ON THE REGULARITY OF MARKOV PROCESSES[1]

BY
FRANK KNIGHT

Introduction

This study is concerned with continuous parameter Markov processes having values in an arbitrary space. More specifically, we shall consider the effect of stopping times on such processes. We first define these objects following [1].

DEFINITION 1. Let X be a space, and let $\mathcal{B}$ be a σ-field of subsets of X. Let $X(t)$ be a stochastic process, defined either for $t \geqq 0$ or for $-\infty < t < \infty$ on a probability space (Ω, F, P), with X as state space and $\mathcal{B}$ as measurable field. Finally, let $\{F(t)\}$ be a family of σ-subfields of F, defined for the same range of t as $X(t)$, such that $F(t_1) \subset F(t_2)$ for $t_1 < t_2$. Then $X(t)$ is a *Markov process relative to the family* $\{F(t)\}$ if (a) for each t and $E \in \mathcal{B}$,

$$\{X(t) \in E\} \in F(t),$$

and (b) for $t_1 < t_2$ and $E \in \mathcal{B}$,

$$P(\{X(t_2) \in E\} | F(t_1)) = P(\{X(t_2) \in E\} | X(t_1)) \quad \text{a.s.}$$

(a.s. abbreviates "almost surely" or "with P-measure 1").

DEFINITION 2. A random variable T on Ω with values in

$$\{\infty\} \cup \{t: X(t) \text{ is defined}\}$$

is a *stopping time in the general sense* for $X(t)$ if $P\{T < \infty\} > 0$ and for each t one has $\{T < t\} \in F(t)$. A *stopping time in the narrow sense* is defined by replacing $\{T < t\}$ by $\{T \leqq t\}$ in the above.

We note first that every stopping time in the narrow sense is also a stopping time in the general sense. Henceforward, *stopping time* will be used to mean stopping time in the general sense.

Along with any stopping time T for $X(t)$ we consider the new probability space derived from (Ω, F, P) by restriction to the set $\{T < \infty\}$.

DEFINITION 3. Let $X_T(t)$ be the process $X(t)$ restricted to $\Omega \cap \{T < \infty\}$, with field composed of the sets $S \cap \{T < \infty\}$, $S \in F$, and probability measure $P(S \cap \{T < \infty\})/P\{T < \infty\}$ for each such set. Let (Ω_T, F_T, P_T) designate this probability space, and let T_T be the restriction of T to Ω_T. For brevity we will omit the subscript in T_T.

Received July 11, 1960; received in revised form April 6, 1961.
[1] This work was supported by an Air Force contract.

It is easily seen that $X_T(t)$ is a Markov process relative to the fields $F_T(t)$ generated by sets $S \cap \{T < \infty\}$, $S \in F(t)$, and that T is a stopping time for $X_T(t)$.

The central regularity property to be considered is the strong Markov property, which we next define. The definition refers to some specified set of conditional probabilities $P(\{X(t + t_1) \in E\} \mid X(t))$, $t_1 > 0$, $E \in \mathcal{B}$, for the process $X(t)$, and to specified families of σ-fields $\{P_T(t), 0 \leqq t < \infty\}$ defining for each stopping time T and $t \geqq 0$ the *past of* $X_T(T + t)$ *up to time* t. Given these conditional probabilities and σ-fields, the *strong Markov property* is the requirement that for every stopping time T:

(i) the process $X_T(T + t)$, $0 \leqq t < \infty$, is again a Markov process on (Ω_T, F_T, P_T), relative to the fields $\{P_T(t)\}$,

(ii) given T and $X_T(T)$ this process is independent of $P_T(0)$, and

(iii) $P(\{X_T(T + t_1) \in E\} \mid X_T(T))$ defines a conditional probability on (Ω_T, F_T, P_T) of $\{X_T(T + t_1) \in E\}$ given $X_T(T)$ for each $t_1 > 0$ and $E \in \mathcal{B}$.

We shall assume at the start that the probabilities

$$P(\{X(t + t_1) \in E\} \mid X(t))$$

are produced by a transition probability function. This is defined to mean that for all t, t_1, and E,

$$P(\{X(t + t_1) \in E\} \mid X(t)) = p(t, X(t), t + t_1, E) \tag{1.1}$$

where $p(t_1, x, t_2, E)$ is a probability measure on $\mathcal{B}$ for fixed (t_1, x, t_2), and a function jointly measurable in (t_1, x, t_2) over $\mathcal{R} \times \mathcal{B} \times \mathcal{R}$ ($\mathcal{R}$ being the field of real Borel sets) for fixed E. The fields $P_T(t)$ are defined following these introductory remarks.

It is well known that not every Markov process satisfies the requirements of the strong Markov property [1]. Somewhat related to this difficulty is the problem of "separability" for a process, which begins with the fact that because of the inadequacy of joint distributions to define a process beyond determination up to sets of probability zero for each t, various intuitively meaningful functions of a process may fail to be measurable, or even to be well defined. A third regularity requirement, referring to the case of processes with transition probabilities stationary in time, is that of convergence of the associated semigroups to the identity at $t = 0$.

The present paper provides, under certain restrictions, a kind of simultaneous resolution to these problems. Starting with a given process $X(t)$ on (Ω, F, P) a related process $Y(t)$ is defined on the same space, such that $Y(t)$ is similar enough to $X(t)$ to replace it for most purposes, but such that $Y(t)$ satisfies, in a sense to be made precise, the above three regularity requirements. The process $Y(t)$ is closely related to the processes defined in [7], and also to those defined in some unpublished work of D. G. Austin. A main difference from [7], however, is that in [7] the starting point is a resolvent family, while here it is a probability space (Ω, F, P) and process $X(t)$. Thus, in the present paper, topologies on X and function space methods are largely absent. For a given stopping time T, we first change the field generated by T and $X_T(T)$ sufficiently to yield automatic conditional independence of the past and $X_T(T + t)$ relative to the changed field. This changed field, called here an entrance boundary field, is generated by T together with an infinitesimally small part of the future of $X_T(t)$ after time $t = T$. Given this new field, the transition probabilities of $X_T(t)$ are expressed as limits of those of $X(t)$, using martingale convergence. The stopping time T is seen to be a "Markov time" for $X_T(t)$ (i. e., T meets the requirements (i), (ii), and (iii) of the strong Markov property) if these transition probabilities are a.s. the same as the given ones. In the other case, and when the transition probabilities are stationary in time and $\mathcal{B}$ satisfies a mild restriction, the process is redefined slightly with the result that for the new process and for every stopping time these limits are a.s. the same as the conditional probabilities defined for the new process. This redefined process is then a "strong Markov process." It is also "separable" in the sense (different from that of [2]) that its paths are determined by their values for a countable dense set of times. Finally, the way in which the new process may replace the original process is revealed in the fact that functionals whose definition does not depend on countable sets of t are naturally redefined on the new process, and their distributions remain unchanged. We note that most functionals can be defined in this way—for example, the value of a process at a fixed time can ordinarily be defined as the limit of an integral average on the path functions as the interval of integration decreases to zero.

Section 1

Commencing the exposition, we first introduce the following σ-fields:

Definition 4. Let $H^+(T)$ be the field of all sets $S \in F_T$ such that

$$S \cap \{T < c\} \in F_T(c)$$

for all c. $H^+(T)$ may be termed the *field of the past up to time* T^+. Moreover, let $H(T)$ be the field consisting of all sets $S \in F_T$ such that

$$S \cap \{T \leqq c\} \in F_T(c)$$

for all c. Then $H(T)$ is the field of the *past up to time* T. It is clear that T is measurable over $H^+(T)$ and over $H(T)$.

Definition 5. The process $X(t)$ is a *strong Markov process* if for each stopping time T it satisfies the requirements (i), (ii), and (iii) of the strong Markov property, with $P_T(t) = H^+(T + t)$. A process $X(T)$ is a *strong Markov process in the narrow sense* if for each stopping time T in the narrow sense it satisfies (i), (ii), and (iii), with $P_T(t) = H(T + t)$.

Definition 6. Let L be the least σ-field containing all sets of the form $\{X(t) \in E\}$, $E \in \mathcal{B}$, all t. L may be termed the *field generated by* $X(t)$. We denote by $L(T)$ the restriction of L to Ω_T, i.e., the σ-field of sets $S \cap \Omega_T$ for $S \in L$.

DEFINITION 7. For each integer $n > 0$ and $-\infty < m < \infty$ we set

$$\alpha(m, n) = m2^{-n}.$$

For any stopping time T, let $T_n(T) = T_n$ be the stopping time defined by $T_n = \alpha(m, n)$ for $\alpha(m-1, n) \leqq T < \alpha(m, n)$, $-\infty < m < \infty$. We denote by $G_n(T)$ the least σ-field containing the sets $\{T_k \in R\}$ and $\{X(T_k) \in E\}$ for $k \geqq n$, $E \in \mathcal{B}$, and $R \in \mathcal{R}$, where $\mathcal{R}$ designates the real Borel sets. Finally, let the *entrance boundary field* $G^+(T)$ be defined by $G^+(T) = \bigcap_{n=1}^{\infty} G_n(T)$.

It is convenient to prove, at this point, that $G^+(T) \subset H^+(T)$. Let $S \in G^+(T)$. Then $S \cap \{T < c\} = \bigcup_{n=1}^{\infty} S \cap \{T_n < c\}$, and it is sufficient to prove that $S \cap \{T_n < c\} \in F_T(c)$. Since $G^+(T) \subset G_n(T)$, this will certainly hold if it holds for the sets $S' = \{T_k \in R\}$ and $S' = \{X(T_k) \in E\}$, $k \geqq n$, which generate $G_n(T)$. But $\{T_k \in R\} \cap \{T_n < c\} \in F_T(c)$ is immediate, while

$$\{X(T_k) \in E\} \cap \{T_n < c\} = \bigcup_{m=-\infty}^{\gamma} (\{T_k = \alpha(m, k)\} \cap \{X(\alpha(m, k)) \in E\} \cap \{T_m < c\})$$

where $\gamma = [2^k c]$ (the greatest integer $\leqq 2^k c$), and this is a union of sets in $F_T(c)$.

