REPRESENTATIONS AND CLASSIFICATIONS
OF STOCHASTIC PROCESSES

BY

DUDLEY PAUL JOHNSON

ABSTRACT. We show that to every stochastic process X one can associate a unique collection $(\Phi, \Phi_+, T(t), E(U), p^*)$ consisting of a linear space Φ, on which is defined a linear functional p^*, together with a convex subset Φ_+ which is invariant under the semigroup of operators $T(t)$ and the resolution of the identity $E(U)$. The joint distributions of X, there being one version for each $\phi \in \Phi_+$, are then given by

$$P_\phi(X(t_1) \in U_1, \cdots, X(t_1 + \cdots + t_n) \in U_n) = p^* E(U_n)T(t_n) \cdots E(U_1)T(t_1)\phi.$$

To each ϕ contained in the extreme points Φ_{++} of Φ_+ and each time t we find a probability measure $P_t^*(\phi, \cdot)$ on Φ_{++} such that $T(t)\phi = \int_{\Phi_{++}} \psi P_t^*(\phi, d\psi)$. P_t^* is the transition probability function of a temporally homogeneous Markov process X^* on Φ_{++} for which there exists a function f such that $X = f(X^*)$. We show that in a certain sense X^* is the smallest of all Markov processes Y for which there exists a function g with $X = g(Y)$. We then apply these results to a class of stochastic process in which future and past are independent given the present and the conditional distribution, on the past, of a collection of random variables in the future.

Introduction. It is the purpose of this paper to find minimal algebraic and Markov process representations of general stochastic processes and then to apply these representations to what we hope will turn out to be an interesting class of stochastic processes. The reader should note that our definition of a stochastic process will not be the standard one. Indeed we will define a stochastic process to be simply a convex set of probability measures on the sample path space which is closed under conditioning on the past. There are several reasons for our doing this. In the first place, a Markov process consists of a convex set of measures on the sample path space, one measure for each initial distribution, and these measures are closed under conditioning on the past. In the second place, suppose that an observer is looking at some random phenomenon which at some initial time is described by a particular measure on the sample path space. If a different

observer were to come along at a later random time, the phenomenon he observed
would no longer be described by the original measure but by some convex combin-
ation of measures conditioned on the past. As a result we feel that it is best to
think of the stochastic process as the collection of all possible measures which
various observers might have to use to describe the phenomenon. Finally, we feel
that there is some reason to believe that every random physical phenomenon is
simply a temporally homogeneous Markov process in disguise. As a result, since
the Markov process is a convex set of measures closed under conditioning on the
past, we feel the original random phenomenon would also have this property.

Our first result is that to any stochastic process one can associate a unique
algebraic structure whose main component consists of a semigroup of operators
on a linear space. This association is well known in the case of temporally homo-
geneous Markov processes. The interested reader can refer to [6] to see how hit-
ting time distributions for general stochastic processes can be found, using the
algebraic representation, as the solution of an abstract Dirichlet problem.

Having defined a stochastic process and having constructed an algebraic
representation for it, we then concern ourselves with the problem of representing
the stochastic process as a function of a temporally homogeneous Markov process.
We feel that this problem is an important one. In the first place, once having re-
·duced a stochastic process to a function of a particular temporally homogeneous
Markov process, we can then use the well-developed theory of Markov processes
to study the original process. In the second place, we feel that in nature stochas-
tic processes are Markovian and that the only reason we observe non-Markovian
processes is that we are not observing enough; that there is some other variable
which, taken together with ours, would make the process Markovian. For example,
the number of cancerous cells in the body, as a function of time, is probably not
Markovian. Yet one feels that there is probably some other observable, such as
the amount of some substance in the body, which together with the number of
cancerous cells would form a temporally homogeneous Markov process. This other
observable would then be treated as the cause of the cancerous growth and its
regulation would then hopefully regulate the growth of the cancer. We feel that
finding a Markov process representation of a given stochastic process is equiv-
alent to finding what is causing the process to behave as it does. Unfortunately
one can find many Markov process representations of a given stochastic process.
The problem is to find the simplest. In [5] we gave a method for constructing
Markov process representations for general stochastic processes, but we did not
show that the resulting representation was the best. In this paper we have divided
stochastic processes into simple and nonsimple ones. We then order all Markov
representations for a given stochastic process, and show first that for simple
processes there is a unique minimal Markov process representation and that it is

precisely the one constructed in [5]. We then show for discrete time stochastic processes, whether simple or not, that the minimal Markov process representations can be constructed using a variation of the construction used in [5]. Unfortunately, this representation need not be unique for nonsimple stochastic processes.

Finally we introduce classes of stochastic processes which we call Ξ processes. These are generalizations of temporally homogeneous Markov processes and include the basic Bush-Mosteller [1] model for learning. Their defining property is that past and future are independent given the present and the conditional distribution, on the past, of a collection Ξ of random variables in the future. We describe the algebraic representations of Ξ processes and then construct explicitly the minimal Markov process representation for a particular example.

1. Definitions and statement of results. Let $\mathfrak{M} = \mathfrak{M}(\Omega, \mathfrak{F})$ be the space of all finite valued signed measures on the measurable space $(\Omega, \mathfrak{F})$ of all functions ω mapping a subset T of $[0, \infty)$ into a measurable space $(\mathcal{S}, \Sigma)$ where $\mathfrak{F}$ is the σ-field generated by the events $X(t)(\omega) = \omega(t) \in U \in \Sigma$ and where $\mathcal{S}$ is a separable compact metric space with Borel sets Σ. Let $T(t)$, $t \in T$, be the semigroup of linear operators on $\mathfrak{M}$ defined by $T(t)\mu(X(t_1) \in U_1, \cdots, X(t_n) \in U_n) = \mu(X(t + t_1) \in U_1, \cdots, X(t + t_n) \in U_n)$; $E(U)$, $U \in \Sigma$, be the resolution of the identity on $\mathfrak{M}$ defined by

$$E(U)\mu(\Lambda) = \mu(X(0) \in U, \Lambda),$$

and p^* be the linear functional on $\mathfrak{M}$ defined by $p^*\mu = \mu(\Omega)$. We define a stochastic process to be any convex subset Φ_+ of probability measures in $\mathfrak{M}$ which has the property that the positive cone which it generates is invariant under $T(t)$, $t \in T$, and $E(U)$, $U \in \Sigma$, and closed in the weak* topology over the continuous functions in the product topology of Ω. We denote by Φ the smallest linear space in $\mathfrak{M}$ containing Φ_+ and we define the algebraic representation of the stochastic process Φ_+ to be the collection $X = (\Phi, \Phi_+, T(t), E(U), p^*)$. Clearly if $\phi \in \Phi_+$, then we can calculate the joint distributions of ϕ from the algebraic representation X using the easily verifiable fact that

$$\phi(X(t_1) \in U_1, \cdots, X(t_1 + \cdots + t_n) \in U_n) = p^* E(U_n) T(t_n) \cdots E(U_1) T(t_1)\phi.$$

Two collections

$$X = (\Phi, \Phi_+, T(t), E(U), p^*) \quad \text{and} \quad \hat{X} = (\hat{\Phi}, \hat{\Phi}_+, \hat{T}(t), \hat{E}(U), \hat{p}^*)$$

will be called isomorphic if there exists a one-to-one linear mapping γ of Φ onto $\hat{\Phi}$ such that

(i) $\gamma\Phi_+ = \hat{\Phi}_+,$

(ii) $\gamma T(t)\phi = \hat{T}(t)\gamma\phi,$

(iii) $\gamma E(U)\phi = \hat{E}(U)\gamma\phi$,

(iv) $p^*\phi = \hat{p}^*\gamma\phi$.

