

SOME NEW DEVELOPMENTS IN MARKOV CHAINS

BY

K. L. CHUNG⁽¹⁾

1. Introduction. The purpose of this paper is to point out some new connections between the sample function behavior and the analytical properties of the transition probability functions of a continuous parameter Markov chain with stationary transition probabilities. The main idea is that of a *post-exit process* derived from the original process by considering its evolution after the exit from a stable state. This leads to various relations between conditional probabilities all of which are "intuitively obvious" but require sometimes painstaking proofs if probability theory like any other branch of mathematics is to be treated as a discipline in logic. These relations imply certain analytical properties of the transition functions, obtained recently by D. G. Austin by purely analytical means, and exhibit them in connection with other quantities introduced in this paper. They also complete some results due to Doob and Lévy. The well-known differential equations of Kolmogorov are seen to be limiting cases of certain more generally valid differential equations involving a continuous parameter. One of these expresses the fundamental transition property of the post-exit process, and the other a similar property of the renewal density of an imbedded renewal process.

2. A conditional probability. So far as possible we use the terminology and notation of [6]. Let $\{x(t), t \geq 0\}$ be a Markov chain with initial distribution $\{p_i\}$ and stationary transition matrix $\{p_{ij}(t)\}$, $t \geq 0$, $i, j = 1, 2, \dots$, such that $\lim_{t \rightarrow 0} p_{ij}(t) = \delta_{ij} = p_{ij}(0)$. Then the $p_{ij}(\cdot)$ are all uniformly continuous in $[0, \infty)$, since $|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(h)$. According to Theorem II.2.6 of [6] we may take $\{x(t)\}$ to be separable relative to the class of closed sets and measurable. Moreover the denumerable set satisfying the conditions of the separability definition will be taken to be the set of rational numbers of the form $r2^{-m}$ (see [6, pp. 56-60]). Separability requires the adjunction of the value ∞ to the range of $x(\cdot, \omega)$. We define this value ∞ as an *adjoined state* with $p_\infty = 0$ and $p_{\infty i}(t) = p_{i\infty}(t) = \delta_{i\infty}$ for all $t \geq 0$. We have then $P\{x(t) = \infty\} = 0$ for every $t \geq 0$. In the sequel, the term "state," unless specified to the contrary, shall exclude the adjoined state. The limit

$$q_i = p'_{ii}(0) = \lim_{t \downarrow 0} \frac{1 - p_{ii}(t)}{t}$$

Presented to the Society, December 29, 1955; received by the editors May 11, 1955.

⁽¹⁾ Research connected with an Air Development and Research Command Project at Syracuse University, OSR Contract No. AF(18)600-760.

exists for every i (Theorem 9 of [3]). Following Lévy the state i is called *stable* or *instantaneous* according as q_i as finite or infinite. We refer to [2] for the foundations of the theory of Markov chains under consideration. Although knowledge of these foundations will be necessary for a thorough understanding of what follows, we shall strive to make the present paper readable by itself.

Let i be a stable state with $q_i > 0$; such a state always exists unless $P\{x(t) \equiv x(0), 0 \leq t < \infty\} = 1$ [8, p. 375]. Suppose that $P\{x(0) = i\} = 1$. Let $\lambda = \lambda_i(w)$ be the "first sojourn time" in the state i , namely the length of the first t -interval in which $x(t, w) \equiv i$ (see Theorem 1 of [2]). Then $P\{\lambda \leq t\} = 1 - e^{-q_i t}$, $t \geq 0$ [3, p. 54]. Let j be an arbitrary state (not ∞ !) and define

$$\alpha = \alpha_{ij}(w) = \inf_{\lambda(w) < t, x(t, w) = j} t.$$

If $j \neq i$, α is the "first entrance time into j "; if $j = i$, α is the "second entrance time into i " (the first being zero by hypothesis). It is easily shown that α is a random variable in the broad sense, namely a measurable w -function defined on a measurable w -set whose probability may be less than one. We define its distribution function in the broad sense by $F_{ij}(t) = P\{\alpha \leq t\}$.

Now it can be proved that the two random variables λ and $\alpha - \lambda$ (which may be zero with positive probability) are independent⁽²⁾. This is a special case of Theorem 5 of [2], but we give a simple proof as follows. Let us first note that the distribution of $\alpha - \lambda$ may be derived as follows. It can be shown that⁽³⁾ $\alpha(w)$ considered as a point on the t -axis is the limit from the right of points of $S_j(w) = \{t: x(t, w) = j\}$; hence we have

$$P\{\alpha - \lambda > u \mid x(0) = i\} = \lim_{h \rightarrow 0} \sum_{n=1}^{\infty} P\{x(\tau h) = i, 0 \leq \tau < n; x(nh) \neq i, \neq j; x(\tau h) \neq j, n + 1 \leq \tau \leq n + 1 + [u/h] \mid x(0) = i\}$$

where $h = 2^{-m}$, $m \rightarrow \infty$. Now for each $s > 0$ define two random variables $\lambda_s = \lambda_s(w)$ and $\alpha_s = \alpha_s(w)$ on the set $\{w: x(s, w) = i\}$ as follows: $\lambda_s(w)$ is the supremum of T such that $x(t, w) \equiv i, s \leq t < s + T$; $\alpha_s(w)$ is the infimum of t such that $t > \lambda_s(w)$ and $x(t, w) = j$. Thus λ_0 and α_0 reduce to the previous λ

⁽²⁾ The random variables $z_1(w), \dots, z_n(w)$, with domains of definition $\Lambda_1, \dots, \Lambda_n$, are said to be independent iff $P\{\bigcap_{k=1}^n \Lambda_k [z_k(w) \leq c_k]\} / P\{\bigcap_{k=1}^n \Lambda_k\} = \prod_{k=1}^n P\{\Lambda_k [z_k(w) \leq c_k]\} / P(\Lambda_k)$ for every real c_1, \dots, c_k .