We shall now establish an extended form of the strong Markov property.

THEOREM 1. *Let $X(t)$ be a Markov process relative to a family of σ-fields $\{F(t)\}$ and such that $X(t)$ is measurable over $\mathcal{R} \times L$ as a function of (t, w) (see ff. (1.1) and Definition 6). Then if T is any stopping time for $X(t)$, $X(T + t)$ is measurable for each $t \geqq 0$, and for $0 < t_1 < \cdots < t_k$, $k > 0$, and $E_1, \cdots, E_k \in \mathcal{B}$,*

$$P_T\left(\bigcap_{i=1}^{k} \{X_T(T + t_i) \in E_i\} \mid H^+(T)\right) = P_T\left(\bigcap_{i=1}^{k} \{X_T(T + t_i) \in E_i\} \mid G^+(T)\right)$$
$$= \lim_{n\to\infty} \int_{E_1} \cdots \int_{E_{k-1}} p(T + t_{k-1}, x_{k-1}, T + t_k, E_k) \cdot p(T + t_{k-2}, x_{k-2}, T + t_{k-1}, dx_{k-1}) \cdots p(T_n, X(T_n), T + t_1, dx_1) \quad a.s.,$$

where for $k = 1$ the last expression is to be replaced by

$$\lim_{n\to\infty} p(T_n, X(T_n), T + t_1, E_1).$$

The proof of Theorem 1 will be carried out for the case $k = 1$ only, since the other cases are proved analogously.

LEMMA 1.1. *For each $n > 0$, $t > \alpha(1, n)$, and $E \in \mathcal{B}$,*

$$P_T(\{X_T(T + t) \in E\} \mid H(T_n)) = P_T(\{X_T(T + t) \in E\} \mid G_n(T)) = p(T_n, X_T(T_n), T + t, E) \quad a.s. \tag{1.2}$$

Proof. Since T_n is countably valued [1] it follows from the definition of $p(t_1, x_1, t_2, E)$ that for each assembly $0 < t_1 < \cdots < t_j$ and $E_1, \cdots, E_j \in \mathcal{B}$, a.s. on the set $\{T_n < t_1\}$,

$$P_T(\cap_{i=1}^{j} \{X_T(t_i) \in E_i\} \mid H(T_n))$$
$$= P_T(\cap_{i=1}^{j} \{X_T(t_i) \in E_i\} \mid G_n(T))$$
$$(1.3) \qquad = \int_{E_1} \cdots \int_{E_{j-1}} p(t_{j-1}, x_{j-1}, t_j, E_j) p(t_{j-2}, x_{j-2}, t_{j-1}, dx_{j-1})$$
$$\cdots p(T_n, X_T(T_n), t_1, dx_1),$$

where in the case $j = 1$ the last expression becomes $p(T_n, X_T(T_n), t_1, E_1)$. We consider this case first, and set $S_m = \{T_n = \alpha(m, n)\}$. For $S \in H(T_n)$ it follows from the representation

$$S \cap S_m = S \cap \{T_n \leqq \alpha(m, n)\} \cap \{\Omega_T - (S \cap \{T_n < \alpha(m-1, n)\})\}$$

that $S \cap S_m \in F(\alpha(m, n))$. Therefore we have for $\alpha(m, n) < t_1$

$$(P\{T < \infty\})^{-1} \int_{S \cap S_m} p(T_n, X_T(T_n), t_1, E_1)\, dP(w)$$
$$(1.4) \qquad = (P\{T < \infty\})^{-1} P(S \cap S_m \cap \{X_T(t_1) \in E_1\})$$
$$= \int_{S \cap S_m} p(T_n, X_T(T_n), t_1, E_1)\, dP_T(w)$$
$$= P_T(S \cap S_m \cap \{X_T(t_1) \in E_1\}).$$

For $j = 1$, (1.3) follows from (1.4) by summing over $\{m : \alpha(m, n) < t_1\}$ and noting that $G_n(T) \subset H(T_n)$. The situation for $j > 1$ is clearly analogous.

To derive the lemma from (1.3) it is necessary to replace the constant time t for $j = 1$ by the random time[2] $T + t$. We first remark that $T + t$ is measurable over $H(T_n)$. To justify this replacement we reduce the problem to one involving a product space. Let (Ω'_T, F'_T, P'_T) and $(\Omega''_T, F''_T, P''_T)$ be two identical replicas of (Ω_T, F_T, P_T). Letting primes indicate the replica in which a point, set, σ-field, or random variable is considered, we use (1.3) to define a measure μ on the space $(\Omega'_{T'} \times \Omega''_{T''}, H'(T'_n) \times L''(T''))$ such that $\mu(S'_1 \times S''_2) = P_T(S_1 \cap S_2)$ where S_i is the copy of S'_i or S''_i in Ω_T, $i = 1, 2$. For sets of the form $S_2 = \cap_{i=1}^{j} \{X_T(t_i) \in E_i\}$ and $S_1 \in H(T_n)$ we define

$$\mu(S'_1 \times S''_2) = P_T(S_1 \cap S_2) = P_T(S_1 \cap S_2 \cap \{t_1 \leqq T_n\})$$
$$(1.5) \qquad + \int_{S_1 \cap \{T_n < t_1\}} \int_{E_1} \cdots \int_{E_{j-1}} p(t_{j-1}, x_{j-1}, t_j, E_j)$$
$$\cdot p(t_{j-2}, x_{j-2}, t_{j-1}, dx_{j-1}) \cdots p(T_n, X_T(T_n), t_1, E_1)\, dP_T(w).$$

[2] This situation, and also the proof, closely resemble those of Lemmas 1, 2 of [9] (as the referee has pointed out).

For the remaining sets of $H'(T'_n) \times L''(T'')$ we define μ as the unique σ-additive extension.

Along with the measure μ we introduce the set mapping M for which $M(S'_1 \times S''_2) = S_1 \cap S_2$. This mapping has a natural extension to $H'(T'_n) \times L''(T'')$. Indeed, let the *diagonal* of any set $S \in H'(T'_n) \times L''(T'')$ be defined as the set of all $w \in \Omega$ such that $(w', w'') \in S$. Then we define

$$(1.6) \qquad M(S) = \text{diagonal of } S.$$

It is clear that M preserves set unions and differences. Let $F(H(T_n), L(T))$ be the least σ-field containing $H(T_n)$ and $L(T)$. It is easily seen that the class of sets S for which $M(S) \in F(H(T_n), L(T))$ contains the ring of sets $S' \times S''$ for $S' \in H'(T'_n)$ and $S'' \in L''(T'')$, and that it is closed under monotone limits. It is thus equal to $H'(T'_n) \times L''(T'')$. Similarly, it follows from (1.5) that

$$(1.7) \qquad \mu(S) = P_T\, M(S).$$

Now let $t^* > 0$ be a real-valued random variable on $H'(T'_n)$ such that

$$\{X''_T(T'_n + t^*) \in E\} \in H'(T'_n) \times L''(T''), \qquad \text{for all } E \in \mathcal{B}.$$

The "section" of $\{X''_T(T'_n + t^*) \in E\}$ in $L''(T'')$ at any point $w' \in \Omega'_T$ for which $t^*(w') = t$ is clearly $\{X''_T(T'_n(w') + t) \in E\}$, and it follows from (1.5) and a standard result for iterated integrals in product spaces that for $S' \in H'(T'_n)$

$$(1.8) \qquad \begin{aligned} &\mu(\{X''_T(T'_n + t^*) \in E\} \cap (S' \times \Omega''_T)) \\ &\qquad = \int_{S'} p(T'_n, X'_T(T'_n), T'_n + t^*, E)\, dP'_T(w'), \end{aligned}$$

the integrand being measurable over $H'(T'_n)$ [9, Lemma 2]. In particular, let $t^* = T' - T'_n + t$. Then if $\{X''_T(T' + t) \in E\} \in H'(T'_n) \times L''(T'')$, it follows from (1.8), after applying the mapping M of (1.6), that for $S \in H(T_n)$

$$(1.9) \qquad P_T(\{X_T(T + t) \in E\} \cap S) = \int_S p(T_n, X_T(T_n), T + t, E)\, dP_T(w),$$

which is equivalent to (1.2).

In order that $\{X''_T(T' + t) \in E\} \in H'(T'_n) \times L''(T'')$ it is sufficient that $X(t)$ be measurable over $\mathcal{R} \times L$, as assumed in the theorem. For we then have $\{(t'', w''): X''_T(t'') \in E\} \in \mathcal{R}'' \times L''(T'')$, and since $T' + t$ is measurable from $H'(T'_n)$ to $\mathcal{R}$ it follows that

$$\{(w', w''): X''_T(T' + t) \in E\} \in H'(T'_n) \times L''(T'').$$

The lemma is thereby proved.

To complete the proof of Theorem 1, we use the fact that the sequence of fields $\{G_n(T)\}$ is nonincreasing. Hence, as is well known,

$$P_T(\{X_T(T + t) \in E\} \mid G_n(T))$$

is a (reversed) martingale in n. By a theorem of Doob it converges a.s. to $P_T(\{X_T(T + t) \in E\} \mid G^+(T))$, implying the theorem for $k = 1$. The situation for $k > 1$ is analogous.

Without some further restrictions it seems difficult to show that the limit in n of Theorem 1 may be brought inside the integral. We shall adopt the following hypothesis:

HYPOTHESIS $*$.[3] (a) The field $\mathcal{B}$ is generated by countably many sets, and wide-sense conditional distributions over $\mathcal{B}$ exist [2]. (b) The transition probabilities $p(t_1, x, t_2, E)$ are stationary in time, i. e.,

$$p(t_1, x, t_2, E) = p(t_2 - t_1, x, E)$$

where $p(t, x, E)$ is jointly measurable in t and x for each $E \in \mathcal{B}$.