With this in mind we have the following theorem.

Theorem I. *A necessary and sufficient condition that a collection $X = (\Phi, \Phi_+, T(t), E(U), p^*)$ be isomorphic to the algebraic representation of a stochastic process is that:*

(1) *Φ is a linear space generated by the convex subset Φ_+.*

(2) *$T(t)$, $t \in T$, is a semigroup of linear operators on Φ with $T(0) = I$.*

(3) *$E(U)$, $U \in \Sigma$, is a family of linear operators on Φ which is a resolution of the identity. That is to say,*

$$E(U \cup V) = E(U) + E(V) \quad if \; U \cap V = \varnothing, \; E(U \cap V) = E(U)E(V) \; and \; E(\mathcal{S}) = I.$$

(4) *p^* is a linear functional on Φ for which $p^*T(t)\phi = p^*\phi$, $p^*E(U_n)T(t_n) \cdots E(U_1)T(t_1)\phi$ is continuous under monotone limits in $U_1, \cdots, U_n$ and $\phi = 0$ whenever $p^*E(U_n)T(t_n) \cdots E(U_1)T(t_1)\phi = 0$ for all $U_1, \cdots, U_n$ in Σ and $t_1, \cdots, t_n$ in T.*

(5) *The positive cone $R_+\Phi_+$ is invariant under the operators $T(t)$, $t \in T$, and $E(U)$, $U \in \Sigma$, and for each $\phi \in \Phi_+$, $p^*\phi = 1$ and $p^*E(U_n)T(t_n) \cdots E(U_1)T(t_1)\phi \geq 0$. Furthermore Φ_+ is closed in the topology generated by the neighbourhoods*

$$N(\epsilon, f_1, \cdots, f_m, t_1, \cdots, t_n)(\phi)$$
$$= \left\{\psi: \left|\int \cdots \int f_i(u_1, \cdots, u_n)p^*E(du_n)T(t_n) \cdots E(du_1)T(t_1)(\phi - \psi)\right| < \epsilon, \right.$$
$$\left. i = 1, \cdots, m\right\}$$

where $f_1, \cdots, f_m$ are continuous functions on $\mathcal{S}^n$, $t_1, \cdots, t_n$ are in T and $\epsilon > 0$.

For example, any temporally homogeneous strong Markov process will be a stochastic process under the above definition. Indeed, if $\mathcal{S}$ is the state space, then up to an isomorphism Φ would be the set of all measures on $\mathcal{S}$, Φ_+ would be the set of all probability measures on $\mathcal{S}$, $T(t)$ wuold be the semigroup of operators on Φ defined by

$$(T(t)\phi)(U) = \int \phi(du)P_t(u, U)$$

where P_t is the transition probability function, $E(U)$ would be the resolution of the identity on Φ defined by $(E(V)\phi)(U) = \phi(U \cap V)$ and p^* would be the linear functional on Φ defined by $p^*\phi = \phi(\mathcal{S})$.

A temporally homogeneous Markov process Y on a measurable state space $(\mathcal{T}, \mathcal{B})$ will be called a perfect representation of the stochastic process X if

there exist measurable functions f and θ mapping $\mathfrak{I}$ into $\mathfrak{S}$ and the set $\mathcal{P}(\mathfrak{I})$ of all probability measures on $\mathfrak{I}$ onto Φ_+ respectively such that, for all $m \in \mathcal{P}(\mathfrak{I})$,

$$\theta m(X(t_1) \in U_1, \cdots, X(t_n) \in U_n) = P_m\{f \circ Y(t_1) \in U_1, \cdots, f \circ Y(t_n) \in U_n\}$$

where P_m is the distribution of $Y(t)$, $t \in T$, when the initial distribution is m. In other words, the measures in Φ_+ can be interpreted as being precisely those measures derived from the temporally homogeneous Markov process Y with various initial distributions. That is to say, any initial distribution of Y yields via f a measure in Φ_+ and any measure in Φ_+ derives from a particular initial distribution of Y. We will call a perfect representation measurable if there exists a measurable subset $\mathfrak{I}'$ of $\mathfrak{I}$ such that θ restricted to $\mathfrak{I}'$ is a one-to-one mapping of $\mathfrak{I}'$ onto the set Φ_{++} of extreme points of Φ_+. The set $\mathcal{P}_X$ of all perfect measurable representations of X can be partially ordered by letting $Y' \leq Y$ if:

(1) The state space $(\mathfrak{I}', \mathfrak{B}')$ of Y' is contained in the state space $(\mathfrak{I}, \mathfrak{B})$ of Y in the sense that $\mathfrak{I}' \subset \mathfrak{I}$ and $\mathfrak{B}' \subset \mathfrak{B}$, f' is the restriction of f to Y' and θ' is the restriction of θ to $\mathcal{P}(\mathfrak{I}')$.

(2) For each $b \in \mathfrak{I} \backslash \mathfrak{I}'$ there exists a probability measure $M(b, \cdot)$ on $\mathfrak{I}'$ such that whenever $a \in \mathfrak{I}'$ and U is a measurable set in $\mathfrak{I}'$, the transition probability functions Q_t and Q_t' of Y and Y' are related by

$$Q_t'(a, U) = Q_t(a, U) + \int_{\mathfrak{I} \backslash \mathfrak{I}'} Q_t(a, db) M(b, U)$$

or, if the time parameter is discrete, by

$$Q_1'(a, U) = Q_1(a, U) + \int_{\mathfrak{I} \backslash \mathfrak{I}'} Q_1(a, db) M(b, U).$$

Intuitively, this means that Y' can be obtained from Y by taking each state $b \in \mathfrak{I} \backslash \mathfrak{I}'$, splitting it into pieces and then assigning each piece to one of the states in $\mathfrak{I}'$, where the proportion of b that is assigned to states in U is $M(b, U)$.

If $X = (\Phi, \Phi_+, T(t), E(U), p^*)$ is a stochastic process with a countable time parameter T and if Φ_+ is a simplex, then we shall say that X is simple. When X is simple it follows from Choquet's theorem (see [5], [7]) that for each ϕ in the set of extreme points Φ_{++} of Φ_+ and each $t \in T$ there exist unique probability measures $P_t^*(\phi, \cdot)$ on Φ_{++} for which

$$T(t)\phi = \int_{\Phi_{++}} \psi \, P_t^*(\phi, d\psi).$$

Since the probability measures $P_t^*(\phi, \cdot)$ are uniquely defined and since $T(s + t) = T(s)T(t)$, it follows that

$$P_{s+t}^*(\phi, \cdot) = \int_{\Phi_{++}} P_s^*(\phi, d\psi) P_t^*(\psi, \cdot)$$

and so we construct a temporally homogeneous Markov process X^* on Φ_{++} with transition probability function P_t^*. If we let f and θ be defined by the equations: $E(\{f\phi\})\phi = \phi$ and $\theta m = \int_{\Phi_{++}} \phi m(d\phi)$, then it follows from [5] that for each $\phi \in \Phi_+$,

$$\theta m\{X(t_1) \in U_1, \cdots, X(t_n) \in U_n\} = P_m\{f \circ Y(t_1) \in U_1, \cdots, f \circ Y(t_n) \in U_n\}.$$

We shall call X^* the dual of X. Clearly X^* is a measurable perfect representation of X.