⁽³⁾ In fact, the set $S_j(w)$ is dense in itself (see [2, §4, (ii)]).

and α . It is then clear that we have by the preceding derivation and stationarity

$$P\{\alpha - \lambda > u \mid x(0) = i\} = P\{\alpha_s - \lambda_s > u \mid x(s) = i\}.$$

Now by definition we have for every $s > 0$,

$$\begin{aligned} P\{\lambda \geq s, \alpha - \lambda > u \mid x(0) = i\} &= P\{x(t) \equiv i, 0 < t < s; \alpha - \lambda > u \mid x(0) = i\} \\ &= P\{x(t) \equiv i, 0 < t \leq s \mid x(0) = i\} P\{\alpha_s - \lambda_s > u \mid x(s) = i\} \\ &= P\{\lambda \geq s \mid x(0) = i\} P\{\alpha - \lambda > u \mid x(0) = i\} \end{aligned}$$

where we have used the fact that the distribution of λ is continuous. This relation is equivalent to the independence of λ and $\alpha - \lambda$ under the hypothesis $x(0) = i$. Hence if $G_{ij}(\cdot)$ denotes the distribution function in the broad sense of $\alpha - \lambda$, then

$$(1) \quad P\{\alpha \leq u \mid \lambda\} = G_{ij}(u - \lambda)$$

for all u with probability one. We now prove that

$$(2) \quad P\{x(t) = j \mid \lambda, \alpha\} = p_{ij}(t - \alpha)$$

for almost all w on the set $\{\alpha \leq t\}$. This is a consequence of Theorem 6 of [2] but can be proved directly as follows. If $i \neq j$, we have, taking both h and δ below to be negative integral powers of 2,

$$\begin{aligned} (3) \quad &P\{\lambda \leq s_1; \alpha \leq u_1; x(t) = j\} \\ &= \lim_{h \rightarrow 0} \sum_{1 \leq n \leq u_1/h} P\{\lambda \leq s_1; (n-1)h < \alpha \leq nh; x(t) = j\} \\ &= \lim_{h \rightarrow 0} \sum_{1 \leq n \leq u_1/h} \lim_{\delta \rightarrow 0} \sum_{(n-1)h/\delta < m \leq nh/\delta} P\{\lambda \leq s_1; x(r\delta) \neq j, \\ &\quad (n-1)h/\delta < r < m; x(m\delta) = j\} p_{ij}(t - m\delta) \\ &= \lim_{h \rightarrow 0} \sum_{1 \leq n \leq u_1/h} P\{\lambda \leq s_1; (n-1)h < \alpha \leq nh\} p_{ij}(t - nh + \theta h) \\ &\quad \text{where } 0 \leq \theta \leq 1, \\ &= \int_0^{u_1} p_{ij}(t - u) d_u P\{\lambda \leq s_1; \alpha \leq u\}. \end{aligned}$$

This is equivalent to equation (2). Now let $i = j$; then $\lambda(w)$ is not the limit from the right of points of $S_i(w)$ since for a stable i the set $S_i(w)$ is the union of disjoint intervals (Theorem 1 of [2]). Hence the first term in (3) can be evaluated as follows:

$$\begin{aligned} & \lim_{h \rightarrow 0} \sum_{2 \leq n \leq [u_1/h]} \sum_{1 \leq m \leq \min(n-1, [s_1/h])} P\{(m-1)h \leq \lambda < mh; \\ & \quad x(rh) \neq i, m \leq r < n, x(nh) = i \mid x(0) = i\} p_{ii}(t-nh) \\ & = \lim_{h \rightarrow 0} \sum_{2 \leq n \leq [u_1/h]} P\{\lambda \leq \min[(n-1)h, s_1], \\ & \quad (n-1)h < \alpha \leq nh \mid x(0) = i\} p_{ii}(t-nh) \end{aligned}$$

which reduces to the same integral as before. In the above evaluations we have used the approximation of Riemann-Stieltjes sums to the Lebesgue-Stieltjes integral of a continuous function, the fact that the w -set $\{\lambda < s\}$ belongs to the Borel field generated by $\{x(t), t < s\}$, and the fact that both λ and α have continuous distributions.

By the rules of superposition of conditional probabilities (see [6, p. 36]), we have

$$P\{x(t) = j \mid \lambda\} = \int_{0-}^{t+0} P\{x(t) = j \mid \lambda, \alpha\} |_{\alpha=u} d_u P\{\alpha \leq u \mid \lambda\}$$

almost everywhere on the set $\{\lambda \leq t\}$. Evaluated at $\lambda = s \leq t$, this becomes, by (1) and (2),

$$\begin{aligned} P\{x(t) = j \mid \lambda = s\} &= \int_{s-0}^{t+0} p_{ij}(t-u) dG_{ij}(u) \\ &= \int_{0-}^{t-s+0} p_{ij}(t-s-v) dG_{ij}(v). \end{aligned}$$

We have thus proved that there exists a non-negative, Borel measurable (in fact, continuous from the right) function

$$(4) \quad r_{ij}(t) = \int_{0-}^{t+} p_{ij}(t-v) dG_{ij}(v)$$

such that

$$(5) \quad P\{x(t) = j \mid \lambda = s\} = r_{ij}(t-s)$$

for each $t > 0$ and all s in $(0, t)$ except a set (depending on t) of measure zero. We remark that the point of the proof above is that the conditional probability in (5) is essentially a function of $t-s$.