Under this hypothesis Theorem 1 assumes the simpler form stated in the following definitions and corollaries.

DEFINITION 7$*$ (replacing Definition 7 under the assumption of Hypothesis $*$). Let $G_n^*(T)$ be the least σ-field containing the sets $\{X_T(T_k) \in E\}$ and $\{T_k - T \in R\}$ for $k \geqq n$, $R \in \mathcal{R}$, and $E \in \mathcal{B}$. Further, let

$$G^{*+}(T) = \bigcap_{n=1}^{\infty} G_n^*(T).$$

COROLLARY 1. *Under Hypothesis $*$, Theorem 1 becomes*

$$\begin{aligned} P_T(\textstyle\bigcap_{i=1}^k \{X_T(T + t_i) \in E_i\} \mid H^+(T)) &= P_T(\textstyle\bigcap_{i=1}^k \{X_T(T + t_i) \in E_i\} \mid G^{*+}(T)) \\ &= \int_{E_1} \cdots \int_{E_{k-1}} p(t_k - t_{k-1}, x_{k-1}, E_k) p(t_{k-1} - t_{k-2}, x_{k-2}, dx_{k-1}) \\ &\qquad \cdots \lim_{n\to\infty} p(t_1 + T - T_n, X_T(T_n), dx_1) \quad a.s. \end{aligned}$$

where $\lim_{n\to\infty} p(t_1 + T - T_n, X_T(T_n), E_1)$ is any wide-sense conditional distribution in E_1 for $X_T(T + t_1)$ given $G^{*+}(T)$, is equal to the indicated limit a.s. for each $E_1 \in \mathcal{B}$, and defines the last expression of Corollary 1 in the case $k = 1$.

We note first that if we replace $G^{*+}(T)$ by $G^+(T)$, then Corollary 1 is a case of Theorem 1. Indeed, for any wide-sense conditional distribution $P(\{X_T(T + t_1) \in (\cdot)\} \mid G^+(T))$ and for any countable field of sets $\{E_i\}$ generating $\mathcal{B}$, we have

$$\lim_{n\to\infty} p(t_1 + T - T_n, X_T(T_n), E_i) = P(\{X_T(T + t_1) \in E_i\} \mid G^+(T))$$

for all E_i a.s. But the integral in Corollary 1 may be defined using only

[3] Wide-sense conditional distributions over $\mathcal{B}$ exist if (for example) $\mathcal{B}$ is generated by the compact sets of a locally compact space X [3].

countably many x_1-sets, for example the sets

$$E(\alpha(m,n)) = E_1 \cap \left\{ x_1 : \int_{E_2} \cdots \int_{E_{k-1}} p(t_k - t_{k-1}, x_{k-1}, E_k) \cdots p(t_2 - t_1, x_1, dx_2) < \alpha(m,n) \right\}$$

and their differences. Choosing these sets among the $\{E_i\}$ we see that a.s. the limit may be brought inside the integral as required. But since the integral in Corollary 1 is already measurable over the field $G^{*+}(T)$ of Definition 7$*$, the conditioning may be limited to this subfield.

DEFINITION 8. For each $t_1 > 0$, let $A(T + t_1)$ be the least σ-field generated by $H^+(T)$ together with the sets $\{X_T(T + t) \in E\}$, $0 \leqq t \leqq t_1$. Thus $A(T + t_1)$ is a *past of* $X_T(T + t)$ *up to time* t_1. We note that when $X(t)$ is measurable over $\mathcal{R} \times L$ we have $A(T + t_1) \subset F_T$.

COROLLARY 2. *For the situation of Corollary* 1, *if* $t_1 > 0$,

$$P_T(\{X_T(T + t_2) \in E_2\} \mid A(T + t_1)) = p(t_2 - t_1, X_T(T + t_1), E_2).$$

Stated in words, $X_T(T + t_1)$ *is a Markov process with transition probabilities* $p(t, x, E)$ *for* $t_1 > 0$, *relative to the fields* $A(T + t_1)$.

This corollary is an immediate consequence of Corollary 1 for $k = 2$.

COROLLARY 3.[4] *If there is a topology on* X *such that* $p(t, x, E)$, *for each* E *in some countable field generating* $\mathcal{B}$, *is jointly continuous in* x *and left continuous in* t, *and for which* $X(t)$ *has right continuous path functions, then* $X(t)$ *is a strong Markov process.*

The proof is again direct from Corollary 1, since the hypotheses imply that

$$\lim_{n\to\infty} X_T(T_n) = X_T(T) \quad \text{a.s.},$$

and that

$$\lim_{n\to\infty} p(t + T - T_n, X_T(T_n), E) = p(t, X_T(T), E) \quad \text{a.s.}$$

for all E in a countable field generating $\mathcal{B}$. Since the measure $p(t, X_T(t), (\cdot))$ is countably additive, it is identical a.s. with any wide-sense conditional

[4] With further assumptions on the topology and on $\mathcal{B}$, left continuity in t is a consequence of continuity in x. Thus for X a compact space and $\mathcal{B}$ the topological Borel field, this corollary reduces to the sufficiency of right continuous paths and semigroup mapping the continuous functions into themselves [4]. Indeed, the semigroup is then weakly right continuous, and this is known to imply strong continuity for $t > 0$ [6, p. 306]. Hence the distributions $p(t + T - T_n, X_T(T_n), (\cdot))$ converge weakly to $p(t, X_T(T), (\cdot))$ a.s. as $n \to \infty$, which is sufficient for the proof to work. Our point of view, however, is that even the assumptions of Corollary 3 are unnaturally topological. Thus any assumption is sufficient which implies that $X(T + t)$ is measurable and that $\lim_{n\to\infty} p(t + T - T_n, X_T(T_n), E) = p(t, X_T(T), E)$ a.s. It is hoped that Section 2 of this paper provides evidence that these are natural requirements to impose on any Markov process.

distribution $P(\{X_T(T + t) \in (\cdot)\} | G^{*+}(T))$, and therefore is itself such a distribution.

For applications to analysis it would be useful to discover how the past up to time T is connected with the field $G^+(T)$. In particular, let $G^-(T)$ be the field defined like $G^+(T)$ but after replacing the stopping times T_n by their counterparts T_n^- "toward the past":

$$T_n^- = \alpha(m - 1, n) \qquad \text{for } \alpha(m - 1, n) < T \leqq \alpha(m, n).$$

It is suggested by Theorem 1 that the field $G^-(T)$ may there suffice for the same purpose as $G^+(T)$, and therefore that the past up to time T may be transmitted to $G^+(T)$ directly from $G^-(T)$. This is not true in general, however, even for stopping times in the narrow sense, as is shown by the following counterexample:

Example 1. Let $X = \{0, \pm i{:}1 \leqq i \leqq 4\}$ with $\mathcal{B}$ the field of all subsets of X. It is known that a process $X(t)$, $t \geqq 0$, with $X(0) = 0$, is defined by assigning to each point $x \in X$ an exponential waiting time parameter $\lambda(x)$, $0 \leqq \lambda(x) \leqq \infty$, and a probability measure $p(x, S)$ on $\mathcal{B}$ for the state first visited after leaving x, provided that there is at most one occurrence in each path of an x such that $\lambda(x) = \infty$ (instantaneous state). We begin with the following assignments:

$$\lambda(0) = \lambda(1) = \lambda(-1) = \lambda(2) = \lambda(-2) = 1,$$

$$\lambda(3) = \lambda(-3) = \infty, \qquad \lambda(4) = \lambda(-4) = 0;$$

$$p(0, 1) = p(0, -1) = \tfrac{1}{2},$$

$$p(1, 2) = p(1, -2) = p(1, 3) = p(1, -3) = \tfrac{1}{4},$$

$$p(-1, 2) = p(-1, -2) = p(-1, 3) = p(-1, -3) = \tfrac{1}{4},$$

$$p(2, 1) = p(2, -1) = p(2, 3) = p(2, -3) = \tfrac{1}{4},$$

$$p(-2, 1) = p(-2, -1) = p(-2, 3) = p(-2, -3) = \tfrac{1}{4}.$$

This is sufficient to define a process up to the first arrival at one of the instantaneous states $+3$ or -3. Let us define the distribution of the next state visited after $+3$ or -3 to depend on the past, as follows: from $+3$ the next state visited is $+4$ if the second state visited was $+1$, and -4 if the second state visited was -1; from -3 the next state visited is $+4$ if the second state visited was -1, and -4 if the second state visited was $+1$. Since $+4$ and -4 are absorbing states, this completes the definition of a process $X(t)$, $0 \leqq t < \infty$. Moreover, $X(t)$ is a Markov process. To convince ourselves of this it is sufficient to note that at any time prior to an arrival in the set $\{+3, -3\}$ the conditional probability of arrival at $+3$ before -3 is $\frac{1}{2}$, independently of the past, and hence so also is the conditional probability of eventual arrival at $+4$ before -4. Now let T be the time of arrival in the set $\{+3, -3\}$. Since

$$\{T \leqq t\} = \{T < t\} \cup \{T = t\} = \{T < t\} \cup \{X(t) = 3\} \cup \{X(t) = -3\},$$

we see that $\{T \leqq t\} \in F(t)$ where $F(t)$ is generated by the sets

$$\{X(t') \in S\}, \qquad t' \leqq t, \quad S \in \mathcal{B},$$

so that T is a stopping time in the narrow sense. Moreover, it is easy to see that $X(t)$ is measurable over $(\mathcal{R} \times L(T))$. However, the past and future are clearly not independent given $G^{-}(T)$, since $G^{-}(T)$ does not contain the set where $+1$ is reached before -1.