Theorem II. *If the stochastic process X is simple, then X^* is, up to an isomorphism, the only minimal measurable perfect representation of X.*

The reader should note that if X is a temporally homogeneous Markov process, then X is simple and Φ_{++} is isomorphic to the set of all probability measures on $\mathcal{S}$ whose support consists of a single point. In this case Φ_{++} can be identified with $\mathcal{S}$ and P_t^*, with this identification, is P_t. In other words, $X = X^*$.

If $X = (\Phi, \Phi_+, T(t), E(U), p^*)$ is a stochastic process which is not simple, then the above procedure might break down because P_t, $t \in T$, need not be a transition probability function. However, we can get around this difficulty when the time parameter is the nonnegative integers by defining the n-step transition probability function P_n^* in terms of the one-step transition probability function P^*. Indeed, we know from Choquet's theorem that to each $\phi \in \Phi_+$ there exists a probability measure Q_ϕ on Φ_{++} such that $\phi = \int_{\Phi_{++}} \psi Q_\phi(d\psi)$. When Q can be chosen so that for each measurable set $U \in \Phi_{++}$, $Q_\phi(U)$ is measurable in ϕ, we will say that X is measurable. It follows that when X is measurable there exist probability measures $P^*(\phi, \cdot)$ on Φ_{++} for which

$$T\phi = \int_{\Phi_{++}} \psi\, P^*(\phi, d\psi), \qquad T = T(1),$$

and for which $P^*(\cdot, U)$ is measurable for each measurable U in Φ_{++}. We define the dual X^* of X to be the temporally homogeneous Markov process on Φ_{++} with one step transition probability function P^*. It should be noted that when X is not simple, X^* is not the same representation that was obtained in [5]. Taking f and θ as before we have

Theorem III. *If X is a measurable discrete time stochastic process, then a necessary and sufficient condition that a measurable perfect representation of X be minimal is that it be isomorphic to a version of X^*.*

If Φ_+ consists of the set of all $\phi \in \Phi$ for which

$$p^*E(U_n)T(t_n) \cdots E(U_1)T(t_1)\phi \geq 0, \qquad p^*\phi = 1,$$

for all $U_1, \cdots, U_n$ in Σ and $t_1, \cdots, t_n$ in T, then X will be called an outer

stochastic process and its algebraic representation will be written as $X = (\Phi, T(t), E(U), p^*)$ since in that case Φ_+ is uniquely determined by Φ, $T(t), E(U)$ and p^*. On the other hand, if there is a $\phi_0 \in \Phi$ such that Φ_+ is generated by the closure of the convex hull of the points $E(U_n)T(t_n) \cdots E(U_1)T(t_1)\phi_0$, where $U_1, \cdots, U_n$ are in Σ and $t_1, \cdots, t_n$ are in T, then X will be called the inner stochastic process of ϕ_0 and its algebraic representation will be written $X = X(\phi_0) = (\Phi, T(t), E(U), p^*, \phi_0)$. To each inner stochastic process one can associate the outer stochastic process $\overline{X} = (\Phi, T(t), E(U), p^*)$. The reader should note that $\overline{\Phi}_{++}$ is not in general equal to Φ_{++} and $\overline{X}^*$ is not in general equal to X^*. An example of this phenomenon is given in §7.

Suppose that Ξ is a collection of bounded measurable real valued functions on $(\Omega, \mathcal{F})$ which includes the identity $1(\omega) = 1$. An outer stochastic process $X = (\Phi, T(t), E(U), p^*)$ will be called a Ξ process if Φ_+ is the weak* closure of the set of all $\phi \in \Phi_+$ for which, for any $t \in T$ and $\Lambda \in \mathcal{F}$,

$$\phi(\Lambda_t^+ | \mathcal{F}_t) = \phi[\Lambda_t^+ | X(t), E(\phi)(\xi_t^+ | \mathcal{F}_t), \quad \xi \in \Xi].$$

Note that $\mathcal{F}_t$ is the σ-field generated by $X(u)$, $u \leq t$, $\xi_t^+(\omega) = \xi(\omega_t^+)$ where $\omega_t^+(u) = \omega(t + u)$ and $\Lambda_t^+ = [\omega: \omega_t^+ \in \Lambda]$. Also $E(\phi)$ denotes the expectation using the measure ϕ.

Theorem IV. *A necessary and sufficient condition that an outer stochastic process* $\hat{X} = (\hat{\Phi}, \hat{T}(t), \hat{E}(U), \hat{p}^*)$ *be a* Ξ *process is that* $\hat{X}$ *be isomorphic to a collection* $X = (\Phi, T(t), E(U), p^*)$ *where*

(1) Φ *is a linear subspace of the set of all functions mapping* $\Sigma \times \Xi$ *into* R *which are such that, for each* $\xi \in \Xi$, $\phi(\cdot, \xi)$ *is a measure.*

(2) $E(V)\phi(U, \xi) = \phi(UV, \xi)$.

(3) $p^*\phi = \phi(\delta, 1)$.

(4) $E(\hat{\phi})(\xi I_U(X_0)) = \phi(U, \xi)$.

Note that if $X = (\Phi, T(n), E(U), p^*)$ is a discrete time Ξ process, then we can define the nonlinear operators $R(x)$, $x \in \delta$, via

$$R(x)\phi(U, \xi) = E(\phi)(\xi_1^+ I_U(X(1)) | X(0) = x)$$

and it then follows that

$$E(\phi)(\xi_{n+1}^+ I_U(X(n+1)) | X(1) = x_1, \cdots, X(n) = x_n) = R(x_n) \cdots R(x_1)\phi(U, \xi)$$

and so

$$\phi(X(0) \in U_0, \cdots, X(n) \in U_n)$$
$$= \int_{U_0} \phi(dx_0, 1) \int_{U_1} R(x_0)\phi(dx_1, 1) \cdots \int_{U_n} R(x_{n-1}) \cdots R(x_0)\phi(dx_n, 1).$$

In other words, X is generated by first picking a point x_0 according to the distribution $\phi(\cdot, 1)$, then picking x_1 according to the distribution $R(x_0)\phi(\cdot, 1)$ and then x_2 according to the distribution $R(x_1)R(x_0)\phi(\cdot, 1)$ and so on. The reader can refer to [1], [3] and [4] for examples of stochastic processes similar to the Ξ processes defined here.

As an example, consider the discrete time Ξ process where $\Xi = \Xi_0$ is the set of all functions ξ of the form $\xi(\omega) = I_U(X(1)(\omega))$ where I_U is the indicator of the Borel set U in R_+. In this case $\delta = R_+$ and Σ is the Borel subset of R_+ and so we can identify Φ with the linear space of functions Φ mapping $\Sigma \times \Sigma \rightarrow R$. We will assume that all of the functions $\phi \in \Phi$ are absolutely continuous in their second varaible with respect to the Lebesgue measure on R_+ and we shall simply write the derivative of $\phi(U, \cdot)$ with respect to the Lebesgue measure on R_+ as $\phi'(U, v)$. We will also assume that $\phi'(U, \cdot) \in C^\infty(R_+)$ for each $U \in \Sigma$ and $\phi \in \Phi$. $\phi'(u, v)$ will denote the density of ϕ with respect to the Lebesgue measure on $R_+ \times R_+$ when it exists. If $\phi(\{a\}, \delta) = 1$ for some $a \in \delta$, then $\phi'(u, v)$ is to be $\phi'(\delta, v)I_{\{a\}}(u)$. Finally, let B be a bounded measurable function mapping $R_+ \times R_+ \times R_+ \rightarrow R_+$ which is a completely monotone function in its third variable and for which $B(x, y, y) = 1$ for all x and y in R_+. It follows that there exist probability measures $\alpha(x, y, \cdot)$ on R_+ such that

$$B(x, y, z) = \int_0^\infty e^{zt}\alpha(x, y, dt).$$

Theorem V. *Let* $X = (\Phi, T, E(U), p^*)$ *be a* Ξ_0 *process with*

$$T\phi(U, V) = -\int_V dv \int_U du \int_0^\infty \frac{\partial}{\partial v}[B(x, u, u + v)\phi'(dx, u + v)].$$