From (5) it follows from the definition of conditional probability that

$$(6) \quad p_{ij}(t) = \int_0^t q_i e^{-q_i(t-s)} r_{ij}(s) ds + \delta_{ij} e^{-q_i t}.$$

In the following an unspecified summation shall range over all the positive integers. Summing (6) over j , we have

$$1 = \int_0^t q_i e^{-q_i(t-s)} \sum_j r_{ij}(s) ds + e^{-q_i t}$$

or

$$\int_0^t q_i e^{-q_i(t-s)} ds = \int_0^t q_i e^{-q_i(t-s)} \sum_j r_{ij}(s) ds.$$

Hence we have

$$(7) \quad \sum_j r_{ij}(s) = 1$$

for almost all $s > 0$. Next, if $t' > 0$ we have by (6),

$$p_{ij}(t + t') = \sum_k p_{ik}(t) p_{kj}(t') = \int_0^t q_i e^{-q_i(t-s)} \sum_k r_{ik}(s) p_{kj}(t') ds + e^{-q_i t} p_{ij}(t').$$

Using (6) in the last term above we have, after a simple transformation,

$$(8) \quad \begin{aligned} p_{ij}(t + t') &= \int_0^t q_i e^{-q_i(t-s)} \sum_k r_{ik}(s) p_{kj}(t') ds \\ &+ \int_{-t'}^0 q_i e^{-q_i(t-s)} r_{ij}(s + t') ds + \delta_{ij} e^{-q_i(t+t')}. \end{aligned}$$

On the other hand, we have directly from (6),

$$(9) \quad p_{ij}(t + t') = \int_{-t'}^t q_i e^{-q_i(t-s)} r_{ij}(s + t') ds + \delta_{ij} e^{-q_i(t+t')}.$$

Equating (8) and (9) we obtain

$$\int_0^t q_i e^{-q_i(t-s)} r_{ij}(s + t') ds = \int_0^t q_i e^{-q_i(t-s)} \sum_k r_{ik}(s) p_{kj}(t') ds.$$

Hence for each t' and t , we have

$$(10) \quad r_{ij}(s + t') = \sum_k r_{ik}(s) p_{kj}(t')$$

for almost all s in $(0, t)$. Therefore for each $t' > 0$, (10) holds for almost all $s > 0$.

The functions on both sides of (10) are measurable functions of the pair (s, t') . Hence the set of (s, t') for which (10) is true is measurable. By Fubini's theorem, there is a set N of measure zero with the following property: if $0 < s \notin N$, then there exists a set N_s of measure zero such that if $t > s$ and $t \notin N_s$, then

$$(11) \quad r_{ij}(t) = \sum_k r_{ik}(s) p_{kj}(t - s).$$

For every $a > 0$, there exists then an $s \in N$, $0 < s < a$ and such that (7) holds.

Now the functions $p_{kj}(\cdot)$, $k=1, 2, \dots$, are all uniformly continuous in $[a, \infty)$. Hence by (7) and (11) the function $r_{ij}(\cdot)$ is uniformly continuous on $[a, \infty) - N_s$. This being true for every $a > 0$ it follows that $r_{ij}(\cdot)$ coincides with a continuous function almost everywhere on $(0, \infty)$. But, by (4), $r_{ij}(\cdot)$ is continuous from the right there, therefore $r_{ij}(\cdot)$ is continuous on $(0, \infty)$. Moreover $r_{ij}(0+)$ exists and equals $G_{ij}(0)$ by (4). Let us define $r_{ij}(0) = r_{ij}(0+)$.

We take this continuous function $r_{ij}(\cdot)$ to be the desired version of the conditional probability in (5); in other words, henceforth we set unequivocally

$$(5 \text{ bis}) \quad P\{x(t) = j \mid \lambda = s\} = r_{ij}(t - s)$$

for all $0 < s < t$.

By Fubini's theorem, the equations (7) and (10) hold for each positive $s \notin M_j$ and positive $t \notin M_j(s)$, where M_j and $M_j(s)$ have measure zero for each j and s :

$$(7 \text{ bis}) \quad \sum_i r_{ij}(s) = 1;$$

$$(10 \text{ bis}) \quad r_{ij}(s + t) = \sum_k r_{ik}(s) p_{kj}(t).$$

For $s \notin M_j$, let $t_n \notin M_j(s)$, $t_n \rightarrow t$. We have, by uniform convergence,

$$r_{ij}(s + t) = \lim_{n \rightarrow \infty} r_{ij}(s + t_n) = \sum_k r_{ik}(s) \lim_{n \rightarrow \infty} p_{kj}(t_n) = \sum_k r_{ik}(s) p_{kj}(t).$$

Hence each $M_j(s)$ is empty. Let $M = \cup_j M_j$ and s be arbitrary. Choose any positive $s' < s$, $s' \notin M$. Then by what was just proved, we have for every k ,

$$(12) \quad r_{ik}(s) = \sum_l r_{il}(s') p_{lk}(s - s').$$

Thus we have, replacing s by $s+t$ in (12),

$$\begin{aligned} r_{ij}(s + t) &= \sum_l r_{il}(s') p_{lj}(s - s' + t) \\ &= \sum_k \sum_l r_{il}(s') p_{lk}(s - s') p_{kj}(t) = \sum_k r_{ik}(s) p_{kj}(t). \end{aligned}$$

Hence M is empty and we have now proved that (10 bis) is true for all positive s and t . Summing (10 bis) over j , we obtain

$$\sum_i r_{ij}(s + t) = \sum_k r_{ik}(s) \sum_j p_{kj}(t) = \sum_k r_{ik}(s).$$

It follows that if (7 bis) is true for a value of s then it is true for all greater values of s ; hence (7 bis) is true for all positive s .

3. Further relations and properties. We now return to (6). The integral in the right member of (6), being that of a continuous function, has a continuous derivative for all $t > 0$, hence so does the left member and we obtain

$$(13) \quad p'_{ij}(t) = q_i[r_{ij}(t) - p_{ij}(t)], \quad t > 0.$$

It follows from (7 bis), now valid for all $s > 0$, that:

$$(14) \quad \sum_j p'_{ij}(t) = 0, \quad t > 0.$$

Substituting (13) into (10 bis), valid for all $s > 0$, $t > 0$, we obtain

$$(15) \quad p'_{ij}(t+s) = \sum_k p'_{ik}(t)p_{kj}(s).$$

The existence and continuity of $p'_{ij}(t)$, and the results (14) and (15), all proved under the sole assumption that $q_i < \infty$, were first established by D. G. Austin in a recent paper [1]. His proof is purely analytical and elementary, and is shorter than ours if we take into account all the preliminary measure-theoretic considerations which lead up to (6). However, once this formula is written down, many will probably accept it on faith. In fact, the formulas (5 bis), (7 bis), and (10 bis) all belong to the category of the "intuitively obvious." A real advantage of our approach is that it ties together probabilistically meaningful quantities in various ways, of which we now proceed to adduce a few more.