Section 2

The next aim is to extend the space X of Theorem 1 and to redefine the process $X(t)$ in such a way that, although differing as little as possible from $X(t)$, the redefined process will be a strong Markov process on the enlarged space. It is to be noted that this may require redefinition at certain t on a set of positive probability, as shown by Example 2 (also [8]).

Example 2. Let $X(t)$, $0 \leqq t < \infty$, have only two path functions, each occurring with probability $\frac{1}{2}$. The first is given by $X(t) = t$, $0 \leqq t \leqq 1$, and $X(t) = 2$, $1 < t < \infty$. The second is given by $X(t) = t$, $0 \leqq t \leqq 1$, and $X(t) = 3$, $1 < t < \infty$. Obviously, $T = \text{g.l.b.}\{t: X(t) = 3\}$ is a stopping time (in the general sense), and $X(T) = 1$ if $T < \infty$. However,

$$P_T(\{X_T(T + t) = 3\} \mid X(T)) = 1$$

while $p(t, X(T), \{3\}) = \frac{1}{2}$. Thus T is not a Markov time, and any redefinition must involve a set of at least probability $\frac{1}{2}$.

This example depends on the fact that T is a stopping time in the general sense but not in the narrow sense. For narrow-sense stopping times to be Markov times relative to the fields $P_T(t) = H(T + t)$ (Definition 5) it is never necessary to redefine the process at any t on a set of probability greater than zero. Indeed, if $\{t_i\}$ is the (countable) set of all t_i for which

$$P\{T = t_i\} > 0,$$

then for each i we have $\{T = t_i\} \in F(t_i)$, and hence by the Markov property, a.s. on the set $\{T = t_i\}$,

$$\begin{aligned} P(\{X(T + t') \in E\} \mid H(T)) &= P(\{X(T + t') \in E\} \mid F(t_i)) \\ &= p(t', X(t_i), E). \end{aligned} \tag{2.1}$$

It follows that T is a Markov time on $\bigcup_i \{T = t_i\}$, and thus no redefinition on this set is necessary. It will be shown below, moreover, that under Hypothesis * the process may be redefined to become a strong Markov process in the narrow sense, without changing the definition at any t on a set of positive probability.

The basis of the redefinition of $X(t)$ to produce a strong Markov process $Y(t)$ is the construction of elements in the state space Y of $Y(t)$. These will be certain families of joint distributions defining Markov processes for $t > 0$ with the original transition probabilities. The construction is carried out in the form of a lemma.

LEMMA 2.1. *Let $X(t)$ be a Markov process relative to fields $F(t)$ with F and each $F(t)$ assumed to be completed for the measure P, and suppose that Hypothesis $*$ holds. Then we have a.s.*

$$\begin{aligned} &P(\{X(t_r(t)) \in E_i\} \mid H^+(t)) \\ (2.2) \qquad &= P(\{X(t_r(t)) \in E_i\} \mid G_c^{*+}(t)) \\ &= \lim_{\tau \downarrow 0, t+\tau \in A} p(t_r(t) - t - \tau, X(t+\tau), E_i) \end{aligned}$$

where $\{E_i\}$ is any field of sets forming a countable basis of $\mathcal{B}$, $A = \{\alpha(m, n)\}$, $H^+(t)$ and $G_c^{*+}(t)$ are the σ-fields $H^+(T)$ and $G^{*+}(T)$ for $T = t$, the latter completed for the measure P, and $t_r(t)$ is any element of the decreasing sequence $\{t_r(t)\}$ with limit t, defined below. Moreover, the limit (2.2) exists a.s. for all t, r, and i simultaneously, defines a conditional probability as indicated, and defines by extension for each t and r a.s. a probability measure on $\mathcal{B}$. These measures satisfy a.s. for all $r_1 < r_2 < \infty$

$$\begin{aligned} &P(\{X(t_{r_1}(t)) \in E\} \mid G_c^{*+}(t)) \\ (2.3) \qquad &= \int_X p(t_{r_1}(t) - t_{r_2}(t), x, E) P(\{X(t_{r_2}(t)) \in dx\} \mid G_c^{*+}(t)). \end{aligned}$$

Wherever (2.2) and its extension exist and (2.3) holds, we define for all $t_1 > 0$ the distributions

$$(2.4) \qquad F(t_1, E; t, w) = \int_X p(t - t_r(t) + t_1, x, E) P(\{X(t_r(t)) \in dx\} \mid G_c^{*+}(t)),$$

where r is any integer for which $t - t_r(t) + t_1 > 0$. For the remaining set of probability 0 at each t we choose an arbitrary but fixed element $\underset{\sim}{x} \in X$ and define

$$(2.5) \qquad F(t_1, E; t, w) = p(t_1, \underset{\sim}{x}, E).$$

Then $F(t_1, (\cdot); t, w)$ is a wide-sense conditional distribution

$$P(\{X(t + t_1) \in (\cdot)\} \mid H^+(t)) = P(\{X(t + t_1) \in (\cdot)\} \mid G_c^{*+}(t))$$

on $\mathcal{B}$, and for $t_1 < t_2$

$$(2.6) \qquad F(t_2, E; t, w) = \int_X p(t_2 - t_1, x, E) F(t_1, dx; t, w).$$

Moreover, let T_1 and T_2 be stopping times with $T_1 \leqq T_2$, and let $E(T_1)$ be a random set (a function of T_1) such that

$$\{X''(t_r(T_1') + t_1) \in E(T_1')\} \in H^{+\prime}(T_2') \times L''(T_2''),$$

where primes indicate replicas in a product space

$$(\Omega'_{T_2} \times \Omega''_{T_2}, H^{+\prime}(T'_2) \times L''(T''_2))$$

as in the proof of Theorem 1. *Then for r and* t_1 *such that, for all* $w \in \Omega_{T_2}$, $t_1 > -(t_r(T_1) - T_2)$, *we have a.s.*

$$F(t_1 + t_r(T_1) - T_2, E(T_1); T_2, w) = P(\{X(t_1 + t_r(T_1)) \in E(T_1)\} \mid H^+(T_2)). \tag{2.7}$$

Proof. We first have to define the functions $t_r(t)$. Consider for each t and $k > 0$ the unique integer $m_k(t)$ such that

$$\alpha(m_k(t), k) \leqq t < \alpha(m_k(t) + 1, k).$$

The sequence $\{t_r(t)\}$ is merely the sequence $\{\alpha(m_k(t) + 1, k)\}$ with repetitions (for different k) eliminated. Thus $\{t_r(t)\}$ is strictly decreasing with limit t.

The expressions $p(2^{-k} - \tau, X(\alpha(m, k) + \tau), E_i)$ form martingales in τ. By a theorem of Doob [2, p. 363] the limits

$$\lim_{\tau \downarrow t-\alpha(m,k), \tau \in A} p(2^{-k} - \tau, X(\alpha(m, k) + \tau), E_i) = \lim_{\tau \downarrow 0, t+\tau \in A} p(\alpha(m + 1, k) - t - \tau, X(t + \tau), E_i) \tag{2.8}$$

exist a.s. for all t such that $\alpha(m, k) \leqq t < \alpha(m + 1, k)$. Thus the limits (2.2) exist a.s. for all t, and by [2, p. 331] they define for each t a conditional probability as required.

But if $\underline{P}(\{X(t_r(t)) \in (\cdot)\} \mid G_c^{*+}(t))$ is any wide-sense conditional distribution as indicated, then the limits (2.2) agree with it a.s. for all E_i, and thus a.s. define by extension a probability measure on $\mathcal{B}$ which is also for each E a conditional probability on Ω. The relations (2.3) and (2.6) are now evident consequences of the Chapman-Kolmogorov equation, and we will see by the same argument as leads from (2.11) to (2.9) below that $F(t_1, (\cdot); t, w)$ is indeed a wide-sense conditional distribution $P(\{X(t + t_1) \in (\cdot)\} | G_c^{*+}(t))$.

Turning now to (2.7) we first show that for E and t fixed, a.s. on the set $\{T_2 < t\}$

$$F(t - T_2, E; T_2, w) = \lim_{\tau \downarrow 0, T_2+\tau \in A} p(t - T_2 - \tau, X(T_2 + \tau), E). \tag{2.9}$$

Since $\{T_2 < t\} = \bigcup_r \{t_r(T_2) < t\}$, it is sufficient to show that (2.9) holds a.s. on a set $\{t_r(T_2) < t\}$. Here we have by (2.6)

$$F(t - T_2, E; T_2, w) = \int_X p(t - t_r(T_2), x, E) F(t_r(T_2) - T_2, dx; T_2, w). \tag{2.10}$$

Since $t - t_r(T_2)$ is countably valued, the integral in (2.10) depends only on countably many sets $\{E'_i\}$ which we shall take to include the $\{E_i\}$. It will be shown that a.s. for all of these sets

$$F(t_r(T_2) - T_2, E'_i; T_2, w) = \lim_{\tau \downarrow 0, T_2+\tau \in A} p(t_r(T_2) - T_2 - \tau, X(T_2 + \tau), E'_i), \tag{2.11}$$

i.e., that the definition of $F(t_r(T_2) - T_2, E_i'; T_2, w)$ is given a.s. by substituting T_2 for t and E_i' for E_i in (2.2). It is clear that the limit in (2.11) is actually a version of $P(\{X(t_r(T_2)) \in E_i'\} \mid G_c^{*+}(T_2))$, so that in particular it exists a.s. for all E_i', and by comparison with any wide-sense conditional distribution $P(\{X(t_r(T_2)) \in (\cdot)\} \mid G_c^{*+}(T_2))$ we see that (2.11) a.s. defines by extension a probability measure. The same facts hold a.s. for all pairs $r_1, r_2, r_1 < r_2$, which means that a.s. $F(t_r(T_2) - T_2, E_i'; T_2, w)$ is defined by (2.4) rather than by (2.5), proving (2.11). By inserting the limit in (2.11) into the integral of (2.10) it is clear that we may bring the limit outside of the integral, upon which (2.9) follows.

