NOTE ON REGULARIZATION OF MARKOV PROCESSES

BY
FRANK KNIGHT

The purpose of this note is to show that for an arbitrary Markov process with a transition function $p(t, x, E)$ measurable in (t, x) on a measurable space $(X, \mathcal{B})$, where $\mathcal{B}$ is generated by countably many sets, there exists a strict Markov process with right-continuous paths on a compact, separable, metric completion of X, with resolvent carrying the continuous functions C into C, with a transition function on the topological Borel field, and which is equal to the original process with probability 1 at each t except for t in a countable set. To this effect we shall use Theorem III of [5], which requires one additional hypothesis: we assume that the family $p(\cdot, x, \cdot)$, $t > 0$, $E \in \mathcal{B}$, separates points in X.[1] It will also be proved that the same result holds for a non-homogeneous process with transition function $p(t_1, x; t_2, E)$ measurable in (t_1, x, t_2). These facts extend the conclusions of [3] under weaker hypotheses. The metric space involved here is different from that of [3], however, and its definition is more complicated.[2] Finally, we give an example to indicate that the present results cannot be strengthened to obtain, in general, a semigroup carrying C into C.

The construction of the metric depends on the lemma which follows.

LEMMA 1. *There exists a countable collection S of $\mathcal{B}$-measurable functions f, $0 \leq f \leq 1$, with the following properties.*

(a) *S contains the indicator functions χ_i of the sets E_i in a countable field generating $\mathcal{B}$.*

(b) *The linear closure of S in the uniform norm is closed under the resolvent operators*

$$R_\lambda f(x) = \int_0^\infty \int_X e^{-\lambda t} p(t, x, dy) f(y)\, dt, \qquad \lambda > 0.$$

(c) *The linear closure of S is closed under multiplication.*

Proof. It is known that a countable collection of sets generates a countable field. Consequently, let $\{E_i\}$ be a countable field generating $\mathcal{B}$, and let S, to begin with, contain the indicator functions χ_i of E_i, $1 \leq i$. Imitating the proceedure in [3, *p.* 327], we proceed inductively as follows. Let $\{\mu_i\}$ be countable dense set in $(0, \infty)$, and let S_0 be $\{\chi_i\}$. Then given S_n, $n \geq 0$, let S_{n+1} be obtained by applying countably often in alternating succession

Received March 23, 1964.

[1] As noted in [3] and [4], this may always be brought about by a preliminary identification of points in X.

[2] Actually, Ray's results hold directly for the situation of [3], as follows from a forthcoming paper of H. Kunita and T. Watanabe. In this case, however, one does not always have resolvents mapping $C \to C$.

the operations of (i) applying $\mu_{n+1} R_{\mu_{n+1}}$ and adjoining the range obtained, then applying successively $\mu_1 R_{\mu_1}, \cdots, \mu_n R_{\mu_n}$ and adjoining their ranges, and (ii) closing the resulting set under the operation of multiplication. It is clear that for each n, S_n is countable, is closed under $\mu_1 R_{\mu_1}, \cdots, \mu_n R_{\mu_n}$, and is closed under multiplication. Let $S = \lim_{n\to\infty} S_n$. Then S is closed under $\mu_n R_{\mu_n}$, $0 \leq n$, and under multiplication, and the elements of S are bounded by 0 and 1. Finally, by the continuity in μ of R_μ in the uniform operator topology, it is seen that the uniform linear closure of S is closed under R_μ, $\mu > 0$, as well as under multiplication. Thus S satisfies the requirements of the lemma.

Turning to the definition of our metric, let $\lambda > 0$ be fixed and $\{g(i)\}$, $g(i) > 0$, $\sum_i g(i) < \infty$ be a sequence of real numbers, and let $\{f_i\}$ be an enumeration of the range $\lambda R_\lambda(S)$ together with all finite products of elements in that range. Then we set

$$d(x, y) = \sum_{i=1} g(i) \left| \frac{f_i(x)}{\sum_{i=1}^{\infty} g(i) f_i(x)} - \frac{f_i(y)}{\sum_{i=1}^{\infty} g(i) f_i(y)} \right|.$$

It follows easily (see [3, p. 325]) that $d(x, y)$ is a metric on X, that convergence is equivalent to simultaneous convergence of the numerators $f_i(x)$, and that the completion $\bar{X}$ of X in this metric is a compact, separable metric space of diameter less than or equal to 1.