From (4) and the continuity of $r_{ij}(\cdot)$ it follows that $G_{ij}(t)$ is continuous for $t > 0$. Recalling that $F_{ij}(t)$, $G_{ij}(t)$, and $1 - e^{-q_i t}$ ($t \geq 0$) are respectively the distribution functions of α , $\alpha - \lambda$, and λ , and that the last two random variables are independent, we have

$$(16) \quad F_{ij}(t) = \int_{0-}^t [1 - e^{-q_i(t-s)}] dG_{ij}(s) = \int_0^t q_i e^{-q_i(t-s)} G_{ij}(s) ds.$$

Since the last member of (16) has a continuous derivative for $t > 0$, so does the first member and we obtain

$$(17) \quad F'_{ij}(t) = q_i[G_{ij}(t) - F_{ij}(t)].$$

This sharpens a result due to Lévy [8, Lemma II.8.1] that $F_{ij}(\cdot)$ is absolutely continuous. Furthermore, it follows from (4) that at a point t where $G'_{ij}(t)$ exists, $r'_{ij}(t)$ also exists and we have

$$r'_{ij}(t) = G'_{ij}(t) + \int_{0-}^t p'_{ij}(t-s) dG_{ij}(s).$$

Consequently by (13) the second derivative $p''_{ij}(t)$ exists wherever $G'_{ij}(t)$ exists, that is almost everywhere since $G_{ij}(\cdot)$ is monotone. To complete the picture we add the following relation, which is implicit in (3):

$$(18) \quad p_{ij}(t) = \int_0^t p_{ij}(t-s) dF_{ij}(s) + \delta_{ij} e^{-q_i t}.$$

Upon differentiation this yields

$$(19) \quad p'_{ij}(t) = F'_{ij}(t) + \int_0^t p'_{ij}(t-s)F'_{ij}(s)ds - \delta_{ij}q_i e^{-q_i t}.$$

Following Doob [3, Theorem 9] we write $q_{ij} = p'_{ij}(0) = \lim_{t \downarrow 0} (p_{ij}(t)/t)$, $i \neq j$; the limit exists and is finite for every i by a result of Kolmogorov [7]. We have from (13), $p'_{ij}(0+) = q_i[r_{ij}(0) - \delta_{ij}]$. The continuity of $p_{ij}(\cdot)$ in $[0, \infty)$, that of $p'_{ij}(\cdot)$ in $(0, \infty)$, and the fact that $p'_{ij}(0+)$ exists as a finite number imply by the mean value theorem that $q_{ij} = p'_{ij}(0+)$; similarly $q_i = -p'_{ii}(0+)$. Thus we have

$$(20) \quad r_{ij}(0) = G_{ij}(0) = \frac{(1 - \delta_{ij})q_{ij}}{q_i}.$$

This extends a result due to Doob [3, Theorem 10] from a stable to an arbitrary j . Since obviously $F_{ij}(0) = 0$ we have by (17) that $F'_{ij}(0+) = (1 - \delta_{ij})q_{ij}$.

Turning to the limits as $t \rightarrow \infty$, we have by (16)

$$(21) \quad F_{ij}(+\infty) = G_{ij}(+\infty).$$

Lévy [7, Theorem II.8.2] has proved that $\lim_{t \rightarrow \infty} p_{ij}(t) = \pi_{ij}$ exists for every i and j ; hence it follows from (4) that $\lim_{t \rightarrow \infty} r_{ij}(t)$ exists and from (6) that $\lim_{t \rightarrow \infty} r_{ij}(t) = \pi_{ij}$. Consequently we have by (13)

$$(22) \quad \lim_{t \rightarrow \infty} p'_{ij}(t) = 0.$$

Letting $t \rightarrow \infty$ in (10 bis) we obtain on account of (7 bis)

$$\pi_{ij} = \sum_k r_{ik}(s)\pi_{ki}, \quad s > 0.$$

The last equation is true if $r_{ik}(\cdot)$ is replaced by $p_{ik}(\cdot)$, hence it is also true if $r_{ik}(\cdot)$ is replaced by $p'_{ik}(\cdot)$, on account of (13).

We summarize some of the results as follows.

Under the sole assumption that $0 < q_i < \infty$:

THEOREM 1. *One version of the conditional probability $P\{x(t) = j | \lambda = s\}$ is a continuous function $r_{ij}(t-s)$ for $0 < s < t < \infty$ which satisfies the equations (6), (7 bis), and (10 bis) for all $s > 0$ and $t > 0$.*

THEOREM 2. *The distribution function $G_{ij}(\cdot)$ of $\alpha_{ij} - \lambda_i$ is a continuous function in $[0, \infty)$ satisfying (20) and (21). The distribution function $F_{ij}(\cdot)$ of α_{ij} has a continuous derivative in $[0, \infty)$ given by (17).*

THEOREM 3. *The function $p_{ij}(\cdot)$ has a continuous derivative in $[0, \infty)$ given by (13) satisfying (14), (15), and (22). The second derivative exists as a finite number almost everywhere.*

THEOREM 4. *The functions $p_{ij}(\cdot)$, $r_{ij}(\cdot)$, $F_{ij}(\cdot)$, and $G_{ij}(\cdot)$ are related by (4), (6), (16), and (18).*

4. The post-exit process. We define a new process $\{y(t), t > 0\}$, as follows:

$$(23) \quad y(t) = y(t, w) = x(\lambda_i(w) + t, w), \quad t > 0.$$

We shall see later that this indeed yields a process which may be taken as separable relative to the closed sets and measurable. This process will be called a *post-exit process, derived from the Markov chain $x(t)$ and the stable state i* . The following theorem shows that as far as its transition is concerned, it is exactly like the original $x(t)$ process. This transparent result proves to be surprisingly elusive of proof; no doubt this is due to the fact the sample function behavior to the immediate right of $\lambda_i(w)$ is not simple, see the Complement to the theorem below.