Then:

(1) $$(R(x)\phi)'(u, v) = -\frac{\partial}{\partial v}[B(x, u, u + v)\phi'(x, u + v)]\Big/\int_0^\infty \phi'(x, z)\,dz.$$

(2) *The mapping*

$$(a, b) \rightarrow \phi_{a,b}(U, V) = \int_V e^{-bv}\,dv, \quad a \in U,$$
$$= 0, \quad a \notin U,$$

is a bijection of $R_+ \times R_+$ *onto* Φ_{++} *and* $f(a, b) = a.$

(3) X^* *has transition probability function*

$$P(a, b: U, V) = \int_U du \int_{(V-b)\cap[0,\infty]} e^{-u(v+b)}\alpha(a, u, dv).$$

(4) *A necessary and sufficient condition that* ϕ *be a version of* X (*i.e.* $\phi \in \Phi_+$) *is that for any* $0 \le x \le y,$

$$\phi(X(n + 1) < x, X(0), \cdots, X(n), X(n + 1) < y < X(n + 1) + X(n + 2))$$

$$= \int_0^x B(X(n), \xi, y) \, d\xi \Big/ \int_0^y B(X(n), \xi, y) \, d\xi.$$

2. Proof of Theorem I. The necessity of conditions (1) through (5) is clear. To prove the sufficiency, suppose that $X = (\Phi, \Phi_+, T(t), E(U), p^*)$ satisfies conditions (1) through (5). To each $\phi \in \Phi_+$ for which $p^*\phi = 1$ we get, via conditions (1) through (5) and the Kolmogorov extension theorem, a probability measure $\gamma\phi \in \mathfrak{M}$ where

$$\gamma\phi(X(t_1) \in U_1, \cdots, X(t_1 + \cdots + t_n) \in U_n) = p^* E(U_n) T(t_n) \cdots E(U_1) T(t_1)\phi.$$

Clearly $\hat{X} = (\hat{\Phi}, \hat{\Phi}_+, \hat{T}(t), \hat{E}(U), \hat{p}^*)$, where $\hat{\Phi} = \gamma\Phi$, $\hat{T}(t)\gamma\phi = \gamma T(t)\phi$, $\hat{E}(U)\gamma\phi = \gamma E(U)\phi$ and $\hat{p}^*\gamma\phi = p^*\phi$ is a stochastic process. Furthermore, γ is an isomorphism of X onto $\hat{X}$. For example, γ is one-to-one since $\gamma\phi = \gamma\psi$ implies that $p^* E(U_n) T(t_n) \cdots E(U_1) T(t_1)(\phi - \psi) = 0$ for all choices of $U_1, \cdots, U_n$ in Σ and $t_1, \cdots, t_n$ in T and so, by condition (4) of Theorem 1, $\phi = \psi$.

3. Proof of Theorem II. We first show that $\leq$ is indeed a partial ordering of P_X. Suppose that $Y \leq Y' \leq Y''$. Then $\mathfrak{J} \subseteq \mathfrak{J}' \subseteq \mathfrak{J}''$ and there exist measures $M(b, \cdot)$ for each $b \in \mathfrak{J}' \backslash \mathfrak{J}$ and $M'(b, \cdot)$ for each $b \in \mathfrak{J}'' \backslash \mathfrak{J}'$ for which

$$Q_t(a, U) = Q_t'(a, U) + \int_{\mathfrak{J}' \backslash \mathfrak{J}} Q_t'(a, db)M(b, U)$$

and

$$Q_t'(a, U) = Q_t''(a, U) + \int_{\mathfrak{J}'' \backslash \mathfrak{J}'} Q_t''(a, db)M'(b, U).$$

Setting $M(b, \cdot) = 0$ for $b \in \mathfrak{J}' \backslash \mathfrak{J}$ and $M'(b, \cdot) = 0$ for $b \in \mathfrak{J}'' \backslash \mathfrak{J}'$ gives us

$$Q_t(a, U) = Q_t''(a, U) + \int_{\mathfrak{J}'' \backslash \mathfrak{J}'} Q_t''(a, db)M'(b, U)$$

$$+ \int_{\mathfrak{J}' \backslash \mathfrak{J}} Q_t''(a, db)M(b, U) + \int_{\mathfrak{J}'' \backslash \mathfrak{J}'} Q_t''(a, db) \int_{\mathfrak{J}' \backslash \mathfrak{J}} M'(b, dc)M(c, U)$$

$$= Q_t''(a, U) + \int_{\mathfrak{J}'' \backslash \mathfrak{J}} Q_t''(a, db)M''(b, U)$$

where

$$M''(b, U) = M'(b, U) + M(b, U) + \int_{\mathfrak{J}' \backslash \mathfrak{J}} M'(b, dc)M(c, U).$$

Thus $Y \leq Y' \leq Y''$ implies that $Y \leq Y''$. Since $\leq$ is reflexive and antisymmetric, it follows that $(\mathcal{P}_X, \leq)$ is a partially ordered system. Suppose that Y is a measurable perfect representation of X. Then Y is a temporally homogeneous Markov process on some state space $(\mathfrak{J}, \mathcal{B})$ with transition probability function P_t, and

there are functions f and θ mapping $\mathcal{J}$ into $\mathcal{S}$ and $\mathcal{P}(\mathcal{J})$ onto Φ_+ respectively such that, for each $m \in \mathcal{P}(\mathcal{J})$,

$$\theta m\{X(t_1) \in U_1, \cdots, X(t_n) \in U_n\} = P_m\{f \circ Y(t_1) \in U_1, \cdots, f \circ Y(t_n) \in U_n\}.$$

Since Y is a measurable representation of X, we know that there exists a measurable subset $\mathcal{J}'$ of $\mathcal{J}$ consisting of exactly one point from each of the sets $\theta^{-1}\phi$, $\phi \in \Phi_{++}$. For each $c \in \mathcal{J}$, let m_c be the probability measure on $(\mathcal{J}, \mathcal{B})$ whose support is $\{c\}$. We know from Choquet's theorem, together with the fact that X is simple, that there exists for each $c \in \mathcal{J}\backslash\mathcal{J}'$ a unique probability measure $Q_c(\cdot)$ on $\mathcal{J}'$ for which

$$\phi_c = \int_{\mathcal{J}'} \phi_b Q_c(db), \qquad \phi_c = \theta m_c.$$

Let P_t' be defined by

$$P_t'(a, U) = P_t(a, U) + \int_{\mathcal{J}\backslash\mathcal{J}'} P_t(a, db)Q_b(U).$$