It will next be shown that the resolvent R_μ may be extended to a resolvent on the space $\bar{C}$ of continuous functions on $\bar{X}$, satisfying the hypotheses of [5, Theorems I and III]. Since the family $\{f_i\}$ is closed under multiplication, it follows from the Stone-Weierstrass theorem that the space $\bar{C}$ is the uniform linear closure of the family $\{\bar{f}_i\}$, where $\bar{f}_i$ is the continuous extension of f_i to $\bar{X}$. Lemma 1 implies that R_μ, $\mu > 0$, carries the uniform linear closure of $\{f_i\}$ into itself, since this closure is contained in Cl S, and since $R_\mu f = R_\lambda(f(\lambda - \mu) R_\mu f)$. We may therefore define a resolvent $\bar{R}_\mu$ on $\bar{C}$ by restricting each $\bar{g} \in \bar{C}$ to $g \in C$, and setting $\bar{R}_\mu \bar{g} = \overline{R_\mu g}$. The family $\lambda R_\lambda(S)$ separates points in X, hence its extension to a family of functions in $\bar{C}$ separates points in $\bar{X}$. It thus satisfies the hypotheses of [5, Theorems I and III].

Returning to the process $X(t)$, we define

$$\bar{X}(t) = \lim_{\tau \downarrow t, \tau \text{ rational}} X(\tau)$$

when this limit exists in the metric $d(x, y)$ for all $t \geq 0$ simultaneously, and we set $\bar{X}(t) = x_0$ for some fixed $x_0 \in X$ otherwise. Since, for $f_i \in \lambda R_\lambda(S)$, $e^{-\lambda\tau} f_i(X(\tau))$ is a supermartingale in τ, and therefore has simultaneous right-hand limits with probability 1, the first case of our definition is seen to have probability 1.

THEOREM 1. *$\bar{X}(t)$ is a regularization of $X(t)$ as described at the start.*

Proof. That $X(t) = \bar{X}(t)$ a.s. except for t in some countable set follows from a theorem of Doob [1, Theorem 11.2] according to which a separable supermartingale has at most countably many fixed discontinuities (see [2, p. 612]). To complete the proof, it is sufficient to establish that $\bar{X}(t)$ is one of the processes "$Y_t(w)$" of [5, Theorem III] corresponding to the resolvent $\bar{R}_\mu$. This follows exactly as in [3, p. 335], but for the reader's convenience we repeat the proof. Let $\bar{p}(t, x, E)$ be the transition function associated with $\bar{R}_\mu$ by [5, Theorem I]. Then $\bar{p}(t, x, E)$ is the unique transition function with this resolvent and such that for $\bar{g} \in \bar{C}$, $\int_{\bar{X}} \bar{p}(t, x, dy)\bar{g}(y)$ is right-continuous in t, $t > 0$, for each x. Since $\bar{X}(t)$ is right-continuous, it need only be shown that $\bar{p}(t, x, E)$ is a transition function for $\bar{X}(t)$. We set $H^+(t) = \bigcap_{\varepsilon>0} \mathfrak{F}(t+\epsilon)$, where $X(t')$, $t' \leq t$, is measurable with respect to $\mathfrak{F}(t)$, and hence $\bar{X}(t')$, $t' \leq t$, is measurable with respect to $H^+(t)$. Then it is sufficient to prove that for $\bar{g} \in \bar{C}$, $t' > 0$, and $A \in H^+(t')$,

$$\int_A \int_0^\infty e^{-\mu t} \int_{\bar{X}} \bar{p}(t, \bar{x}(t'), dx)\bar{g}(x)\, dt\, dP = \int_A \int_0^\infty e^{-\mu t}\bar{g}(\bar{X}(t' + t))\, dt\, dP.$$

If A is in $\mathfrak{F}(t')$ and $P\{\bar{X}(t') = X(t')\} = 1$, then this follows since

$$\int_A \int_0^\infty e^{-\mu t} \int_X p(t, x(t'), dx)g(x)\, dt\, dP = \int_A \int_0^\infty e^{-\mu t}g(X(t' + t))\, dt\, dP.$$