THEOREM 5. *Let $\{y(t); t > 0\}$ be the post-exit process defined by (23). Let $0 < t_1 < t_2 < \dots < t_l$, and j_1, j_2, \dots, j_l be arbitrary states. Then we have*

$$P\{y(t_\nu) = j_\nu, 1 \leq \nu \leq l\} = r_{ij_1}(t_1) \prod_{\nu=1}^{l-1} p_{j_\nu j_{\nu+1}}(t_{\nu+1} - t_\nu).$$

Proof. We need some general information regarding the sample functions of the $x(t)$ process. Detailed theorems of this kind can be found in [2], but the relevant facts here can be briefly described as follows. Let $S_j(w) = \{t: x(t, w) = j\}$. First, it is true with probability one that every point of $S_j(w)$ is a limit point of $S_j(w)$ ⁽⁴⁾. Next, by a fundamental theorem of Doob [3, Theorem 12], as $t \rightarrow \tau(w)$ ⁽⁵⁾ from one side there is at most one finite limiting value of $x(t, w)$, with probability one. Moreover, it is possible to make the separable process $x(t)$ so that $x(\tau(w), w) = +\infty$ if and only if $\lim_{t \rightarrow \tau(w)} x(t, w) = +\infty$. Consequently if $\tau(w)$ is a limit point of $S_i(w)$ from both sides, then $x(\tau(w), w) = i$. We have thus proved the following inclusions with probability one

$$\begin{aligned} & \bigcap_m \bigcap_{\nu=1}^l \{S_{j_\nu} \cap (\lambda + t_\nu - m^{-1}, \lambda + t_\nu) \neq \emptyset, S_{j_\nu} \cap (\lambda + t_\nu, \lambda + t_\nu + m^{-1}) \neq \emptyset\} \\ & \quad \subseteq \bigcap_{\nu=1}^l \{x(\lambda + t_\nu) = j_\nu\} \\ & \quad \subseteq \bigcap_m \bigcap_{\nu=1}^l \{S_{j_\nu} \cap (\lambda + t_\nu - m^{-1}, \lambda + t_\nu + m^{-1}) \neq \emptyset\}. \end{aligned}$$

⁽⁴⁾ See footnote (3).

⁽⁵⁾ For lack of a better notation, we write $\tau(w)$ to denote a generic abscissa in the "graph" $(\tau, x(\tau, w))$ corresponding to the generic w . Thus it is neither fixed irrespective of w nor properly speaking a function of w .

We prove that the last-written w -set above is measurable as follows. Let $f(t, w) = 1$ if $t \in S_j(w)$ and $= 0$ otherwise; let $g(t, w) = 1$ if $\lambda(w) + t_\nu - m^{-1} < t < \lambda(w) + t_\nu + m^{-1}$ and $= 0$ otherwise. Then both f and g are measurable functions of the pair (t, w) . Thus the function fg is measurable in the pair and by Fubini's theorem for almost all s , $f(s, w)g(s, w)$ is a measurable function of w . Thus there exists a denumerable and everywhere dense t -set $\{s_n\}$ such that $\sum_n f(s_n, w)g(s_n, w)$ is a measurable function of w . This set $\{s_n\}$ may be taken to be the set satisfying the conditions of the separability definition of the $x(t)$ process, since $\lim_{t \rightarrow 0} p_{ij}(t) = \delta_{ij}$ [6, Theorem II.2.2]. We have

$$\{S_{j\nu} \cap (\lambda + t_\nu - m^{-1}, \lambda + t_\nu + m^{-1}) \neq \emptyset\} = \left\{ \sum_n f(s_n, w)g(s_n, w) > 0 \right\}$$

which is therefore a measurable w -set, as was to be proved.

Let $0 < s < t_1$. We have, setting $h = 2^{-p}$, $p \rightarrow \infty$,

$$\begin{aligned} &P \left\{ \bigcap_m \bigcap_{\nu=1}^l [S_{j\nu} \cap (\lambda + t_\nu - m^{-1}, \lambda + t_\nu + m^{-1}) \neq \emptyset] \right\} \\ &\leq \lim_{m \rightarrow \infty} \lim_{h \rightarrow 0} \sum_{n=0}^{\infty} \sum_j P \{nh \leq \lambda < (n+1)h, x(nh+h+s) = j\} \\ &\quad \cdot P \{S_{j\nu} \cap (t_\nu + nh - m^{-1}, t_\nu + nh + h + m^{-1}) \neq \emptyset, 1 \leq \nu \leq l \mid x(nh+h+s) = j\} \\ &= \lim_{m \rightarrow \infty} \lim_{h \rightarrow 0} \sum_{n=0}^{\infty} \sum_j P \{nh \leq \lambda < (n+1)h, x(nh+h+s) = j\} \\ &\quad \cdot P \{S_{j\nu} \cap (t_\nu - s - h - m^{-1}, t_\nu - s + m^{-1}) \neq \emptyset, 1 \leq \nu \leq l \mid x(0) = j\} \\ &\leq \lim_{h \rightarrow 0} \sum_{n=0}^{\infty} \sum_j P \{nh \leq \lambda < (n+1)h, x(nh+h+s) = j\} \\ &\quad \cdot \lim_{m \rightarrow \infty} P \{S_{j\nu} \cap (t_\nu - s - 2m^{-1}, t_\nu - s + m^{-1}) \neq \emptyset, 1 \leq \nu \leq l \mid x(0) = j\}. \end{aligned}$$