Then for any $a \in \mathcal{J}'$, we have

$$p^* E(U_n)T(t_n) \cdots E(U_1)T(t_1) \int_{\mathcal{J}} \phi_b P_t(a, db)$$

$$= \int_{\mathcal{J}} p^* E(U_n)T(t_n) \cdots E(U_1)T(t_1)\phi_b P_t(a, db)$$

$$= \int_{\mathcal{J}} \phi_b(X(t_1) \in U_1, \cdots, X(t_1 + \cdots + t_n) \in U_n) P_t(a, db)$$

$$= \int_{\mathcal{J}} P_{m_b}(f \circ Y(t_1) \in U_1, \cdots, f \circ Y(t_1 + \cdots + t_n) \in U_n) P_t(a, db)$$

$$= P_{m_a}(f \circ Y(t + t_1) \in U_1, \cdots, f \circ Y(t + t_1 + \cdots + t_n) \in U_n)$$

$$= \phi_a(X(t + t_1) \in U_1, \cdots, X(t + t_1 + \cdots + t_n) \in U_n)$$

$$= p^* E(U_n)T(t_n) \cdots E(U_1)T(t_1)T(t)\phi_a$$

$$= p^* E(U_n)T(t_n) \cdots E(U_1)T(t_1) \int_{\Phi_{++}} \psi P_b^*(\phi_a, d\psi).$$

Thus $\int_{\mathcal{J}} \phi_b P_t(a, db) = \int_{\Phi_{++}} \psi P_t^*(\phi_a, d\psi)$ and so

$$\int_{\mathcal{J}'} \phi_b P_t'(a, db) = \int_{\mathcal{J}'} \phi_b P_t(a, db) + \int_{\mathcal{J}'} \phi_b \left[\int_{\mathcal{J}\backslash\mathcal{J}'} P_t(a, dc)Q_c(db) \right]$$

$$= \int_{\mathcal{J}'} \phi_b P_t(a, db) + \int_{\mathcal{J}\backslash\mathcal{J}'} \left[\int_{\mathcal{J}'} \phi_b Q_c(db) \right] P_t(a, dc)$$

$$= \int_{\mathcal{J}'} \phi_b P_t(a, db) + \int_{\mathcal{J}\backslash\mathcal{J}'} \phi_c P_t(a, dc)$$

$$= \int_{\mathcal{J}} \phi_b P_t(a, db) = \int_{\Phi_{++}} \psi P_t^*(\phi_a, d\psi).$$

Since Φ_+ is a simplex, it follows that $P_t'(a, db) = P_t^*(\phi_a, d\phi_b)$ and so $X^* = Y'$ $\leq Y$. Since X^* is a perfect representation of X, see [5], it follows that X^* is the only minimal perfect representation of X and the theorem is proved.

4. Proof of Theorem III.

Since, for each $m \in \mathcal{P}(\Phi_{++})$,

$$
\begin{aligned}
(\theta m)(X(0) \in U_0, \cdots, X(n) \in U_n) &= p^* E(U_n) T \cdots E(U_1) T E(U_0)(\theta m) \\
&= p^* E(U_n) T \cdots E(U_1) T E(U_0) \int_{\Phi_{++}} \psi_0 \, m(d\psi_0) \\
&= p^* E(U_n) T \cdots E(U_1) T \int_{f^{-1}U_0} \psi_0 \, m(d\psi_0) \\
&= \int_{f^{-1}U_0} m(d\psi_0) p^* E(U_n) T \cdots E(U_1) T \psi_0 \\
&= \int_{f^{-1}U_0} m(d\psi_0) p^* E(U_n) T \cdots E(U_1) \int_{\Phi_{++}} \psi_1 P^*(\psi_0, d\psi_1) \\
&= \int_{f^{-1}U_0} m(d\psi_0) p^* E(U_n) T \cdots E(U_2) T \int_{f^{-1}U_1} \psi_1 P^*(\psi_0, d\psi_1) \\
&= \int_{f^{-1}U_0} m(d\psi_0) \int_{f^{-1}U_1} P^*(\psi_0, d\psi_1) p^* E(U_n) T \cdots E(U_2) T \psi_1 \\
&= \int_{f^{-1}U_0} m(d\psi_0) \int_{f^{-1}U_1} P^*(\psi_0, d\psi_1) \cdots \int_{f^{-1}U_n} P^*(\psi_{n-1}, d\psi_n) \\
&= P_m(f \circ X^*(0) \in U_0, \cdots, f \circ X^*(n) \in U_n)
\end{aligned}
$$

it follows that X^* is a perfect representation of X. Suppose that X^* is not minimal. Then there exists a subset Φ_{++}' of Φ_{++} and a probability measure $M(\phi, \cdot)$ for each $\phi \in \Phi_{++} \backslash \Phi_{++}'$ such that the temporally homogeneous Markov process X' on Φ_{++}' with one step transition probability function

$$
P'(\phi, \cdot) = P^*(\phi, \cdot) + \int_{\Phi_{++} \backslash \Phi_{++}'} P^*(\phi, d\psi) M(\psi, \cdot)
$$

is a perfect representation of X. That is to say,

$$
\theta m(X(t_1) \in U_1, \cdots, X(t_n) \in U_n) = P_m(f' \circ Y(t_1) \in U_1, \cdots, f' \circ Y(t_n) \in U_n)
$$

where f' and θ' are the restrictions of f and θ to Φ_{++}' and $\mathcal{P}(\Phi_{++}')$ respectively. Let $\phi \in \Phi_{++} \backslash \Phi_{++}'$. Then

$$
\begin{aligned}
p^* E(U_n) T \cdots E(U_1) T E(U_0)(\theta m) \\
&= \int_{f'^{-1}U_0} m(d\psi_0) \int_{f'^{-1}U_1} P'(\psi_0, d\psi) \cdots \int_{f'^{-1}U_n} P'(\psi_{n-1}, d\psi_n) \\
&= \int_{f'^{-1}U_0} m(d\psi_0) \psi_0(X(0) \in U_0, \cdots, X(n) \in U_n) \\
&= \int_{f'^{-1}U_0} m(d\psi_0) p^* E(U_n) T \cdots E(U_1) T E(U_0) \psi_0 \\
&= p^* E(U_n) T \cdots E(U_1) T E(U_0) \int_{f'^{-1}U_0} \psi_0 m(d\psi_0).
\end{aligned}
$$

Since this holds for all choices of n, $U_0, \cdots, U_n$, it follows that

$$\phi = \int_{f'^{-1}U_0} \psi\, m(d\psi).$$

But $\phi \in \Phi_{++}$ and $\phi \notin f'^{-1}U_0 \subset \Phi_{++}\backslash\Phi'_{++}$ and so we have a contradiction. Thus X^* is minimal.

On the other hand, suppose that Y is a measurable representation of X. Then there exists a measurable subset $\mathfrak{J}'$ of $\mathfrak{J}$ consisting of exactly one point from each of the sets $\theta^{-1}\phi$, $\phi \in \Phi_{++}$. For each $c \in \mathfrak{J}$, let m_c be the probability measure on $(\mathfrak{J}, \mathfrak{B})$ whose support is $\{c\}$. Since X is measurable we know that for each $c \in \mathfrak{J}\backslash\mathfrak{J}'$ we can find a probability measure $Q_c(\cdot)$ on $\mathfrak{J}'$ which is measurable in c and which has the property that

$$\phi_c = \int_{\mathfrak{J}'} \phi_b Q_c(db), \qquad \phi_c = \theta m_c.$$

Let P' be the one step transition probability function

$$P'(a, U) = P(a, U) + \int_{\mathfrak{J}\backslash\mathfrak{J}'} P(a, db)Q_b(U).$$