In the contrary case, we may choose a sequence $t_n \downarrow t'$ for which the above conditions hold, A being in $\mathfrak{F}(t_n)$ automatically, and then let $n \to \infty$, using right-continuity of $\bar{X}(t)$ and continuity of $\bar{R}_\mu \bar{g}$, to derive the identity in the limit. This completes the proof.[3]

The construction will now be extended to the nonhomogeneous case. Let $X(t)$ be any Markov process on $(X, \mathfrak{B})$ with a transition function $p(t_1, x; t_2, E)$, and form the "space-time process" $(X(t), t)$ on

$$(X \times (0, \infty), \mathfrak{B} \times \mathfrak{B}(0, \infty)),$$

where $\mathfrak{B}(0, \infty)$ is the real Borel field on $(0, \infty)$.[4] We assume that for each

[3] The connection between the topological σ-field of $\bar{X}$ and $\mathfrak{B}$ is derived in [3, Theorem 1.3]. In short, the restriction of this field to X coincides with the subfield of $\mathfrak{B}$ generated by the measurable family $p(t, \cdot, E)$, $t > 0$, $E \in \mathfrak{B}$. At this point, we call attention to two errors in [3]. In Theorem 1.1, the field $G^{*+}(T)$ must be replaced by the field $G^+(T) = \bigcap_{n=1}^\infty G_n(T)$, where $G_n(t)$ is generated by the sets $\{X(t_k) \in E\}$, and $\{t_k \in R\}$, R in the real Borel sets, and in (1.2),

$$F(t_1, dx_1; T, w)$$

must be replaced by

$$F(t_1 - T, dx_1; T, w).$$

Secondly, in the proof of Theorem 1.3, p. 331, line–6 ff., it is necessary to replace "simple function," by "simple function measurable over $\mathfrak{I} \times \mathfrak{B}$, where $\mathfrak{I}$ is the ring of finite unions of intervals". The approximation must then be almost uniformly rather than uniformly, and the boundedness of the convergence must be used in passing to the limit.

[4] The author is indebted to Professor S. Orey for the idea of using the space-time process for this purpose.

t_1, $p(t_1, \cdot; t_2, E)$, $t_2 > t_1$, $E \in \mathcal{B}$, separates points in X. The transition function of $(X(t), t)$ is defined by

$$p(t, (x, t'), E \times \{t' + t\}) = p(t', x; t' + t, E).$$

The earlier construction may now be applied to the process $(X(t), t)$. For the sets E_i, we choose the field generated by the sets $A_j \times I_k$, where the A_j generate $\mathcal{B}$ and the I_k are all intervals with rational endpoints. Since the space $X \times (0, \infty)$ is contained in the state space of the regularized process, we may reidentify the points with the same x-coordinate, and introduce the corresponding non-homogeneous transition function

$$\bar{p}(t_1, x; t_2, E) = \bar{p}(t_2 - t_1, y, E \times R) \qquad \text{for } y = (x, t_1), \text{ and}$$

$$\bar{p}(t_1, y; t_2, E) = \bar{p}(t_2 - t_1, y, E \times R) \qquad \text{for } y \text{ not of the form } (x, t).$$

It is clear that the Markov property is retained, and that the resulting process with the points (x, t) identified to x satisfies the regularity assertions (except, of course, for the one involving a resolvent). However, it is no longer clear what connection there is between the topological σ-field, following the identification of points, and the original field $\mathcal{B}$. According to [3, Theorem 1.3, corollary] there is for each set in the field generated by

$$p(t, (x, t'), E \times R), t > 0, E \in \mathcal{B}, R \in \mathcal{B}(0, \infty),$$

a set in the topological σ-field whose restriction to $X \times (0, \infty)$ is equal to the given set. But the field generated by $p(t, (x, t_2), E \times R)$ contains the products of sets in that generated by $p(t_2, x; t_3, E)$, $0 \leq t_2 \leq t_3$, $E \in \mathcal{B}$, with the point t_2, since in particular $p(t, (x, t_2), E \times (0, \infty))$ must be measurable, and we may set $t = t_3 - t_2$. Hence for each set B in the field generated by $p(t_2, x; t_3, E)$ there is a set $\bar{B}$ for the regularized process $\bar{X}$ such that, if $P\{X(t_2) = \bar{X}(t_2)\} = 1$, then

