$

Proof. We note first the following complement to (5) or (6):

$$p_{ij}(t) \geqq \int_0^t \phi_{ij}^{(n)}(t - s) \, d\pi_i^{(n)}(s) \tag{10}$$

which is immediate by the sample function interpretation. We can therefore integrate (6) under the limit sign by dominated convergence and obtain

$$P_{ij}(t) = \lim_{n\to\infty} \int_0^t \Phi_{ij}^{(n)}(t - s) \, d\pi_i^{(n)}(s). \tag{11}$$

It follows from (8) that given any $\varepsilon > 0$, there exists an $n_0(\varepsilon)$ independent of s such that for all $n_2 > n_1 > n_0(\varepsilon)$ we have

$$\Phi_{ij}^{(n_1)}(s) \geqq (1 - \varepsilon)\Phi_{ij}^{(n_2)}(s) \tag{12}$$

for all $s \geqq 0$. The $\Phi_{ij}^{(n)}$ and G_{ij} are continuous functions and $\Phi_{ij}^{(n)}$ converges to G_{ij} by 1° and 2°. These are the hypotheses in Dini's theorem on uniform convergence except that the condition of monotonicity is weakened to (12). The usual proof of Dini's theorem carries over without ado, establishing that the convergence of $\Phi_{ij}^{(n)}(s)$ is uniform with respect to s in $[0, t]$ for any finite t. Now simultaneous passage to the limit of the integrand and integrator in (11) is permitted, and (9) follows since

$$\lim_{n\to\infty} \pi_i^{(n)}(t) = \int_0^t p_{ii}(s)\, ds.$$

4°. *The G_{ij}'s satisfy the following system of equations:*

$$G_{ij}(s + t) - G_{ij}(t) = \textstyle\sum_k G_{ik}(s)\; {}_ip_{kj}(t) \tag{13}$$

for all $s \geqq 0$, $t \geqq 0$.

Proof. Integrating (10) we see that

$$P_{ij}(t) \geqq \int_0^t \Phi_{ij}^{(n)}(t - s)\, d\pi_i^{(n)}(s) \geqq \Phi_{ij}^{(n)}(s)\pi_i^{(n)}(t - s)$$

for any $s \in (0, t)$. Since $\sum_j P_{ij}(t) = t$, it follows that $\sum_j \Phi_{ij}^{(n)}(s)$ converges uniformly with respect to n (and also with respect to s in any finite interval). We have by definition

$$\sum_k \Phi_{ik}^{(n)}(s)\; {}_ip_{kj}(t) = \int_0^s \sum_k \phi_{ik}^{(n)}(u)\; {}_ip_{kj}(t)\, du$$
$$= \int_0^s \phi_{ij}^{(n)}(u + t)\, du = \Phi_{ij}^{(n)}(s + t) - \Phi_{ij}^{(n)}(t).$$

Letting $n \to \infty$, we obtain (13) on account of the stated uniform convergence.

5°. *The G_{ij}'s have continuous derivatives in $[0, \infty)$ satisfying*

$$G'_{ij}(s + t) = \textstyle\sum_k G'_{ik}(s)\; {}_ip_{kj}(t) \tag{14}$$

for all $s > 0$, $t > 0$.

Proof. An application of Fubini's theorem on differentiation to the system (13) yields the system (14) for each $t > 0$ and almost all s (depending on t). Hence it also holds if $s \notin Z$ and $t \notin Z_s$ where Z and Z_s are sets of measure zero. If we consider G'_{ij} ($\geqq 0$) as the right-hand lower derivate, then we see directly from (13), upon taking the proper difference quotients and using Fatou's lemma, that the inequality obtained from (14) by changing " = " into " $\geqq$ " holds for all positive s and t. Fix an $s \notin Z$; if the left member of (14) is strictly greater than the right member for a certain value of t, then the same is true for all greater values of t by the semigroup property of $(({}_ip_{kj}))$. This being impossible by a previous assertion, the equation (14) is true for $s \notin Z$ and all $t > 0$. Now for arbitrary positive s and t write $s + t = s' + t'$ where $0 < s' < s$, $s' \notin Z$. Applying what we have just proved to $G'_{ij}(s' + t')$ and using the semigroup property again, we see that (14) is true for all positive s and t. The continuity of G'_{ij} is then a consequence of the system (14); see

Theorem 2.3 of [1]. Finally, this implies the continuous differentiability of G_{ij} by another well-known theorem of Dini.