The last step is justified as follows:

$$\begin{aligned} &\sum_{n=0}^{\infty} \sum_j P \{nh \leq \lambda < (n+1)h, x(nh+h+s) = j\} \\ &= \sum_{n=0}^{\infty} \sum_j \int_{nh}^{(n+1)h} q_j e^{-q_i u} r_{ij}(nh+h+s-u) du \\ &= \sum_j \int_0^{\infty} q_j e^{-q_i u} r_{ij}([uh^{-1}]h+h+s-u) du. \end{aligned}$$

Since $r_{ij}(\cdot)$ is non-negative and continuous, the convergence in (7bis) is uniform with respect to s in every $[a, b]$ where $0 < a < b < \infty$. Hence the last series above converges uniformly with respect to h in $(0, 1)$. This is sufficient for the interchange of limits in the last step.

Similarly, we have

$$\begin{aligned}
 P \left\{ \bigcap_{m=1}^l \bigcap_{\nu=1}^l [S_{i_\nu} \cap (\lambda + t_\nu - m^{-1}, \lambda + t_\nu) \neq 0; S_{i_\nu} \cap (\lambda + t_\nu, \lambda + t_\nu + m^{-1}) \neq 0] \right\} \\
 \geq \lim_{h \rightarrow 0} \sum_{n=0}^{\infty} \sum_j P \{ nh \leq \lambda < (n+1)h, x(nh + h + s) = j \} \\
 \cdot \lim_{m \rightarrow \infty} P \{ S_{i_\nu} \cap (t_\nu - s - m^{-1}, t_\nu - s - 2^{-1}m^{-1}) \neq 0, \\
 S_{i_\nu} \cap (t_\nu - s, t_\nu - s + 2^{-1}m^{-1}) \neq 0, 1 \leq \nu \leq l \mid x(0) = j \}.
 \end{aligned}$$

Now for fixed t_ν , it follows from a theorem of Doob [4, p. 457] (see Theorem 2 of [2]) that the two limits in m at the end of the last two sets of inequalities have the same value given by

$$p_{i i_1}(t_1 - s) \prod_{\nu=1}^{l-1} p_{i_\nu i_{\nu+1}}(t_{\nu+1} - t_\nu) = p_{i i_1}(t_1 - s)Q$$

where Q stands for the product in ν . Hence we have, since our basic probability is tacitly assumed to be a complete measure:

$$\begin{aligned}
 P \{ y(t_\nu) = j_\nu \} \\
 = \lim_{h \rightarrow 0} \sum_{n=0}^{\infty} \sum_j P \{ nh \leq \lambda < (n+1)h, x(nh + h + s) = j \} p_{i i_1}(t_1 - s)Q.
 \end{aligned}$$

Using (5) and (10 bis), we have

$$\begin{aligned}
 P \{ y(t_\nu) = j_\nu \} &= Q \lim_{h \rightarrow 0} \sum_{n=0}^{\infty} \sum_j \int_{nh}^{(n+1)h} q_i e^{-q_i u} r_{i j}(nh + h + s - u) p_{i i_1}(t_1 - s) du \\
 &= Q \lim_{h \rightarrow 0} \sum_j \int_0^{\infty} q_i e^{-q_i u} r_{i j}([uh^{-1}]h + h + s - u) p_{i i_1}(t_1 - s) du \\
 &= Q \lim_{h \rightarrow 0} \int_0^{\infty} q_i e^{-q_i u} r_{i j}([uh^{-1}]h + h + t_1 - u) du \\
 &= Q \int_0^{\infty} q_i e^{-q_i u} r_{i j_1}(t_1) du = r_{i j_1}(t_1)Q
 \end{aligned}$$

since $0 \leq r_{ij}(\cdot) \leq 1$ and $r_{ij}(\cdot)$ is continuous. This proves the theorem.

COROLLARY 1. *There is a standard modification of the $y(t)$ process which is separable relative to the class of closed sets and measurable.*

Proof. By Theorem II.2.6 of [6], it is sufficient to prove that for every $t > 0$ and $\epsilon > 0$

$$\lim_{s \rightarrow t} P \{ | y(t) - y(s) | < \epsilon \} = 1.$$

Now we have by the theorem, if $0 < \epsilon < 1$,

$$\begin{aligned} \lim_{s \uparrow t} P \{ |y(t) - y(s)| < \epsilon \} &= \lim_{s \uparrow t} \sum_j P \{ y(s) = y(t) = j \} \\ &= \lim_{s \uparrow t} \sum_j r_{ij}(s) p_{ij}(t - s) \geq \sum_j r_{ij}(t) = 1 \end{aligned}$$

by (7 bis); similarly for $s \downarrow t$.

COROLLARY 2. *For each $t > 0$, $P \{ y(t) = +\infty \} = 0$. Thus the $y(t)$ process as well as the $x(t)$ process has no "fixed infinities."*

COMPLEMENT TO THEOREM 5. *For every stable j , we have*

$$P \left\{ \lim_{t \downarrow 0} y(t) = j \right\} = \frac{(1 - \delta_{ij})q_{ij}}{q_i};$$

for every instantaneous j , we have

$$\begin{aligned} P \left\{ \liminf_{t \downarrow 0} y(t) = j < +\infty = \limsup_{t \downarrow 0} y(t) \right\} &= \frac{q_{ij}}{q_i}; \\ P \left\{ \lim_{t \downarrow 0} y(t) = +\infty \right\} &= 1 - \frac{\sum_{i \neq j} q_{ij}}{q_i}. \end{aligned}$$

The Complement follows from (20) and Theorem 4 of [2]. Examples of all three kinds of behavior can be found in Kolmogorov [7] and Lévy [8, pp. 365-367].

5. A counterpart. The exact counterpart to a post-exit process would be a *pre-entrance process*. We shall not introduce this formally but choose the following shorter course.