Then for any $a \in \mathfrak{J}'$ we have

$$p^* E(U_n)T \cdots E(U_1)TE(U_0)\int_{\mathfrak{J}} \phi_b P(a, db)$$

$$= \int_{\mathfrak{J}} p^* E(U_n)T \cdots E(U_1)TE(U_0)\phi_b P(a, db)$$

$$= \int_{\mathfrak{J}} \phi_b (X(0) \in U_0, \cdots, X(n) \in U_n) P(a, db)$$

$$= \int_{\mathfrak{J}} P_{m_b}(f \circ Y(0) \in U_0, \cdots, f \circ Y(n) \in U_n) P(a, db)$$

$$= P_{m_a}(f \circ Y(1) \in U_0, \cdots, f \circ Y(n + 1) \in U_n)$$

$$= \phi_a(X(1) \in U_0, \cdots, X(n + 1) \in U_n) = p^* E(U_n)T \cdots E(U_0)T\phi_a.$$

Since our algebraic representation is reduced we can conclude that $T\phi_a = \int_{\mathfrak{J}} \phi_b P(a, db)$. Hence

$$\int_{\mathfrak{J}'} \phi_b P'(a, db) = \int_{\mathfrak{J}'} \phi_b P(a, db) \int_{\mathfrak{J}'} \phi_b \left[\int_{\mathfrak{J}\backslash\mathfrak{J}'} P(a, dc)Q_c(db)\right]$$

$$= \int_{\mathfrak{J}'} \phi_b P(a, db) + \int_{\mathfrak{J}\backslash\mathfrak{J}'} \left[\int_{\mathfrak{J}'} \phi_b Q_c(db)\right] P(a, dc)$$

$$= \int_{\mathfrak{J}'} \phi_b P(a, db) + \int_{\mathfrak{J}\backslash\mathfrak{J}'} \phi_c P(a, dc)$$

$$= \int_{\mathfrak{J}} \phi_b P(a, db) = T\phi_a.$$

Since $\mathfrak{J}' \subset \Phi_{++}$ we conclude that Y' is a version of X^* and the proof is completed.

5. Proof of Theorem IV. To prove the sufficiency, suppose that the stochastic process $X = (\Phi, T(t), E(U), p^*)$ satisfies the conditions (1) through (4). To show that X is a Ξ process, we must show for all $\Lambda \in \mathcal{F}$, $\phi \in \Phi_+$ and $t \in T$ that

$$\phi(\Lambda_t^+ | \mathcal{F}_t) = \phi[\Lambda_t^+ | X(t), E(\phi)(\xi_t^+ | \mathcal{F}_t), \xi \in \Xi].$$

But for any choice of $U_1, \cdots, U_{m+n}$ in Σ and $s_1, \cdots, s_{m+n}$ in T, letting $\psi(k) = E(U_k)T(s_k) \cdots E(U_1)T(s_1)\phi$, we have

$$\phi(X(s_1 + \cdots + s_{m+n}) \in U_{m+n}, \cdots, X(s_1 + \cdots + s_m + s_{m+1})$$

$$\in U_{m+1} | X(s_1 + \cdots + s_m) \in U_m, \cdots, X(s_1) \in U_1)$$

$$= p^* \psi(m + n) / p^* \psi(m)$$

$$= p^* E(U_{m+n})T(s_{m+n}) \cdots E(U_{m+1})T(s_{m+1}) \left[\frac{\psi(m)}{p^*\psi(m)}\right].$$

Using (4),

$$\frac{\psi(m)(U, \xi)}{p^*\psi(m)} = E_{\{\psi(m)/p^*\psi(m)\}}[\xi I_U(X(0))]$$

$$= \frac{E_{\{\psi(m)\}}[\xi I_U(X(0))]}{p^*\psi(m)}$$

$$= \frac{E(\phi)[\xi_{s_1 + \cdots + s_m}^+ I_U(X(s_1 + \cdots + s_m)), X(s_1) \in U_1, \cdots, X(s_1 + \cdots + s_m) \in U_m]}{\phi(X(s_1) \in U_1, \cdots, X(s_1 + \cdots + s_m) \in U_m)}$$

$$= E(\phi)[\xi_{s_1 + \cdots + s_m}^+ I_U(X(s_1 + \cdots + s_m)) | X(s_1) \in U_1, \cdots, X(s_1 + \cdots + s_m) \in U_m].$$

Thus

$$\phi(\Lambda_{s_1 + \cdots + s_m}^+ | X(s_1) \in U_1, \cdots, X(s_1 + \cdots + s_m) \in U_m)$$

$$= \phi[\Lambda_{s_1 + \cdots + s_m}^+ | E(\phi)(\xi_{s_1 + \cdots + s_m}^+ I_U(X(s_1 + \cdots + s_m))) |$$

$$X(s_1) \in U_1, \cdots, X(s_1 + \cdots + s_m) \in U_m), \xi \in \Xi, U \in \Sigma]$$

or

$$\phi(\Lambda_t^+ | \mathcal{F}_t) = \phi[\Lambda_t^+ | E(\phi)(\xi_t^+ I_U(X(t)) | \mathcal{F}_t), \xi \in \Xi, U \in \Sigma]$$

$$= \phi[\Lambda_t^+ | I_U(X(t))E(\phi)(\xi_t^+ | \mathcal{F}_t), \xi \in \Xi, U \in \Sigma']$$

$$= \phi[\Lambda_t^+ | X(t), E(\phi)(\xi_t^+ | \mathcal{F}_t), \xi \in \Xi]$$

and so X is a Ξ process.

On the other hand, if $X = (\hat{\Phi}, \hat{T}(t), \hat{E}(U), \hat{p}*)$ is a Ξ process, then $(\hat{\Phi}, \hat{T}(t), \hat{E}(U), \hat{p}*)$ is isomorphic to $(\Phi, T(t), E(U), p*)$ via the mapping $\hat{\phi} \rightarrow \phi$ where

(1) Φ is the set of all functions ϕ mapping $\Sigma \times \Xi$ into $\mathcal{S}$ defined by $\phi(U, \xi) = E(\hat{\phi})(\xi I_U(X(0)), \hat{\phi} \in \hat{\Phi}.$

(2) $T(t)\phi(U, \xi) = E(\hat{\phi})(\xi_t^+ I_U(X(t))).$

(3) $E(V)\phi(U, \xi) = \phi(UV, \xi).$

(4) $p*\phi = \phi(\mathcal{S}, 1).$

Clearly $T(t)$ is well defined on Φ since if $\hat{\phi} = \hat{\psi}$, then $\phi = \psi$ since X is a Ξ process. Also one clearly has

$$p^* E(U_n)T(t_n) \cdots E(U_1)T(t_1)\phi$$

$$= \phi(X(t_1) \in U_1, \cdots, X(t_1 + \cdots + t_n) \in U_n) = \hat{p}^* \hat{E}(U_n)\hat{T}(t_n) \cdots \hat{E}(U_1)\hat{T}(t_1)\hat{\phi}$$

and so the isomorphism is proved.