We are now ready to state the following result.

THEOREM 1. *If $i \neq j$, then*

$$p_{ij}(t) = \int_0^t p_{ii}(s) g_{ij}(t - s)\, ds, \qquad 0 \leqq t < \infty, \tag{15}$$

where g_{ij} is the limit of $\phi_{ij}(\delta; s)$ as $\delta \downarrow 0$, uniformly with respect to s in any finite interval. The system (14) holds with $G'_{ij} = g_{ij}$; and

$$g_i(t) = \sum\nolimits_{j \neq i} g_{ij}(t)$$

is continuous and nonincreasing for $t > 0$. We have $g_{ij}(0+) = p'_{ij}(0)$ and $g_i(0+) = -p'_{ii}(0)$.

Proof. It is permissible to differentiate (9) under the integral sign; the result is (15) with g_{ij} replaced by G'_{ij}. Using this result, we have by (14)

$$\begin{aligned}\phi_{ij}(\delta; s) &= \frac{1}{\delta} \sum_k \int_0^\delta p_{ii}(\delta - u) G'_{ik}(u)\, {}_ip_{kj}(s)\, ds \\ &= \frac{1}{\delta} \int_0^\delta p_{ii}(\delta - u) G'_{ij}(u + s)\, ds.\end{aligned}$$

When $\delta \downarrow 0$, the left member tends to $g_{ij}(s)$, and the right member tends uniformly to $G'_{ij}(s)$ by continuity. Hence $g_{ij} \equiv G'_{ij}$, and (15) is proved. It now follows from (14) that if $g_i(s) < \infty$, then for all $t > 0$,

$$g_i(s + t) = \sum\nolimits_{k \neq i} g_{ik}(s)[1 - F_{ki}(t)] \leqq g_i(s) < \infty.$$

Summing (15) over $j \neq i$ we see that $g_i(s) < \infty$ for a.a. s, hence indeed for all $s > 0$; and furthermore g_i is continuous there by the equation above, since all F_{ki} are. It is not difficult to show that g_i is absolutely continuous and that for each s and a.a. t we have

$$g'_i(s + t) = -\sum\nolimits_{k \neq i} g_{ik}(s) f_{ki}(t),$$

where $f_{ki} = F'_{ki}$. The last assertion of the theorem follows from (15).

Theorem 1 has been proved by Jurkat [2]. His treatment is algebraic-analytical and does not require the "row condition" that $\sum_j p_{ij}(t) = 1$ for the transition matrix. The above proof is new and shows more directly the relation to the "movement" of the Markov chain. The probabilistic significance becomes clearer in the next statement.

THEOREM 2. *For each t, $\Gamma_{ij}(s, t)$ as defined in (4) has a continuous derivative with respect to s given by $p_{ii}(s) g_{ij}(t - s)$; and the distribution $\Gamma_i(s, t)$ as defined in (3) has a continuous density in s given by $p_{ii}(s) g_i(t - s)$. One version of the conditional probability*

$$P\{x(t, w) = j \mid x(0, w) = i, \gamma_i(t, w) = s\}$$

is equal to $g_{ij}(t-s)/g_i(t-s)$; *thus as a function of* (s,t) *it is a function of the difference* $t-s$ *only.*

Proof. We have by substituting from (15) and using (14),

$$\Gamma_{ij}(s,t) = \sum_{k\neq i}\int_0^s p_{ii}(u)g_{ik}(s-u)\ {}_ip_{kj}(t-s)\,du$$
$$= \int_0^s p_{ii}(u)g_{ij}(t-u)\,du.$$

Summing over $j \neq i$ we have

$$\Gamma_i(s,t) = \int_0^s p_{ii}(u)g_i(t-u)\,du.$$

These formulas establish the first two assertions of the theorem. The last assertion follows from (15) written in the form

$$p_{ij}(t) = \int_0^t \frac{g_{ij}(t-s)}{g_i(t-s)}\,d_s\Gamma_i(s,t).$$

The dual of Theorem 1 is well known. We give it here for the sake of comparison.