Let j be a stable state. Whatever the initial distribution of $x(0)$ may be, the t -set $S_j(w) = \{t : x(t, w) = j\}$ consists of a sequence of disjoint intervals, called the j -intervals, of which there is a finite number in every finite t -interval, with probability one. The total number of j -intervals may be finite or infinite. The beginning of the n th j -interval is a random variable (in the broad sense) denoted by $\tau_j^n = \tau_j^n(w)$ and called the n th entrance time into the state j ; the length of the n th j -interval is a random variable (in the broad sense) denoted by $\lambda_j^n = \lambda_j^n(w)$ and called the n th sojourn time in the state j . If $P \{ x(0) = i \} = 1$ then we have $P \{ \tau_j^1 \leq t \} = F_{ij}(t)$; $P \{ \tau_j^{n+1} - \tau_j^n \leq t \} = F_{jj}(t)^{(6)}$, $n = 1, 2, \dots$; and $P \{ \lambda_j^n \leq t \} = 1 - e^{-q_i t}$, $t \geq 0$. The random variables τ_j^n and λ_j^n are independent; and the sequence of random variables $\{ \tau_j^{n+1} - \tau_j^n \}$, $n = 0, 1, 2, \dots$, where $\tau_j^0 = 0$, are independent. All the preceding statements can be found in Lévy's work [8]; proofs can be found in [2].

We have, clearly, if $t > 0$,

⁽⁶⁾ Note that if $q_j = 0$ then $F_{jj}(\cdot) \equiv 0$.

$$\begin{aligned}
 P\{x(t) = j\} &= \sum_{n=1}^{\infty} P\{\tau_j^n \leq t < \tau_j^{n+1}, x(t) = j\} \\
 &= \sum_{n=1}^{\infty} P\{\tau_j^n \leq t, \lambda_j^n > t - \tau_j^n\}.
 \end{aligned}$$

Hence we have, by the independence of τ_j^n and λ_j^n ,

$$(24) \quad P\{x(t) = j\} = \sum_{n=1}^{\infty} \int_0^{t+0} e^{-q_j(t-s)} dP\{\tau_j^n \leq s\}.$$

This formula has been given by Lévy [8, (II.8.1)]. Let $P\{x(0) = i\} = 1$, then

$$\sum_{n=1}^{\infty} P\{\tau_j^n \leq s\} = U_{ij}(s),$$

where

$$U_{ij}(s) = \delta_{ij}\epsilon(s) + \sum_{n=0}^{\infty} F_{ij}(s) * F_{ij}^{n*}(s)$$

and F^{n*} denotes the n -fold convolution of F with itself, $F_{jj}^{0*}(s)$ being the unitary distribution function $\epsilon(s)$. The function $U_{ij}(s)$, which is finite for finite s and nondecreasing, is well-known in renewal theory and may be interpreted as the expected number of entrances into j in time s (see [5]).

So far the initial state i is arbitrary. We now assume that i is a stable state. Then by Theorem 2 both $F_{ij}(\cdot)$ and $F_{jj}(\cdot)$ have continuous derivatives. It follows easily that the derivative

$$(25) \quad u_{ij}(s) = [U_{ij}(s) - \delta_{ij}\epsilon(s)]' = \int_0^s F'_{ij}(s-v) d\left[\sum_{n=0}^{\infty} F_{ij}^{n*}(v)\right]$$

exists and is non-negative and continuous for all $s > 0$. We can now rewrite (24) as

$$(26) \quad p_{ij}(t) = \int_0^t e^{-q_j(t-s)} u_{ij}(s) ds + \delta_{ij}e^{-q_j t}.$$

Differentiating, we obtain

$$(27) \quad p'_{ij}(t) = u_{ij}(t) - q_j p_{ij}(t)$$

which is the counterpart to (13). Note however that (27) is proved under the double condition: $q_i < \infty$ and $q_j < \infty$.

Henceforth we shall assume the more stringent condition that $q_k < \infty$ for every k . Writing k for i in (26), multiplying by $p_{ik}(t')$ and summing over k , we obtain, for every $t' > 0$:

$$\begin{aligned}
 \sum_k p_{ik}(t') p_{kj}(t) &= \int_0^t e^{-q_j(t-s)} \sum_k p_{ik}(t') u_{kj}(s) ds + p_{ij}(t') e^{-q_j t} \\
 (28) \qquad \qquad \qquad &= \int_0^t e^{-q_j(t-s)} \sum_k p_{ik}(t') u_{kj}(s) ds \\
 &\quad + \int_{-t'}^0 e^{-q_j(t-s)} u_{ij}(t' + s) ds + \delta_{ij} e^{-q_j(t'+t)}.
 \end{aligned}$$

On the other hand, we have

$$(29) \qquad p_{ij}(t' + t) = \int_{-t'}^t e^{-q_j(t-s)} u_{ij}(t' + s) ds + \delta_{ij} e^{-q_j(t'+t)}.$$

Equating (28) and (29), we have

$$(30) \qquad \sum_k p_{ik}(t) u_{kj}(s) = u_{ij}(t + s)$$

for each $t > 0$ and almost all $s > 0$. Since both members of (30) are measurable functions of the pair (t, s) , it follows by Fubini's theorem that there are sets N and N_s of measure zero such that if $0 < s \notin N$ and $0 < t \notin N_s$ then (30) holds for all i and j . Furthermore since the functions $u_{kj}(\cdot)$ are non-negative and continuous, it follows from (30) that

$$(31) \qquad \sum_k p_{ik}(t) u_{kj}(s) \leq u_{ij}(t + s)$$

for all positive t and s . Suppose that for some t and s strict inequality holds in (31), then we have since $p_{ii}(t') > 0$ for every $t' > 0$,

$$\begin{aligned}
 \sum_k p_{ik}(t' + t) u_{kj}(s) &= \sum_l p_{il}(t') \sum_k p_{lk}(t) u_{kj}(s) \\
 &< \sum_l p_{il}(t') u_{lj}(t + s) \leq u_{ij}(t' + t + s).
 \end{aligned}$$