It is next to be shown that t in (2.9) may be replaced by $t_1 + t_r(T_1)$ and E by $E(T_1)$, which will prove (2.7). This situation is analogous to the earlier replacement of t_1 by $T + t$ in (1.4). We first note that on $\{T_2 < t_1\}$ we have

$$(2.12) \qquad F(t_1 + t_2 - T_2, E; T_2, w) = \int_X p(t_2, x, E) F(t_1 - T_2, dx; T_2, w).$$

Since, by (2.9), $F(t - t_2, E; T_2, w)$ is a version of $P_{T_2}(\{X(t) \in E\} \mid H^+(T_2))$ on $\{T_2 < t\}$, we can use $F(t - T_2, E; T_2, w)$ and (2.12) to generate a measure μ on a product space $H^{+\prime}(T_2') \times L''(T_2'')$ such that

$$(2.13) \qquad \begin{aligned} \mu(S' \times \{X''(t) \in E\}) = \int_{S \cap \{T_2 < t\}} & F(t - T_2, E; T_2, w)\, dP_{T_2}(w) \\ & + P_{T_2}(S \cap \{X(t) \in E\} \cap \{T_2 \geqq t\}) \end{aligned}$$

where S' is the replica of S in Ω', and $S \in H^+(T_2)$.

Since $S \cap \{X(t) \in E\} \cap \{T_2 \geqq t\} \in H^+(T_2)$, it follows that (2.13) defines, along with (2.12), a σ-additive set function on a field generating

$$H^{+\prime}(T_2') \times L''(T_2'').$$

We define μ on $H^{+\prime}(T_2') \times L''(T_2'')$ as the σ-additive extension of this function. Clearly, μ is a probability measure.

Recalling the mapping M of a set onto its diagonal (1.6), we also have for $K \in H^{+\prime}(T_2') \times L''(T_2'')$

$$(2.14) \qquad \mu(K) = P_{T_2}(M(K)).$$

Under the assumption that $\{X''(t_r(T_1') + t_1) \in E(T_1')\} \in H^{+\prime}(T_2') \times L''(T_2'')$ it follows that

$$(2.15) \qquad M\{X''(t_r(T_1') + t_1) \in E(T_1')\} = \{X(t_r(T_1) + t_1) \in E(T_1)\},$$

and that the right side is in the least σ-field containing $H^+(T_2)$ and $L(T_2)$. On the other hand, we have by (2.13), since $t_r(T_1) + t_1 - T_2 > 0$, that

$$(2.16) \qquad \begin{aligned} & \mu((S' \times \Omega_{T_2}'') \cap \{X''(t_r(T_1') + t_1) \in E(T_1')\}) \\ & \qquad = \int_S F(t_r(T_1) + t_1 - T_2, E(T_1); T_2, w)\, dP_{T_2}(w), \end{aligned}$$

as the section of $\{X''(t_r(T_1') + t_1) \in E(T_1')\}$ in Ω'' at $T_1' = t_2$, $T_2' = t_3$ has conditional measure $F(t_r(t_2) + t_1 - t_3, E(t_2); t_3, w)$. Thus by (2.14) and (2.15)

$$P_T(S \cap \{X(t_r(T_1) + t_1) \in E(T_1)\}) \tag{2.17}$$
$$= \int_S F(t_r(T_1) + t_1 - T_2, E(T_1); T_2, w)\, dP_{T_2}(w).$$

It follows from (2.9) and (2.16) that $F(t_r(T_1) + t_1 - T_2, E(T_1); T_2, w)$ is measurable over $H^+(T_2)$. Hence by (2.17) it is a version of

$$P(\{X(t_1 + t_r(T_1)) \in E(T_1)\} \mid H^+(T_2))$$

as asserted.

By means of (2.4), (2.5), and (2.6) there are defined functions $Y(t, w)$ whose values are the Markovian families of joint distributions determined by the expressions

$$\int_{E_1} \cdots \int_{E_{k-1}} p(t_k - t_{k-1}, x_{k-1}, E_k) p(t_{k-1} - t_{k-2}, x_{k-2}, dx_{k-1}) \cdots F(t_1, dx_1; t, w) \tag{2.18}$$

for all finite assemblies $0 < t_1 < \cdots < t_k$; $E_1, \cdots, E_k \in \mathcal{B}$. We will show that $Y(t, w)$ may be interpreted as a Markov process on (Ω, F, P), and that it possesses certain regularity properties. As above in the case of $X(t)$, we will write in place of $Y(t, w)$ simply $Y(t)$, the dependence on w being understood.

Definition 9. Let $\prod_{t>0} X_t$ be the product space of X indexed by the real parameter $t > 0$, and let $\prod_{\{\alpha(m,n)\}} \mathcal{B}_t$ be the least σ-field on $\prod_{t>0} X_t$ containing all sets of the form $(\prod_{0<t<\alpha(m,n)} X_t)(E_{\alpha(m,n)})(\prod_{\alpha(m,n)<t} X_t)$ for $E_{\alpha(m,n)} \in \mathcal{B}_{\alpha(m,n)}$, $0 < m < \infty, 0 < n < \infty$. Let Y be the space of all probability measures on $\prod_{\{\alpha(m,n)\}} \mathcal{B}_t$, and let $\mathcal{B}(Y)$ be the least σ-field on Y containing all sets of the form $\{y \in Y : P^y(Q) \in R\}$ where P^y is the measure defining the point y, $Q \in \prod_{\{\alpha(m,n)\}} \mathcal{B}_t$, and $R \in \mathcal{R}_{[0,1]}$ (the real Borel sets on $[0, 1]$). We define $Y(t)$ to be the function of (t, w) with values in Y determined, under Hypothesis $*$, by (2.18) for all assemblies $t_i = \alpha(m_i, n_i)$, $1 \leqq i \leqq k$. It is clear that this determination is unique.

Theorem 2. *$Y(t)$ is a strong Markov process on (Ω, F, P) with state space Y and $\mathcal{B}(Y)$ as measurable field, relative to the fields $H^+(t)$ and the (stationary in time) conditional probabilities* (2.19 a, b) *defined below.*

Remarks. It is easily seen that $F(t) \subset H^+(t)$ (see Definition 4). Also, the sets $\{Y(t') \in Q\}$ for $t' \leqq t$ are in $H^+(t)$, as noted below. Lastly, since the distributions determined by (2.18) all have the transition probabilities $p(t, x, E)$, it is seen that measurability of $\mathcal{B}(Y)$ implies measurability of $\{y \in Y : P^y(Q) \in R\}$ for any Q in $\prod_{0<t} \mathcal{B}_t$ if we extend the definition of Y and

$Y(t)$ to measures on $\prod_{0<t} \mathcal{B}_t$ by using (2.18) for arbitrary t_i in the latter case.

Proof. The measurability of $\mathcal{B}(Y)$ for $Y(t)$ over $H^+(t)$ follows from that of $\mathcal{R}_{[0,1]}$ for (2.18). This in turn is a consequence of the definition of $F(t_1, E; t, w)$ (see (2.4) and (2.2)) and the completeness of $H^+(t)$.

To define the conditional probabilities of $Y(t)$ it is expedient to introduce an auxiliary process $X^y(t)$, $0 < t < \infty$, with state space X and transition probabilities $p(t, x, E)$, corresponding to each point $y \in Y$ actually assumed by $Y(t)$ for some t. The process is to have as distributions those defining the point y through (2.18). We leave the sample space of $X^y(t)$ undefined for the present. Its existence is assured by the fact that $X^y(t)$ has transition probabilities. Now let

$$M(t_1, \cdots, t_k; E_1, \cdots, E_k; R)$$

for $0 < t_1 < \cdots < t_k$, $t_i = \alpha(m_i, n_i)$, $E_i \in \mathcal{B}$, $1 \leqq i \leqq k$, and $R \in \mathcal{R}_{[0,1]}$, designate $\{y \in Y: P^y(E_1(t_1) \times \cdots \times E_k(t_k)) \in R\}$, where $E_i(t_i) \in \mathcal{B}_{t_i}$ and the factors X_t have been omitted for the sake of brevity. We introduce the function $p_Y(t, y, M(t_1, \cdots, t_k; E_1, \cdots, E_k; R))$ defined for points y in the range of $Y(t)$ by

$$\begin{aligned} &p_Y(t, y, M(t_1, \cdots, t_k; E_1, \cdots, E_k; R)) \\ &\qquad = P^y\{P^y(\cap_{i=1}^k \{X^y(t + t_i) \in E_i\} | G_c^{y*+}(t)) \in R\}, \end{aligned} \tag{2.19a}$$

where P^y is the measure for the process $X^y(t)$ and $G_c^{y*+}(t)$ is the indicated field for $X^y(t)$. To extend (2.19a) to arbitrary $M \in \mathcal{B}(Y)$ we can make use of Lemma 2.1 applied to the process $X^y(t)$ (it is clear that the limitation of the process to $t > 0$ does not invalidate the lemma, and Hypothesis $*$ holds for $X^y(t)$). By Lemma 2.1 we may choose the conditional probabilities

$$P^y(\cap_{i=1}^k \{X^y(t + t_i) \in E_i\} | G_c^{y*+}(t))$$

in such a way as to define for t and y fixed a measure on the least subfield of $\prod_{\{\alpha(m,n)\}} \mathcal{B}_t$ containing the cylinder sets determined by the

$$E_1(t_1) \times \cdots \times E_k(t_k).$$

Since, moreover, this measure has the transition probabilities $p(t, x, E)$, it has a unique σ-additive extension to $\prod_{\{\alpha(m,n)\}} \mathcal{B}_t$. Designating this extension by $P^y(Q | G_c^{y*+}(t))$ for $Q \in \prod_{\{\alpha(m,n)\}} \mathcal{B}_t$, we define for

$$M = \{y \in Y: P^y(Q) \in R\}$$

$$p_Y(t, y, M) = P^y\{P^y(Q | G_c^{y*+}(t)) \in R\}. \tag{2.19b}$$

It follows that $P^y(Q | G_c^{y*+}(t))$ is indeed a conditional probability of the event in the future of $X^y(t)$ after time t defined by the set Q, and in particular that (2.19b) is well defined. We will next show that $p_Y(t, Y(t), M)$ is measurable over $H_Y(t)$ (the past of $Y(t)$ up to time t), and hence a fortiori measurable over $H^+(t)$. For the case (2.19a) with $k = 1$ we use the fact that

$$(2.20)\quad \begin{aligned} &P^y(\{X^y(t + t_1) \in E_1\} | G_c^{y*+}(t)) \\ &\qquad = \lim_{r\to\infty} p(t_1 + t - t_r(t), X^y(t_r(t)), E_1) \quad \text{a.s.} \end{aligned}$$