6. **Proof of Theorem V.** We first note that T satisfies condition 4 of Theorem IV since if $\xi = I_V(X(1))$, then

$$E(\phi)(\xi I_U(X(0)) = E(\phi)(I_V(X(1))I_U(X(0)) = \phi(X(0) \in U, X(1) \in V)$$

$$= p^* E(V)TE(U)\phi = E(V)TE(U)\phi(\mathcal{S}, \mathcal{S}) = TE(U)\phi(V, \mathcal{S})$$

$$= -\int_0^\infty dv \int_V du \int_0^\infty \frac{\partial}{\partial v} [B(x, u, u + v)E(U)\phi'(dx, u + v)]$$

$$= -\int_0^\infty dv \int_V du \int_U \frac{\partial}{\partial v} [B(x, u, u + v)\phi'(dx, u + v)]$$

$$= \int_V du \int_U \phi'(dx, u) = \phi(U, V) = \phi(U, \xi).$$

Thus X is a Ξ_0 process provided that Φ_+ is closed. We will show this later. Part (1) of the theorem is obvious. To prove part (2), we note that

$$TE(U_n) \cdots TE(U_0)\phi(U, V)$$

$$= (-1)^n \int_V dv \int_U du \int_{U_n} \cdots \int_{U_0} \frac{\partial}{\partial v} B(x_n, v, v + u)$$

$$\frac{\partial}{\partial v} B(x_{n-1}, x_n, x_n + u + v) \frac{\partial}{\partial v} B(x_{n-2}, x_{n-1}, x_{n-1} + x_n + u + v) \cdots$$

$$\cdot \frac{\partial}{\partial v} B(x_0, x_1, x_1 + \cdots + x_n + u + v)\phi'(dx_0, x_1 + \cdots + x_n + u + v)dx_1 \cdots dx_n$$

and so a necessary and sufficient condition that $\phi \in \Phi_+$ is that $p*\phi = 1$ and for all $U, U_0, \cdots, U_n$

$$p^* E(U) TE(U_n) \cdots TE(U_0)\phi = TE(U_n) \cdots TE(U_0)\phi(U, \delta)$$

$$(1) \quad = (-1)^{n-1} \int_U du \int_{U_n} \cdots \int_{U_0} \frac{\partial}{\partial u} B(x_{n-1}, x_n, x_n + u) \frac{\partial}{\partial u} B(x_{n-2}, x_{n-1} + x_n + u)$$

$$\cdots \frac{\partial}{\partial u} B(x_0, x_1, x_1 + \cdots + x_n + u + v)\phi'(dx_0, x_1 + \cdots + x_n + u)dx_1 \cdots dx_n \geq 0$$

or equivalently that

$$(-1)^{n-1} \int_{U_0} \frac{\partial}{\partial u} B(x_{n-1}, x_n, x_n + u) \frac{\partial}{\partial u} B(x_{n-2}x_{n-1}x_{n-1} + x_n + u)$$

$$\cdots \frac{\partial}{\partial u} B(x_0, x_1, x_1 + \cdots + x_n + u)\phi'(dx_0, x_1 + \cdots + x_n + u) \geq 0.$$

Letting $x_2 = \cdots = x_n = u = 0$ we see that a necessary condition that $\phi \in \Phi_+$ is that for all nonnegative integers n,

$$[-\partial/\partial x]^n \phi'(U, x) \geq 0.$$

Thus $\phi'(U, \cdot)$ is a completely monotone function on R_+. Conversely, if $\phi'(U, \cdot)$ is a completely monotone function, then it is clearly in Φ_+ since the completely monotone functions are closed under products and the operators $-\partial/\partial x$. It follows, therefore, that Φ_{++} consists of the functions

$$\phi_{a,b}(U, V) = \int_V e^{-bv}, \quad a \in U,$$
$$= 0, \quad a \notin U,$$

and so the proof of part (2) is complete. Part (3) follows, letting $\alpha_b(a, y, U) = \alpha_{a,y}(U - b)$, from the fact that

$$\int_0^\infty e^{-\eta v} P(a, b; U, d\eta) = \int_0^\infty \int_0^\infty \phi_{\xi,\eta}(U, v) P(a, b; d\xi, d\eta)$$

$$= (T\phi_{a,b})'(U, v) = - \int_U \frac{\partial}{\partial v} B(a, u, u + v)e^{-b(u+v)} dy$$

$$= - \int_U \frac{\partial}{\partial v} \int_0^\infty e^{-(b+\eta)(u+v)}\alpha(a, u, d\eta) du$$

$$= \int_U \int_0^\infty (b + \eta)e^{-(b+\eta)(u+v)}\alpha(a, u, d\eta) du$$

$$= \int_U \int_b^\infty \eta e^{-\eta u} e^{-\eta v} \alpha_b(a, u, d\eta) du$$

$$= \int_0^\infty e^{-\eta v} I_{[b,\infty)}(\eta) \int_U e^{-\eta u}\alpha_b(a, u, d\eta) du$$

which implies that

$$P(a,\ b;\ U,\ V) = \int_{V \cap [b,\infty)} \int_U e^{-\eta u} \alpha_b(a,\ u,\ d\eta)\ du$$

$$= \int_U du \int_{(V-b) \cap [0,\infty)} e^{-u(\eta+b)} \alpha(a,\ u,\ d\eta).$$

To prove part (4), we assume first that ϕ is a version of X. Then for all $0 \leq x \leq y$

$$\phi(X(n+1) < x | X(0) = x_0,\ \cdots,\ X(n) = x_n,\ X(n+1) < y < X(n+1) + X(n+2))$$

$$= \frac{(-1)^n \int_0^x dx_{n+1} \int_{y-x_{n+1}}^\infty \frac{\partial}{\partial x_{n+2}} B(x_n,\ x_{n+1},\ x_{n+1}+x_{n+2}) \cdots \phi'(dx_0,\ x_1 + \cdots + x_{n+2})\,dx_{n+2}}{(-1)^n \int_0^y dx_{n+1} \int_{y-x_{n+1}}^\infty \frac{\partial}{\partial x_{n+2}} B(x_n,\ x_{n+1},\ x_{n+1}+x_{n+2}) \cdots \phi'(dx_0,\ x_1 + \cdots + x_{n+2})\,dx_n}$$

$$= \frac{\int_0^x B(x_n,\ x_{n+1},\ y) \frac{\partial}{\partial y} B(x_{n-1},\ x_n,\ x_n+y) \cdots \phi'(dx_0,\ x_1 + \cdots + x_n + y)\,dx_{n+1}}{\int_0^y B(x_n,\ x_{n+1},\ y) \frac{\partial}{\partial y} B(x_{n-1},\ x_n,\ x_n+y) \cdots \phi'(dx_0,\ x_1 + \cdots + x_n + y)\,dx_{n+1}}$$

$$= \int_0^x B(x_n,\ x_{n+1},\ y)\,dx_{n+1} \Big/ \int_0^y B(x_n,\ x_{n+1},\ y)\,dx_{n+1}.$$

On the other hand, suppose that

$$\phi(X(n+1) < x | \mathcal{F}_n,\ X(n+1) < y < X(n+1) + X(n+2))$$

$$= \int_0^x B(x_n,\ \xi,\ y)\,d\xi \Big/ \int_0^y B(x_n,\ \xi,\ y)\,d\xi.$$

Then

$$\frac{\int_0^x B(X(n),\ \xi,\ y)\,d\xi}{\int_0^y B(X(n),\ \xi,\ y)\,d\xi} = \frac{\phi(X(n+1) < x < y < X(n+1) + X(n+2) | \mathcal{F}_n)}{\phi(X(n+1) < y < X(n+1) + X(n+2) | \mathcal{F}_n)}$$

$$= \frac{E[\phi(X(n+1) < x < y < X(n+1) + X(n+2) | \mathcal{F}_{n+1}) | \mathcal{F}_n]}{E[\phi(X(n+1) < y < X(n+1) + X(n+2) | \mathcal{F}_{n+1}) | \mathcal{F}_n]}$$

$$= \frac{\int_0^x \phi(X(n+2) > y - \xi | \mathcal{F}_n,\ X(n) = \xi) \phi(X(n+1) = d\xi | \mathcal{F}_n)}{\int_0^y \phi(X(n+2) > y - \xi | \mathcal{F}_n,\ X(n+1) = \xi) \phi(X(n+1) = d\xi | \mathcal{F}_n)}$$

$$= \frac{\int_0^x \phi(X(n+2) > y - \xi | \mathcal{F}_n,\ X(n+1) = \xi) \phi'(X(n+1) = \xi | \mathcal{F}_n)\,d\xi}{\int_0^y \phi(X(n+2) > y - \xi | \mathcal{F}_n,\ X(n+1) = \xi) \phi'(X(n+1) = \xi | \mathcal{F}_n)\,d\xi}.$$