THEOREM 3. *If* $i \neq j$, *then*

$$p_{ij}(t) = \int_0^t p_{jj}(t-s)f_{ij}(s)\,ds, \qquad 0 \leqq t < \infty, \tag{16}$$

where f_{ij} *is the continuous derivative (density) of the first entrance time distribution* F_{ij}. *We have*

$$f_{ij}(s+t) = \textstyle\sum_k {}_jp_{ik}(s)f_{kj}(t), \qquad s > 0, t > 0. \tag{17}$$

Proof. That

$$p_{ij}(t) = \int_0^t p_{jj}(t-s)\,dF_{ij}(s) \tag{18}$$

is a special case of Theorem II. 11.8 of [1]; that F_{ij} has a continuous derivative f_{ij} satisfying (17) can be shown (oral communication by D. G. Austin) by differentiating the following identity:

$$1 - F_{ij}(s+t) = \textstyle\sum_k {}_jp_{ik}(s)[1 - F_{kj}(t)] \tag{19}$$

as in 5° above.

3. Since the basic formula (18) can be proved by a probabilistic argument relying on a special case (where the optional time is the first entrance time) of the strong Markov property, it is natural to ask if this genre of reasoning can also be dualized to yield a proof of Theorem 1, at least in the form corresponding to (18). This will now be shown.

For a given $T > 0$, we define the *reversed Markov chain in* $[0, T]$ as follows:

$$x_T^*(t, w) = x(T - t, w), \qquad 0 \leqq t \leqq T. \tag{20}$$

This has the state space I and the *nonstationary* transition probabilities given by

$$p_T^*(s,t;j,i) = P\{x_T^*(t,w) = i \mid x_T^*(s,w) = j\} = \frac{p_i(T-t)}{p_j(T-s)} p_{ij}(t-s), \tag{21}$$

where

$$p_i(t) = \textstyle\sum_k p_k \, p_{ki}(t)$$

is the *absolute distribution of* x_t. The *first entrance time distribution from* j *to* i *and starting at time* s is given by

$$\begin{aligned} F_T^*(s, t; j, i) &= 1 - P\{x_T^*(u, w) \neq i, s \leqq u \leqq t \mid x_T^*(s, w) = j\} \\ &= 1 - P\{x(T - u, w) \neq i, s \leqq u \leqq t \mid x(T - s, w) = j\} \\ &= 1 - \frac{1}{p_j(T-s)} \sum_k p_k(T - t) \, {}_ip_{kj}(t - s) \\ &= 1 - \frac{1}{p_j(T-s)} \sum_k p_k \Gamma_{kj}(T - t, T - s). \end{aligned} \tag{22}$$

We shall prove the following first entrance formula for a Markov chain with nonstationary transition probabilities, corresponding to (18).

THEOREM 4.[2] *Let* $\{y_t, 0 \leqq t \leqq T\}$ *be a measurable Markov chain with the state space* I *and transition probability function*

$$p(s, t; i, j) = P\{y(t, w) = j \mid y(s, w) = i\}, \qquad 0 \leqq s \leqq t \leqq T, i, j \in I.$$

The following assumptions are made:

(i) *for fixed* t, i, *and* j, $p(s, t; i, j)$ *is right continuous in* $s \leqq t$;

(ii) *there exist a set* Ω *of probability one and a denumerable dense set* $R \in [0, T]$ *such that if* $w \in \Omega$, *then for all* $t \in [0, T)$, $y(t, w)$ *is a limiting value of* $y(r, w)$ *as* $r \downarrow t$, $r \in R$.

(iii) *for each* t *and* j, $P\{t \in \overline{S_j(w)} - S_j(w)\} = 0$.