(by the martingale convergence theorem, as in (2.2)). Let $f(x)$ be a continuous function on [0, 1]. Since (2.19a) defines a measure in $R \in \mathcal{R}_{[0,1]}$, it follows from (2.20) that

$$(2.21)\quad \begin{aligned} &\int_0^1 f(x) p_Y(t, y, M(t_1, E_1, dx)) \\ &\qquad = \lim_{r\to\infty} \int_0^1 f(x) P^y\{p(t_1 + t - t_r(t), X^y(t_r(t)), E_1) \in dx\}. \end{aligned}$$

Indeed, let $E(r, R) = \{x : p(t_1 + t - t_r(t), x, E_1) \in R\}$. Then since $E(r, R) \in \mathcal{B}$, (2.21) is immediate. Moreover, we have

$$\begin{aligned} P^{Y(t)}\{p(t_1 + t - t_r(t), X^{Y(t)}(t_r(t)), E_1) \in R\} &= P^{Y(t)}\{X^{Y(t)}(t_r(t)) \in E(r, R)\} \\ &= F(t_r(t), E(r, R); t, w), \end{aligned}$$

which is measurable not only over $H_Y(t)$ but over the field generated by $Y(t)$ as the inverse of $\mathcal{B}(Y)$. It follows that for each r,

$$\int_0^1 f(x) P^{Y(t)}\{p(t_1 + t - t_r(t), X^{Y(t)}(t_r(t)), E_1) \in dx\}$$

is measurable over the same field, whence by (2.21) so is

$$\int_0^1 f(x) P^{Y(t)}\{P^{Y(t)}(\{X^{Y(t)}(t + t_1) \in E_1\} \mid G_c^{Y(t)*+}(t)) \in dx\}.$$

This implies the measurability of (2.19a) with $y = Y(t)$ when $k = 1$. Since the measures (2.19a) for $k > 1$ are connected by the transition probabilities $p(t, x, E)$, the measurability for $k > 1$ follows immediately. Measurability of (2.19b) now follows from the fact that measurability is preserved under monotone pointwise limits.

It is next to be shown that $Y(t)$ is a Markov process with (2.19a, b) as conditional probabilities given $H^+(t)$. Let $S \in H^+(t)$, and let $M(t_1, E_1, R)$ be the set defined in (2.19a). For $f(x)$ continuous on [0, 1] we then have

$$(2.22)\quad \begin{aligned} &\int_S \int_0^1 f(x) p_Y(t, Y(t'), M(t_1, E_1, dx))\, dP(w) \\ &\quad = \int_S \lim_{r\to\infty} \int_0^1 f(x) P^{Y(t')}\{X^{Y(t')}(t_r(t)) \in E(r, dx)\}\, dP(w) \\ &\quad = \int_S \lim_{r\to\infty} \int_0^1 f(x) F(t_r(t), E(r, dx); t', w)\, dP(w) \\ &\quad = \lim_{r\to\infty} \int_0^1 f(x) P(S \cap \{p(t_1 + t - t_r(t), X(t' + t_r(t)), E_1) \in dx\}) \\ &\quad = \int_0^1 f(x) P(S \cap \{Y(t' + t) \in M(t_1, E_1, dx)\}), \end{aligned}$$

where for the last step we need to observe that in the definition of

$$P(\{X(t' + t + t_1) \in E\} | G^{*+}(t' + t))$$

by means of $\{t_r(t' + t)\}$ it is possible to replace $t_r(t' + t)$ by $t' + t_r(t)$, the martingale limits being a.s. the same for either case. From (2.22) it follows that $p_Y(t, Y(t'), M(t_1, E_1, R))$ is the required conditional probability given $H^+(t')$. The extension of this result to

$$p_Y(t, Y(t'), M(t_1, \cdots, t_k; E_1, \cdots, E_k; R))$$

is completely analogous, using in place of (2.20)

$$\begin{aligned} & P^y(\cap_{i=1}^k \{X^y(t + t_i) \in E_i\} \mid G_c^{y*+}(t)) \\ (2.23) \quad & = \lim_{r\to\infty} \int_{E_1} \cdots \int_{E_{k-1}} p(t_k - t_{k-1}, x_{k-1}, E_k) p(t_{k-1} - t_{k-2}, x_{k-2}, dx_{k-1}) \\ & \qquad \cdots p(t_1 + t - t_r(t), X^y(t_r(t)), dx_1) \quad \text{a.s.} \end{aligned}$$

Finally, the extension to $p_Y(t, Y(t'), M)$ for arbitrary $M \in \mathcal{B}(Y)$ is valid since, by the Lebesgue monotone convergence theorem, the property of being a conditional probability of $\{Y(t + t') \in M\}$ given $H^+(t')$ is preserved under simultaneous monotone limits of sets M and probabilities $p_Y(t, Y(t'), M)$. This completes the proof that $Y(t)$ is a Markov process with the conditional probabilities (2.19a, b).

We next let T be a stopping time for $X(t)$ and show that the requirements of the strong Markov property for $Y(t)$ are satisfied by T. This will prove that $Y(t)$ is a strong Markov process if it is also shown that any stopping time for $Y(t)$ is also a stopping time for $X(t)$. But if $\{T < t\} \in H^+(t)$ for all t, then $\{T < t\} = (\cup_{k=1}^\infty \{T < t - 2^{-k}\}) \in F(t)$ for all t, and the second fact thus follows from Definition 2. We will need to use (2.7) of Lemma 2.1 with $T_1 = T$ and $E(T)$, depending also on t_1, E_1, r, and R, defined by

$$(2.24) \qquad E(T; r, R) = \{x : p(t_1 + T - t_r(T), x, E_1) \in R\}.$$

It is therefore necessary to show that for $T \leqq T_2$

$$\{X''(t_r(T') + t) \in E(T'; r, R)\} \in H^{+\prime}(T_2') \times L''(T_2'')$$

(we replace the t_1 of (2.7) by t). We have immediately

$$\begin{aligned} & \{X''(t_r(T') + t) \in E(T'; r, R)\} \\ & \quad = \{p(t_1 + T' - t_r(T'), X''(t_r(T') + t), E_1) \in R\} \\ (2.25) \quad & \quad = \cup_{\{\alpha(m,n)\}} [(\{t_r(T') = \alpha(m, n)\} \times \Omega''_{T_2}) \\ & \qquad \cap \{p(t_1 + T' - \alpha(m, n), X''(\alpha(m, n) + t), E_1) \in R\}]. \end{aligned}$$

By the measurability of $p(t, x, E_1)$ in (t, x) it follows that each set in the union is in $H^{+\prime}(T_2') \times L''(T_2'')$ as required. Recalling the definition of $M(t_1, E_1, R)$, we now have for any continuous function $f(x)$ on [0, 1] a.s.

$$\begin{aligned}
&\int_0^1 f(x)p_Y(t, Y(T), M(t_1, E_1, dx)) \\
&\qquad = \lim_{r\to\infty} \int_0^1 f(x)F(t_r(T) + t - T, E(T; r, dx); T, w) \\
&\qquad = \lim_{r\to\infty} \lim_{n\to\infty} \int_0^1 f(x)F(t_r(T) + t - T_n, E(T; r, dx); T_n, w) \qquad (2.26) \\
&\qquad = \lim_{n\to\infty} \lim_{r\to\infty} \int_0^1 f(x)F(t_r(T) + t - T_n, E(T; r, dx); T_n, w) \\
&\qquad = \lim_{n\to\infty} \int_0^1 f(x)p_Y(t - (T_n - T), Y(T_n), M(t_1, E_1, dx)).
\end{aligned}$$

Indeed, the first equality follows from (2.21) and (2.24). The second equality follows from (2.7) with $T_1 = T_2 = T$, together with the definition of $F(t_r(T) + t - T_n, E(T; r, R); T_n, w)$ as a version of

$$P(\{X(t + t_r(T)) \in E(T; r, R)\} | H^+(T_n)),$$

which follows in turn from (2.7) with $T_1 = T$, $T_2 = T_n$. For, in view of these equivalences, we have

$$\begin{aligned}
&\int_0^1 f(x)F(t_r(T) + t - T, E(T; r, dx); T, w) \qquad (2.27a) \\
&\qquad = E(f(p(t_1 + T - t_r(T), X(t_r(T) + t), E_1)) \mid H^+(T)) \quad \text{a.s.}, \\
&\int_0^1 f(x)F(t_r(T) + t - T_n, E(T; r, dx); T_n, w) \qquad (2.27b) \\
&\qquad = E(f(p(t_1 + T - t_r(T), X(t_r(T) + t), E_1)) \mid H^+(T_n)) \quad \text{a.s.},
\end{aligned}$$

and the equality in question follows by martingale convergence in n.