Thus if (30) does not hold for certain s and t , then it does not hold for the same s and all greater t . Thus the sets N_s for $s \notin N$ must be all empty. Let s and t be positive but arbitrary. Choose $s' < s$ such that $s - s' \notin N$. Then we have just proved that

$$(32) \qquad u_{kj}(t + s - s') = \sum_l p_{kl}(t') u_{lj}(s - s')$$

for all k and j . Hence we have, using (32) with s' resp. $t + s'$ for t ,

$$\begin{aligned}
 \sum_k p_{ik}(t) u_{kj}(s) &= \sum_l \sum_k p_{ik}(t) p_{kl}(s') u_{lj}(s - s') \\
 &= \sum_l p_{il}(t + s') u_{lj}(s - s') = u_{ij}(t + s).
 \end{aligned}$$

Therefore (30) holds in fact for all positive t and s . The relation (30) may be interpreted as a transition property of the "entrance frequency" $u_{ij}(s)$.

Substituting (27) into (30), we obtain⁽⁷⁾

$$(33) \quad p'_{ij}(t + s) = \sum_k p_{ik}(t) p'_{kj}(s).$$

This is the exact counterpart to (15), proved however under the more stringent condition that all $q_k < \infty$.

We summarize the results as follows.

THEOREM 6. *If $q_i < \infty$ and $q_j < \infty$ then the function $u_{ij}(s)$ given by (25) is continuous for all $s > 0$ and satisfies (26) and (27). If $q_k < \infty$ for every k then the relations (30) and (33) hold.*

6. The generalized Kolmogorov equations. The pair of equations (15) and (33):

$$p'_{ij}(t + s) = \sum_k p'_{ik}(s) p_{kj}(t), \quad p'_{ij}(t + s) = \sum_k p_{ik}(t) p'_{kj}(s)$$

may be called the *generalized Kolmogorov equations*. The first equation (obtained first by Austin) is proved under the hypothesis that $q_i < \infty$, the second under the hypothesis that $q_k < \infty$ for every k ; they are both valid for all positive t and s . Since $p'_i(\cdot)$ is a continuous function, we have as $s \downarrow 0$:

$$p'_{ij}(t) = \lim_{s \downarrow 0} \sum_k p'_{ik}(s) p_{kj}(t), \quad t > 0,$$

$$p'_{ij}(t) = \lim_{s \downarrow 0} \sum_k p_{ik}(t) p'_{kj}(s), \quad t > 0.$$

The historical approach is equivalent to attempts to take the limits above underneath the summation sign, and it is found necessary to impose further conditions to make this legitimate (see [4]). From our point of view the equations (15) and (33) are consequences of (10 bis) and (30), respectively, which are probabilistically more meaningful. They have the further analytical advantage of involving only non-negative quantities. To illustrate this we prove the following theorem which sharpens a result of Doob [4, p. 461].

THEOREM 7. *Suppose that $q_i = -p'_i(0) < \infty$. If the equation*

$$(34) \quad p'_{ij}(t) = \sum_k p'_{ik}(0) p_{kj}(t)$$

is true for some $t > 0$ and every j , then

$$(35) \quad \sum_k p'_{ik}(0) = 0;$$

⁽⁷⁾ Although the equation (33) is given in Austin's paper [1], according to an oral communication his proof is incomplete.

conversely if (35) is true, then (34) is true for every $t > 0$ and j . Furthermore, either assertion is equivalent to the following assertion regarding the post-exit process (23):

$$P \left\{ \lim_{t \downarrow 0} y(t) = +\infty \right\} = 0.$$

Proof. (34) and (35) are equivalent, by means of (13), to the following limit forms of (10 bis) and (7 bis):

$$(36) \quad r_{ij}(t) = \sum_k r_{ik}(0) p_{kj}(t),$$

$$(37) \quad \sum_k r_{ik}(0) = 1.$$

If (36) is true for some $t > 0$ and every j , then summing over j and using (7 bis) we have $1 = \sum_j r_{ij}(t) = \sum_k r_{ik}(0) \sum_j p_{kj}(t) = \sum_k r_{ik}(0)$. Conversely, suppose (37) is true. By letting $s \rightarrow 0$ in (10 bis) we have for every $t > 0$

$$(38) \quad r_{ij}(t) \geq \sum_k r_{ik}(0) p_{kj}(t).$$

Hence we have as before

$$1 = \sum_j r_{ij}(t) \geq \sum_k r_{ik}(0) = 1.$$

Thus we must have equality in (38) for every j .

The last statement follows from the Complement to Theorem 5, since (35) is the same as $\sum_{j \neq i} q_{ij} = q_i$.

REFERENCES

1. D. G. Austin, *On the existence of derivatives of Markoff transition probability functions*, Proc. Nat. Acad. Sci. U.S.A. vol. 41 (1955) pp. 224–226.
2. K. L. Chung, *Foundations of the theory of continuous parameter Markov chains*, to appear in the Third Berkeley Symposium.
3. J. L. Doob, *Topics in the theory of Markoff chains*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 37–64.
4. ———, *Markoff chains—denumerable case*, Trans. Amer. Math. Soc. vol. 58 (1945) pp. 455–473.
5. ———, *Renewal theory from the point of view of the theory of probability*, Trans. Amer. Math. Soc. vol. 63 (1948) pp. 422–438.
6. ———, *Stochastic processes*, New York, 1953.
7. A. N. Kolmogorov, *On the differentiability of the transition probabilities in stationary Markov processes with a denumerable number of states*, Uchenye Zapiski Moskovskogo Gosudarstvennogo Universiteta. Matematika vol. 4 (1951) pp. 53–59 (Russian).
8. P. Lévy, *Systèmes Markoviens et stationnaires. Cas dénombrable*, Ann. École Norm. vol. 68 (1951) pp. 327–381.