We have then if $i \neq j$, $0 \leqq s < t < T$,

$$p(s, t; i, j) = \int_s^t p(u, t; j, j) \, d_u F(s, u; i, j), \tag{23}$$

where

$$F(s, u; i, j) = P\{y(t, w) = j \text{ for some } t \in [s, u] \mid y(s, w) = i\}.$$

Remark. The condition (ii) may be roughly described as "*right separable with respect to* R." Any process has such a version.

[2] This theorem can be easily modified to yield the strong Markov property for the process. While this property in the nonstationary case has been discussed by other authors, I was unable to find a result which would cover the situation here.

Proof. Let R be enumerated as $\{r_m\}$, and let $\{r_m, 1 \leqq m \leqq n\}$ be ordered as $r_1^{(n)} < \cdots < r_n^{(n)}$. Define

$$\beta(s, w) = \inf\{t : t > s, y(t, w) = j\},$$

$$\beta^{(n)}(s, w) = \inf\{r_m^{(n)} : r_m^{(n)} > \beta(s, w); y(r_m^{(n)}, w) = j\}.$$

Clearly, β and $\beta^{(n)}$ are optional random variables (see [1] for the definition which is also valid in the nonstationary case). We shall prove that for almost all w in the set $\{w : \beta(s, w) \leqq t\}$, we have

$$P\{y(t, w) = k \mid \beta(s, w)\} = p(\beta(s, w), t; j, k) \tag{24}$$

for all k. This implies (23) when $k = j$.

To prove (24) we have only to repeat the argument in the stationary case used in [1]. It follows from (ii) that $\beta^{(n)}(s, w) \downarrow \beta(s, w)$ for almost all w. Hence if $s \leqq t' \leqq t$,

$$\begin{aligned} P\{y(t, w) = k; \beta(s, w) < t'\} &= \lim_{n\to\infty} P\{y(t, w) = k; \beta^{(n)}(s, w) < t'\} \\ &= \lim_{n\to\infty} \sum_{r_m^{(n)} < t'} P\{\beta^{(n)}(s, w) = r_m^{(n)}\} P\{y(t, w) = k \mid y(r_m^{(n)}, w) = j\} \\ &= \lim_{n\to\infty} \sum_{r_m^{(n)} < t'} P\{\beta^{(n)}(s, w) = r_m^{(n)}\} p(r_m^{(n)}, t; j, k) \\ &= \lim_{n\to\infty} \int_{\{w:\beta^{(n)}(s,w)<t'\}} p(\beta^{(n)}(s, w), t; j, k) P(dw) \\ &= \int_{\{w:\beta(s,w)<t'\}} p(\beta(s, w), t; j, k) P(dw). \end{aligned}$$

The truth of this for all $t' \leqq t$ is equivalent to (24), in view of (iii).

We now apply Theorem 4 to the reversed Markov chain $\{x_T^*(t), 0 \leqq t \leqq T\}$ defined in (20). A glance at (21) shows that condition (i) is satisfied. As for condition (ii) we need only take the version of $\{x_t, 0 \leqq t < \infty\}$, denoted by $\{x_-(t), 0 \leqq t < \infty\}$ in [1], which has the property that

$$x_-(t, w) = \liminf_{r \uparrow t, r \in R} x(r, w), \qquad 0 < t < \infty;$$

then we have

$$x_T^*(t, w) = \liminf_{r \downarrow t, r \in R_T} x_T^*(r, w)$$

where R_T consists of the numbers $T - r$ where $r \in R$, $0 \leqq r \leqq T$. Thus (ii) is satisfied. It is known that (iii) is true for $\{x_-(t)\}$, hence also for $\{x_T^*(t)\}$.