The interchange of limits for the third equality is justified by (2.27b), from which we see that the convergence in r is uniform in the martingale convergence in n except for a set whose probability is small when r is large. In particular, this proves the existence of the limit in n. The last equality, like the first, is a result of (2.21). An extension of (2.26) to

$$M(t_1, \cdots, t_k; E_1, \cdots, E_k; R)$$

proceeds the same way, using (2.23) in the (new) definition of $E(T; r, R)$.

If Theorem 1 is applied to the process $Y(t)$, leaving aside the question of measurability of $Y(t)$, one formally obtains from (2.26) that

$$\begin{aligned}
&\lim_{n\to\infty} \int_0^1 f(x)P_T(\{Y(T + t) \in M(t_1, E_1, dx)\} \mid H^+(T_n)) \\
&\qquad = \int_0^1 f(x)P_T(\{Y(T + t) \in M(t_1, E_1, dx)\} \mid H^+(T)) \qquad (2.28) \\
&\qquad = \int_0^1 f(x)P_T(\{Y(T + t) \in M(t_1, E_1, dx)\} \mid G_Y^{*+}(T)) \\
&\qquad = \int_0^1 f(x)p_Y(t, Y(T), M(t_1, E_1, dx)) \quad \text{a.s.},
\end{aligned}$$

where $G_Y^{*+}(T)$ is the field $G^{*+}(T)$ for the process $Y(t)$. Similarly, one would obtain (2.28) with $M(t_1, \cdots, t_k; E_1, \cdots, E_k; dx)$ in place of $M(t_1, E_1, dx)$. From this it then follows that

$$P_T(\{Y(T+t) \in M(t_1, \cdots t_k; E_1, \cdots, E_k; R)\}|H^+(T)) = p_Y(t, Y(T), M(t_1, \cdots, t_k; E_1, \cdots, E_k; R)) \quad \text{a.s.,} \tag{2.29}$$

and (2.29) extends to arbitrary $M \in \mathcal{B}(Y)$ by taking monotone limits. Hence $Y(t)$ will satisfy the requirements of the strong Markov property for T, and if T can be chosen arbitrarily, then $Y(t)$ is a strong Markov process.

To complete the proof of Theorem 2 it remains only to show that Theorem 1 may be applied to the process $Y(t)$ for any stopping time T.

It is easily seen that the probabilities (2.19a, b) can replace transition probability functions in the proof. The missing property is measurability of $Y(t)$ over $\mathcal{R} \times L_Y$, where L_Y is the σ-field generated by the sets $\{Y(t) \in M\}$, $M \in \mathcal{B}(Y)$. The only point at which this measurability is used in the proof is in showing that $\{(w', w''): Y''(T' + t) \in M\} \in H'_Y(T'_n) \times L''_Y(T'')$, where $H'_Y(T'_n) = H'^+(T'_n)$ because of the definition of the fields of the past for $Y(t)$. A check of the proof reveals, in fact, that it is sufficient that the above set be only in the completion $(H'_Y(T'_n) \times L''_Y(T''))_c$ with respect to μ since $H'^+(T'_n)$ is complete with respect to P'_T We now require the following lemma:

LEMMA 2.2.[5] *Let $\{t_n\}$ be a sequence decreasing to 0. Then for each continuous function $f(x)$ on [0, 1] and each stopping time T*

$$\lim_{t_n \downarrow 0} \int_0^1 f(x) p_Y(t_n, Y(T), M(t_1 - t_n, E_1, dx)) = f(F(t_1, E_1; T, w)) \quad a.s.$$

Proof. From the definition of $p_Y(t, y, M)$ we have

$$\begin{aligned}
&\lim_{t_n \downarrow 0} \int_0^1 f(x) p_Y(t_n, Y(T), M(t_1 - t_n, E_1, dx)) \\
&\quad = \lim_{t_n \downarrow 0} \int_0^1 f(x) P^{Y(T)}\{P^{Y(T)}(\{X^{Y(T)}(t_1) \in E_1\} \mid G^{Y(T)*+}(t_n)) \in dx\} \\
&\quad = \int_0^1 f(x) P^{Y(T)}\{P^{Y(T)}(\{X^{Y(T)}(t_1) \in E_1\} \mid G^{Y(T)*+}(0)) \in dx\}.
\end{aligned} \tag{2.30}$$

Let a.s.$_y$ designate a.s. for the process $X^y(t)$. It remains to prove that

$$P^{Y(T)}(\{X^{Y(T)}(t_1) \in E_1\}|G^{Y(T)*+}(0)) = F(t_1, E_1; T, w) \quad \text{a.s.}_{Y(T)}, \quad \text{a.s.} \tag{2.31}$$

By applying Lemma 2.1 to any process $X^{Y(T)}(t)$ it follows that

$$P^{Y(T)}(\{X^{Y(T)}(t) \in (\cdot)\}|G^{Y(T)*+}(0))$$

may be chosen to be a probability on $\mathcal{B}$ for each t, and to have transition prob-

[5] This lemma expresses the sense in which the regularity property of convergence of the semigroup mentioned in the introduction to the identity at $t = 0$ is satisfied by $Y(t)$.

abilities $p(t, x, E)$. It is now sufficient to prove for this choice of conditional probabilities that for any t, a.s. on the set $\{T < t \leqq T + t_1\}$,

$$(2.32)\qquad \begin{aligned} &P^{Y(T)}(\{X^{Y(T)}(t - T) \in (\cdot)\} | G^{Y(T)*+}(0)) \\ &\qquad = F(t - T, (\cdot); T, w) \quad \text{a.s.}_{Y(T)} . \end{aligned}$$

For (2.31) then follows by integrating (2.32) with $p(T + t_1 - t, x, E_1)$ as integrand and taking a union over rational t. If (2.32) holds for a particular process $X^{Y(T)}(t)$ for a fixed $y = Y(T)$, then it holds for every process with the same joint distributions. We will show that for almost all $Y(T)$ a process $X^{Y(T)}(t)$ may be defined for which (2.32) holds. Let us consider the definition of $X(t)$ extended to the space $(\Omega \cup \bigcup_{t'}(\prod_\tau X_\tau)_{t'})$ i.e., to Ω together with a replica of $\prod_\tau X_\tau$ corresponding to each t'. For $w \in (\prod_\tau X_\tau)_{t'}$ we define $X(t, w) = w_{\tau=t}$, i.e., the coordinate of w in X_t. We extend the measure by setting $P(S) = 0$ for all $S \subset \bigcup_{t'} (\prod_\tau X_\tau)_{t'}$, and stipulate that $S \in F(t)$ for all $S \subset \bigcup_{t' \leqq t} (\prod_\tau X_\tau)_{t'}$, for each t. If we construct $Y(t)$ on this new space, and extend the definition of T by setting $T = t'$ on $(\prod_\tau X_\tau)_{t'}$, it is clear that (2.32) will then hold a.s. on $\{T < t \leqq T + t_1\}$ if and only if it already holds for the original space and process $Y(t)$. Indeed, the original space is imbedded in the new space in such a way that $Y(T)$ in the new space is a.s. in the image of the original space. Let us now define the sets $\Omega(T, Y(T))$ by

$$\begin{aligned} \Omega(T, Y(T)) = \{w \in (\Omega \cup \textstyle\bigcup_{t'}(\prod_\tau X_\tau)_{t'}) &: \text{for } T < t < \infty, \\ F(t - T, (\cdot); T, w) &= F(t - T, (\cdot); T, w_{Y(T)})\} \end{aligned}$$

where $w_{Y(T)}$ is any point such that $T(w_{Y(T)}) = T$ and $Y(T, w_{Y(T)}) = Y(T)$. In other words, $\Omega(T, Y(T))$ is the class of all points w defining the same T and $y \in Y$ as appear in $Y(T)$. Moreover, define the processes $X^{Y(T)}(t)$ by $X^{Y(T)}(t) = X(T + t)$ on $\Omega(T, Y(T))$. It is possible to introduce the distributions defining $X^{Y(T)}(t)$ on the space $\Omega(T, Y(T))$ since, because of the extension, $\Omega(T, Y(T))$ contains points of every cylinder set

$$\{X(T + t_1) \in E_1, \cdots, X(T + t_k) \in E_k\},$$

and is large enough, in fact, so that these distributions may be extended to a probability measure. We will show that (2.32) holds a.s. on $\{T < t \leqq T + t_1\}$ for these processes $X^{Y(T)}(t)$. Indeed, for these processes we have a.s. on $(\Omega \cup \bigcup_{t'}(\prod_\tau X_\tau)_{t'}) \cup \{T < t\}$, using (2.9),

$$(2.33)\qquad \begin{aligned} &\lim_{r\to\infty} p(t - t_r(T), X^{Y(T)}(t_r(T) - T), E) \\ &\qquad = \lim_{r\to\infty} p(t - t_r(T), X(t_r(T)), E) \\ &\qquad = P(\{X(t) \in E\} | G^{*+}(T)) \\ &\qquad = F(t - T, E; T, w). \end{aligned}$$

Since the first term is equal to

$$P^{Y(T)}(\{X^{Y(T)}(t - T) \epsilon E\}|G^{Y(T)*+}(0)) \quad \text{a.s.}_{Y(T)},$$

for each $Y(T)$, and since also

$$\begin{aligned} P(\{\lim_{r\to\infty} p(t - t_r(T), X^{Y(T)}(t_r(T) - T), E) \\ = F(t - T, E; T, w)\} \mid Y(T)) \\ = P^{Y(T)}(\{P^{Y(T)}(\{X^{Y(T)}(t - T) \epsilon E\}|G^{Y(T)*+}(0)) \\ = F(t - T, E; T, w)\}) \quad \text{a.s.}, \end{aligned}$$

by using (2.9) once more, the proof of (2.32) is completed by letting E range through a countable field generating $\mathcal{B}$. This proves Lemma 2.2.