$$P\{(X(t_2) \in B) \triangle (\bar{X}(t_2) \in \bar{B})\} = 0.$$

In the homogeneous case, it would be desirable if the completion discussed above could be further enlarged in such a way that the resulting semigroups T_t would map $C \to C$. One knows that T_t maps the closure of the range of R_μ into itself, but this will not, in general, be all of C. The following example, moreover, is designed to show that a completion large enough to result in a semigroup mapping C into C may destroy the possibility of any identification between the original process and the "regularized" process. In this example the latter differs from the former with positive probability at each t. It was suggested by [4, Example IV] of a process in which R_λ maps C into C but T_t does not.

Example 1. *Let $X_{x'}(t)$ be a Markov process on the space $(0, 1] \cup \{2\}$, with $X_{x'}(0) = x'$, and the transition function*

$$
\begin{aligned}
&p(t, 2, 2) = 1; && \\
&p(t, x, x + t) = 1 && \text{for } x + t \leq 1; \\
&p(t, x, x + t - 1) = \tfrac{1}{2} = p(t, x, 2) && \text{for } 1 < x + t \leq 2; \\
&p(t, x, x + t - n) = 2^{-n}, \qquad p(t, x, 2) = 1 - 2^{-n} && \\
& && \text{for } n < x + t \leq n + 1, n \geq 1.
\end{aligned}
$$

The process consists of uniform translation to the right on $(0, 1]$ followed by return to 0^+ or absorption at 2, each with probability $\frac{1}{2}$. It is not difficult to see that for any countable field $\{E_i\}$ generating $\{2\} \cap \{$the Borel sets of $(0, 1]\}$ and any countable family $\{f_i\}$ of functions containing the indicators χ_i of E_i and generating a metric in which the functions $R_\lambda f_i$ are uniformly continuous, as constructed above, the metric topology is identical to the Euclidean topology on $(0, 1] \cup \{2\}$, and the completion is $[0, 1] \cup \{2\}$. For any continuous function $\bar{g}$ on this space such that $\bar{g}(1) \neq \frac{1}{2}(\bar{g}(0) + \bar{g}(2))$, and each $x \in (0, 1)$, the semigroup $\bar{T}_t$ of the regularized process $\bar{X}_{x'}(t)$ (which has right-continuous paths and thus never reaches the point 1) is discontinuous at x when $t = 1 - x$. In order to make $\bar{T}_t$ carry continuous functions into themselves, therefore, each point $x \in (0, 1)$ must be split into two points $x_1(x)$ and $x_2(x)$ corresponding to limits from the left and from the right of $\bar{T}_t f$ at $t = 1 - x$. But since

$$\bar{T}_{1-x}\, \bar{g}(x_1(x)) = g(1), \quad \text{while} \quad \bar{T}_{1-x}\, \bar{g}(x_2(x)) = \tfrac{1}{2}(\bar{g}(0) + \bar{g}(2)),$$

in order to identify the original process $X_{x'}(t)$ with the process $\bar{X}_{x'}(t)$ one would have to identify $x \in (0, 1)$ with $x_1(x)$. On the other hand, right-continuity of path for $\bar{X}_{x'}(t)$ requires that $\bar{X}_{x'}(t)$ take on the value $x_2(x)$ whenever $X_{x'}(t) = x$. Thus the process $X_{x'}(t)$, $x' \neq 2$, would differ from $\bar{X}_{x'}(t)$ at all t up to and including $\inf_t : X_{x'}(t) = 2$. This concludes the example.

References

1. J. L. Doob, *Stochastic processes*, New York, Wiley, 1953.
2. F. Knight, *On the regularity of Markov processes*, Illinois J. Math., vol. 5 (1961), pp. 591–613.
3. ———, *Markov processes on an entrance boundary*, Illinois J. Math., vol. 7 (1963), pp. 322–335.
4. D. Ray, *Stationary Markov processes with continuous paths*, Trans. Amer. Math. Soc., vol. 82 (1956), pp. 452–493.
5. ———, *Resolvents, transition functions, and strongly Markovian processes*, Ann. of Math. (2), vol. 70 (1959), pp. 43–72.

University of Illinois
Urbana, Illinois