Differentiating both sides of this eqaution with respect to x and then letting $x=y$ yields

$$\int_0^y \phi(X(n+2) > y - \xi | \mathcal{F}_n, X(n+1) = \xi) \phi'(X(n+1) = \xi | \mathcal{F}_n) \, d\xi$$

$$= \phi'(X(n+1) = y | \mathcal{F}_n) \int_0^y B(X(n), \xi, y) \, d\xi.$$

Thus

$$\int_0^y B(X(n), \xi, y) \, d\xi = \frac{\int_0^x \phi(X(n+2) > y - \xi | \mathcal{F}_n, X(n+1) = \xi)\phi'(X(n+1) = \xi | \mathcal{F}_n) \, d\xi}{\phi'(X(n+1) = y | \mathcal{F}_n)}.$$

Differentiating with respect to x and then replacing $y - x$ by x_{n+2} and x by x_{n+1} yields the equation

$$\phi(X(n+2) > x_{n+2} | \mathcal{F}_n, X(n+1) = x_{n+1})$$

$$= B(X(n), x_{n+1}, x_{n+1} + x_{n+2}) \frac{\phi'(X(n+1) = x_{n+1} + x_{n+2} | \mathcal{F}_n)}{\phi'(X(n+1) = x_{n+1} | \mathcal{F}_n)}$$

or

$$\phi'(X(n+2) = x_{n+2} | X(0) = x_0, \cdots, X(n+1) = x_{n+1})$$

$$= -\frac{\partial}{\partial x_{n+2}} B(x_n, x_{n+1}, x_{n+1} + x_{n+2}) \frac{\phi'(X(n+1) = x_{n+1} + x_{n+2} | X(0) = x_0, \cdots, X(n) = x_n)}{\phi'(X(n+1) = x_{n+1} | X(0) = x_0, \cdots, X(n) = x_n)}.$$

It therefore follows by induction on n that

$$\phi'(X(0) = x_0, \cdots, X(n) = x_n)$$

$$= \phi'(X(0) = x_0)\phi'(X(1) = x_1 | X(0) = x_0) \cdots \phi'(X(n) = x_n | X(0) = x_0, \cdots, X(n-1) = x_{n-1})$$

$$= -\phi'(X(0) = x_0)\phi'(X(1) = x_1 | X(0) = x_0) \cdots \phi'(X(n-2) = x_{n-2} | X(0) = x_0, \cdots, X(n-3) = x_{n-3})$$

$$\cdot \frac{\partial}{\partial x_n} B(X(n-2), x_{n-1}, x_{n-1} + x_n)\phi'(X(n-1) = x_{n-1} + x_n | X(0) = x_0, \cdots, X(n-2) = x_{n-2})$$

$$= -\frac{\partial}{\partial x_n} B(x_{n-2}, x_{n-1}, x_{n-1} + x_n)\phi'(X(0) = x_0, \cdots, X(n-2) = x_{n-2}, X(n-1) = x_{n-1} + x_n)$$

$$= (-1)^n \frac{\partial}{\partial x_n} B(x_{n-2}, x_{n-1}, x_{n-1} + x_n) \frac{\partial}{\partial x_n} B(x_{n-3}, x_{n-2}, x_{n-2} + x_{n-1} + x_n)$$

$$\cdots \frac{\partial}{\partial x_n} B(x_0, x_1, x_1 + \cdots + x_n)\phi'(X(0) = x_0, X(1) = x_1 + \cdots + x_n)$$

and so, comparing with equation (1) we see that ϕ is a Ξ_0 process. Finally, we note that Φ is weak* closed since Φ_{++} is weak* closed. That is to say, if ϕ_{a_n, b_n} converges weak* to ϕ, then for any continuous function f mapping $R_+ \times R_+ \to R$ one has

$$\int_0^\infty \int_0^\infty f(u, v) p^* E(dv) T E(du) [\phi_{a_m, b_m} - \phi_{a_n, b_n}]$$

$$= \int_0^\infty \int_0^\infty f(u, v) \phi_{a_m, b_m}(du, dv) - \int_0^\infty \int_0^\infty f(u, v) \phi_{a_n, b_n}(du, dv)$$

$$= \int_0^\infty f(a_m, v) e^{-b_m v} dv - \int_0^\infty f(a_n, v) e^{-b_n v} dv \to 0.$$

Clearly this implies the existence of a and b such that $a_n \to a$ and $b_n \to b$ and $\phi_{a_n, b_n} \to \phi_{a, b} = \phi$.

7. **Example where $\bar{X}^* \neq X^*$.** Consider the discrete time temporally homogeneous Markov process Y on $\{1, 2, 3\}$ whose transition probability matrix is

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix}$$

and whose initial distribution is $(0, 0, 1)$. Let θ be the mapping of $\{1, 2, 3\}$ onto $\{1, 2\}$ defined by $f(1) = f(2) = 1$ and $f(3) = 2$, and let μ be the probability measure generated by θY on $(\Omega, \mathcal{F})$ where Ω is the set of all functions mapping the nonnegative integers into $\{1, 2\}$. The reader can easily check that $X(\mu) = (\Phi, T, E(U), p^*, \phi_0)$ where

(1) $$\Phi = R^3,$$

(2) $$T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix},$$

(3) $$E(\{1\}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E(\{2\}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(4) $$p^*(x_1, x_2, x_3) = x_1 + x_2 + x_3,$$

(5) $$\phi_0 = (0, 0, 1).$$

Furthermore,

$$\Phi_{++} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

and $X(\phi_0)^*$ is the temporally homogeneous Markov process on Φ_{++} with transition probability matrix

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix}.$$

Thus identifying 1 with $(1, 0, 0)$, 2 with $(0, 1, 0)$ and 3 with $(0, 0, 1)$ we see that $X(\phi)^* = Y$. On the other hand, $\overline{X}^*$ lives on $\overline{\Phi}_{++} = \{(1, 0, 0), (-1, 2, 0), (0, 0, 1)\}$ and has the transition probability matrix

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

REFERENCES

1. Robert R. Bush and Frederick Mosteller, *Stochastic models for learning*, Wiley, New York; Chapman & Hall, London, 1955.   MR 16, 1136.

2. N. Dunford and J. T. Schwartz, *Linear operators. I: General theory*, Pure and Appl. Math., vol. 7, Interscience, New York, 1958.   MR 22 #8302.

3. Dudley Paul Johnson, *On a class of stochastic processes and its relationship to infinite particle gases*, Trans. Amer. Math. Soc. 132 (1968), 275–295.   MR 41 #1108.

4. ———, *On a class of stochastic processes with two states and continuous time parameter*, Trans. Amer. Math. Soc. 142 (1969), 257–267.   MR 39 #7687.

5. Dudley Paul Johnson, *Markov process representations of general stochastic processes*, Proc. Amer. Math. Soc. 24 (1970), 735–738.   MR 41 #6303.

6. ———, *Hitting time distribution for general stochastic processes*, Ann. Math. Statist. 43 (1972), 1716–1718.

7. P. A. Meyer, *Probability and potentials*, Blaisdell, Waltham, Mass., 1966. MR 34 #5119.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALGARY, CALGARY, ALBERTA, CANADA