Let us note at this point that Lemma 2.2 and the proof remain valid if $M(t_1, \cdots, t_k; E_1, \cdots, E_k; dx)$ is substituted for $M(t_1, E_1, dx)$, and at the same time

$$\int_{E_1} \cdots \int_{E_{k-1}} p(t_k - t_{k-1}, x_{k-1}, E_k) p(t_{k-1} - t_{k-2}, x_{k-2}, dx_{k-1}) \cdots F(t_1, dx_1; T, w)$$

is substituted for $F(t_1, E_1; T, w)$. From Lemma 2.2 and (2.26) it follows now that $f(F(t_1, E_1; T', w''))$ is measurable over $(H'^{+}(T'_n) \times L''_Y(T''))_c$. Indeed, we see from (2.24) that $E(t; r, R)$ defines the section at $t = T$ of a measurable set in $\mathcal{R} \times \mathcal{B}$. It follows that

$$F(t_r(T') + t - T'_n, E(T'; r, R); T'_n, w'')$$

is measurable over $H'^{+}(T'_n) \times L''_Y(T'')$, whence the conclusion follows from the first and third lines of (2.26) upon applying Lemma 2.2. The same result holds after replacement of $F(t_1, E_1; T', w'')$ by (2.18), by the extension of (2.26) to that case. Again, the same result holds with $T + t$ in place of T, the former being again a stopping time measurable over $H^{+}(T_n)$. We therefore have, since $f(x)$ is arbitrary, that

$$\begin{aligned} \{(w', w''): Y''_T(T' + t) \epsilon M(t_1, \cdots, t_k; E_1, \cdots, E_k; R)\} \\ \epsilon (H'^{+}(T'_n) \times L''_Y(T''))_c. \end{aligned} \tag{2.34}$$

But (2.34) then holds for arbitrary $M \epsilon \mathcal{B}(Y)$ by closure under monotone limits. This completes the proof of Theorem 2.

It remains to consider the connection between the processes $X(t)$ and $Y(t)$. This connection will be established by identifying a point $y \epsilon Y$ with a point $x \epsilon X$ whenever for all t and E

$$P^y\{X^y(t) \epsilon E\} = p(t, x, E). \tag{2.35}$$

We shall assume that for $x_1 \neq x_2$ the functions $p(\cdot, x_1, \cdot)$ and $p(\cdot, x_2, \cdot)$

are never identical. This may be brought about, if necessary, by a preliminary identification of the points of X into equivalence classes [7, p. 45].

THEOREM 3. *Under the above assumption, and after the identification* (2.35), $Y(t)$ *and* $X(t)$ *agree except on at most a set of probability* 0 *for each* t *and a countable set of* t *for each* w.

Proof. According to a theorem of Doob [2, p. 363] the limits (2.8) exist from the left, as well as from the right, a.s. for all t. It follows that a.s. there are for each $\varepsilon > 0$ only finitely many discontinuities of oscillation greater than ε in (2.8) for fixed $\alpha(m, k)$ and E_i. Consequently, a.s. the limits as $\tau \uparrow 0$, $t + \tau \in A$, in (2.2) agree with the limits as $\tau \downarrow 0$, $t + \tau \in A$, except for a countable set of t depending on w. For each t this common limit is equal to $p(t_r(t) - t, X(t), E_i)$ a.s. Indeed, we may adjoin t to the (countable) set A. The martingale (2.2) on the enlarged parameter set will be continuous at t a.s. where the right and left limits on A are equal at t [2, Theorem 11.5]. Except for sets of the type stated in Theorem 3, therefore, the limits (2.2) are found for all E_i by substitution of $\tau = 0$. We then have by extension of the measure, outside of these exceptional sets, $F(t_1, E; t, w) = p(t_1, X(t), E)$ for all $E \in \mathcal{B}$ and $t > 0$, when $t_1 > t_r(t) - t$. Letting r become large in (2.4) and then letting t_1 decrease to 0 through a sequence of values, it follows from the definition of $P^y\{X^y(t_1) \in E\}$ that $Y(t) = X(t)$ in the sense of (2.35) outside of sets of the type stated in the theorem.

As a concluding remark, we have the statement referred to following (2.1). For $x \in X$, let $L(x)$ designate the point of Y defined from x by (2.35).

COROLLARY.[6] *There exists a strong Markov process* $\underset{\sim}{X}(t)$ *in the narrow sense on* (Ω, F, P) *relative to the fields* $F(t)$, *with state space* Y *and measurable field* $\mathcal{B}(Y)$, *such that* $P\{\underset{\sim}{X}(t) = L(X(t))\} = 1$ *for all* t.

Proof. It is clear that $P\{Y(t) = L(X(t))\} = 1$ unless t is a fixed point of discontinuity for one of the martingales (2.8), and the set D of these fixed discontinuities is at most countable [2, Theorem 11.2]. Accordingly, we define $\underset{\sim}{X}(t)$ as follows:

$$\underset{\sim}{X}(t) = \begin{cases} Y(t) & \text{if } t \notin D, \\ L(X(t)) & \text{if } t \in D. \end{cases} \tag{2.36}$$

Then $P\{\underset{\sim}{X}(t) = L(X(t))\} = 1$ for all t.

To define the conditional probabilities for $\underset{\sim}{X}(t)$ it is necessary to treat separately the transitions starting or terminating at a time $t_d \in D$. For these, we define

$$p(\{\underset{\sim}{X}(t_d + t) \in M\} | \underset{\sim}{X}(t_d)) = p(t, X(t_d), L^{-1}(M)),$$

[6] Let $\mathcal{B}^\star$ be the σ-field generated by the sets $\{x : p(\alpha(m, n), x, E) \in R\}$, $m > 0, n > 0$. Then $X(t)$ is already a Markov process relative to $\mathcal{B}^\star$ (see the remark following Theorem 2). Moreover, we can write $\mathcal{B}^\star = L^{-1}(\mathcal{B}(Y))$. It follows by the completeness of $F(t)$ that for t fixed the sets of $L(\mathcal{B}^\star)$ are measurable for $\underset{\sim}{X}(t)$. However it is not necessarily true that $L(\mathcal{B}^\star) \subset \mathcal{B}(Y)$.

and for $t < t_d$,

$$p(\{\underline{X}(t_d) \in M\} \mid \underline{X}(t)) = \begin{cases} 0 & \text{if } \underline{X}(t) \neq L(X(t)), \\ p(t_d - t, X(t), L^{-1}(M)) & \text{otherwise.} \end{cases}$$

Since $L^{-1}(\mathcal{B}(Y)) \subset \mathcal{B}$, these definitions are meaningful. For the remaining cases we define

$$p(\{\underline{X}(t + t_1) \in M\} | \underline{X}(t)) = p_Y(t_1, \underline{X}(t), M).$$

The functions thus defined are not stationary in time, but this has not been required.

To show that $\underline{X}(t)$ is a strong Markov process in the narrow sense relative to these conditional probabilities, three cases must be considered. For a given narrow-sense stopping time T, we consider first the set $\{T \in D\}$. On this set T is a Markov time by (2.1) and the fact that $\underline{X}(t) = L(X(t))$ a.s. for each t. Next, consider the set $\{T = t_i, t_i \notin D, P\{T = t_i\} > 0\}$, in which at most countably many t_i can be involved. On this set $Y(T) = L(X(T)) = \underline{X}(T)$ a.s., and T is a Markov time again by (2.1). Finally, consider the set $S = \{T \notin D, T = t, P\{T = t\} = 0\}$, in which uncountably many t may be involved. On this set T is a Markov time because of the strong Markov property for $Y(t)$, provided it is shown that $p_Y(t_1, Y(T), M)$ on S is measurable over $H(T)$, as required by the definition of the strong Markov property in the narrow sense (measurability over $H^+(T)$ was proved above). However, we have

$$\begin{aligned} &S \cap \{T \leqq t\} \cap \{p_Y(t_1, Y(T), M) \in R\} \\ &\quad = \textstyle\bigcup_{n=1}^{\infty} (S \cap \{T \leqq t - \alpha(1, n)\} \cap \{p_Y(t_1, Y(T), M) \in R\}) \cup \Phi, \end{aligned} \tag{2.37}$$

where $\Phi \subset \{T = t\}$. Each set in the first union is in $F(t)$, and $F(t)$ is complete for P. Thus $S \cap \{p_Y(t_1, Y(T), M) \in R\} \in H(T)$, and the corollary is proved.

References

1. R. M. Blumenthal, *An extended Markov property*, Trans. Amer. Math. Soc., vol. 85 (1957), pp. 52–72.
2. J. L. Doob, *Stochastic processes*, New York, Wiley, 1953.
3. L. E. Dubins, *Conditional probability distributions in the wide sense*, Proc. Amer. Math. Soc., vol. 8 (1957), pp. 1088–1092.
4. E. Dynkin and A. Yushkevič, *Strong Markov processes*, Teor. Veroyatnost. i Primenen., vol. 1 (1956), pp. 149–155 (in Russian, English summary).
5. P. R. Halmos, *Measure theory*, New York, Van Nostrand, 1950.
6. E. Hille and R. S. Phillips, *Functional analysis and semi-groups*, rev. ed., Providence, American Mathematical Society, 1957.
7. D. Ray, *Resolvents, transition functions, and strongly Markovian processes*, Ann. of Math. (2), vol. 70 (1959), pp. 43–72.
8. A. A. Yushkevič, *On strong Markov processes*, Teor. Veroyatnost. i Primenen., vol. 2 (1957), pp. 187–213 (in Russian, English summary).
9. ———, *On the definition of a strong Markov process*, Teor. Veroyatnost. i Primenen., vol. 5 (1960), pp. 237–243 (in Russian, English summary).

University of Minnesota
Minneapolis, Minnesota