We take $p_i = 1$ for $\{x_t, t \geqq 0\}$ so that $p_k(t) \equiv p_{ik}(t)$ for all k. Applying (23) with $s = 0$, interchanging i and j, and substituting from (21) and (22), we obtain after a trivial simplification:

$$p_{ij}(t) = \int_0^t p_{ii}(t - u) \frac{-d_u \Gamma_{ij}(T - u, T)}{p_{ii}(T - u)} = \int_0^t p_{ii}(t - u)\, dG^T(u), \tag{25}$$

say. Without using the results in §2, we can proceed as follows. Each G^T is a nondecreasing continuous function in $[0, T]$ which may be normalized by making $G^T(0) = 0$. For each positive integer n, and all $T \geqq n$, we have from (25)

$$G^T(n) \leqq \frac{p_{ij}(n)}{\min_{0 \leqq t \leqq n} p_{ii}(t)} < \infty.$$

Applying Helly's selection principle first in each $[0, n]$, and then diagonalizing, we see that there exist a subsequence $G^{(n_m)}$, and a G_{ij} nondecreasing in $[0, \infty)$ and bounded in every finite interval, such that

$$\lim_{m \to \infty} G^{(n_m)}(u) = G(u), \qquad 0 \leqq u < \infty.$$

Passing to the limit in (25), we obtain

$$p_{ij}(t) = \int_0^t p_{ii}(t - u)\, dG_{ij}(u), \qquad 0 \leqq t < \infty, \tag{26}$$

for which it now appears that G_{ij} must be continuous since p_{ij} and p_{ii} are. It is possible to start from (26) and derive further properties of G_{ij}.

However, we shall look back at Theorem 2 and observe at once that

$$\frac{-1}{p_{ii}(T - u)} \frac{\partial}{\partial u} \Gamma_{ij}(T - u, T) = g_{ij}(u).$$

Hence G^T is actually independent of T, and we have identified (26) with (15). For given p_{ij} and p_{ii}, the uniqueness of the G_{ij} in (26) is of course a known fact, as is best seen by taking Laplace transforms. Finally, we remark that the generalization of (19) to the reversed Markov chain leads to the equation (13).

4. In an interesting special case the treatment of §3 can be simplified. This case in rather restrictive for the purposes of this paper, but there a "*duality principle*" holds which simplifies the preceding considerations.

Consider a *class* C of mutually communicating states containing two distinct states 1 and 2. We set

$$e_i = \int_0^\infty {}_2p_{1i}(t)\, dt + \int_0^\infty {}_1p_{2i}(t)\, dt, \qquad i \in C. \tag{27}$$

It can be shown that $0 < e_i < \infty$ for every i, and

$$\sum_j e_j\, p_{ji}(t) \leqq e_i, \qquad i \in C. \tag{28}$$

This is proved in [1; Theorem II. 13.5] for a recurrent class with the inequality strengthened into an equality; for an arbitrary class the proof requires only an obvious change which necessitates the inequality in general. The existence of a positive solution $\{e_i\}$ for the system of inequalities (28) has been shown independently by D. G. Kendall [3] without the explicit formula (27). It is known that equality holds in (28) for all i if and only if the class is recurrent, and then there is a unique positive solution apart from a constant factor. In a recurrent-positive class we may take $e_i = \lim_{t\to\infty} p_{ii}(t)$; then $\sum_{i\in C} e_i = 1$, and $\{e_i\}$ yields the stationary distribution of the Markov chain (on C). If this is taken to be the initial distribution, then the reversed chain in §3 will have stationary transition probabilities. The case under discussion here is more general in the sense that $\{e_i\}$ plays the role of a stationary pseudo-distribution even though $\sum_{i\in C} e_i$ may diverge.

We define the *dual matrix* $((p^*_{ij}))$ as follows:

$$(29) \qquad p^*_{ij}(t) = (e_j/e_i)p_{ji}(t), \qquad i, j \in C; \quad t \geqq 0.$$

The Markov (semigroup) property of the dual matrix is at once verified, but (28) shows that it is only substochastic. We may however make it stochastic by the usual device of adjoining a new state θ and setting

$$p^*_{i\theta}(t) = 1 - \textstyle\sum_{j\in C} p^*_{ij}(t); \qquad p^*_{\theta\theta}(t) \equiv 1; \qquad p^*_{\theta j}(t) \equiv 0, \quad j \in C.$$

Thus enlarged, the dual matrix becomes a stochastic transition matrix with which we may associate a Markov chain[3] $\{x^*_t, t \geqq 0\}$ with state space $C^* = C \cup \{\theta\}$ such that

$$p^*_{ij}(t) = P\{x^*(t, w) = j \mid x^*(0, w) = i\}.$$

Taking a well-separable and measurable version of the *dual chain* $\{x^*_t, t \geqq 0\}$, we can introduce the *taboo probabilities*

$$(30) \quad {}_Hp^*_{ij}(t) = P\{x^*(t, w) = j; x^*(s, w) \notin H, 0 < s < t \mid x^*(0, w) = i\},$$

where H is an arbitrary subset of C. Similarly,

$$F^*_{ij}(t) = P\{x^*(s, w) = j \quad \text{for some} \quad s \in [0, t] \mid x^*(0, w) = i\}.$$

Now we have, generalizing (29),

$$(31) \qquad {}_Hp^*_{ij}(t) = (e_j/e_i)\, {}_Hp_{ji}(t).$$

This follows from an analytic way of defining the probability in (30); we have in fact

$${}_Hp^*_{ij}(t) = \lim_{n\to\infty} \sum_{i_1\notin H} \sum_{i_2\notin H} \cdots \sum_{i_{n-1}\notin H} p^*_{ii_1}(t/n)p^*_{i_1i_2}(t/n) \cdots p^*_{i_{n-1}j}(t/n).$$

Hence (31) follows from (29).

Let us write the first entrance formula for the dual chain corresponding to (16):

$$p^*_{ji}(t) = \int_0^t p^*_{ii}(t - s)f^*_{ji}(s)\, ds, \qquad i, j \in C,$$

where f^*_{ji} is the continuous derivative of F^*_{ji}. Substituting from (29) we ob-

[3] This is not to be confused with the $x^*_T(t)$ in §3.

tain (15) with

(32) $$g_{ij}(s) = (e_j/e_i)\, f^*_{ji}(s).$$

The formula corresponding to (17) is

$$f^*_{ji}(s + t) = \sum_k {}_ip^*_{jk}(s) f^*_{ki}(t).$$

By (31) and (32) this is equivalent to

$$(e_i/e_j) g_{ij}(s + t) = \sum_k (e_k/e_j)\, {}_ip_{kj}(s)(e_i/e_k) g_{ik}(t)$$

which is (14).

While the above treatment is not adequate for the general results in §§2–3, it seems worthwhile mentioning that it is already sufficient as a basis for a derivation of Ornstein's differentiability theorem [5]. Given the state i, consider the class C in which i belongs. Applying (14) to all $j \in C - \{i\}$, for which the method of this section suffices, and summing, we obtain

(33) $$p_{ii}(t) + \int_0^t p_{ii}(t - s) c_i(s)\, ds = d_i(t)$$

where $c_i = \sum_{j \in C - \{i\}} g_{ij}$ is a continuous function and $d_i = \sum_{j \in C} p_{ij}$ is a non-decreasing, continuously differentiable function. The integral equation (33) for p_{ii} may be used as the starting point to prove the continuous differentiability of p_{ii} ; see [2] or [1].

References

1. K. L. Chung, *Markov chains with stationary transition probabilities*, Heidelberg, 1960.
2. W. B. Jurkat, *On the analytic structure of semi-groups of positive matrices*, Math. Zeitschrift, vol. 73 (1960), pp. 346–365.
3. D. G. Kendall, *Unitary dilations of one-parameter semigroups of Markov transition operators, and the corresponding integral representations for Markov processes with a countable infinity of states*, Proc. London Math. Soc. (3), vol. 9 (1959), pp. 417–431.
4. A. N. Kolmogorov, *On some problems concerning the differentiability of the transition probabilities in a temporally homogeneous Markov process having a denumerable number of states*, Moskov. Gos. Univ. Učenye Zapiski. Mat., vol. 148 (4) (1951), pp. 53-59 (in Russian).
5. D. S. Ornstein, *The differentiability of transition functions*, Bull. Amer. Math. Soc., vol. 66 (1960), pp. 36–39.

Syracuse University
Syracuse, New York