The Annals of Probability
1993, Vol. 21, No. 4, 2246–2268

INVARIANT MEASURES AND EVOLUTION EQUATIONS FOR MARKOV PROCESSES CHARACTERIZED VIA MARTINGALE PROBLEMS

BY ABHAY G. BHATT[1] AND RAJEEVA L. KARANDIKAR

Indian Statistical Institute

We extend Echeverria's criterion for invariant measures for a Markov process characterized via martingale problems to the case where the state space of the Markov process is a complete separable metric space. Essentially, the only additional conditions required are a separability condition on the operator occurring in the martingale problem and the well-posedness of the martingale problem in the class of progressively measurable solutions (as opposed to well-posedness in the class of r.c.l.l. solutions, i.e. solutions with paths that are right continuous and have left limits, in the locally compact case). Uniqueness of the solution to the (measure valued) evolution equation for the distribution of the Markov process (as well as a perturbed equation) is also proved when the test functions are taken from the domain of the operator of the martingale problem.

1. Introduction. Suppose A is the generator of a Markov process. Then a probability measure μ is an invariant measure for the Markov process if

$$\int (Af)\, d\mu = 0 \qquad \forall f \in \text{domain } A. \tag{1.1}$$

Also, for an initial distribution μ_0, if μ_t is the law of the process at time t, then $(\mu_t)_{t\geq 0}$ is the unique solution to the measure valued evolution equation

$$\int f d\nu_t = \int f d\mu_0 + \int_0^t \left(\int Af d\nu_s \right) ds \qquad \forall f \in \text{domain } A. \tag{1.2}$$

However, typically, domain A is difficult to describe even for, say, finite dimensional diffusions arising as solutions to Itô stochastic differential equations.

On the other hand, if one starts with an operator A and if the martingale problem for A is well-posed, then under suitable conditions, it gives rise to a Markov process. Echeverria (1982) showed that when the state space is a locally compact metric space, then under suitable conditions on A, (1.1) still implies that μ is invariant. From this, one can deduce that (1.2) admits a unique solution as well. The advantage here is that one can choose domain A suitably.

Received June 1991; revised June 1992.
[1]Supported by National Board for Higher Mathematics, Bombay, India.
AMS 1991 *subject classifications*. Primary 60J25; secondary 60J35, 60G05, 60G44.
Key words and phrases. Markov process, Martingale problem, invariant measure, evolution equation.

In this article, we look at these questions when A is an operator on $C_b(E)$, where E is a complete, separable metric space. We are able to prove that (1.1) implies that μ is an invariant measure, assuming that: (i) domain A is an algebra that separates points and vanishes nowhere; (ii) a separability condition holds (Hypothesis I), and (iii) every progressively measurable solution to the martingale problem for A admits an r.c.l.l. modification. It may be noted that in the compact case, (i) is assumed, and (ii) and (iii) are consequences of compactness of E and (i). It is proved that for Hilbert space valued diffusions as in Yor (1974), these conditions are satisfied. The presentation here closely follows that in Ethier and Kurtz (1986). We first prove existence of a stationary solution to the (A, μ) martingale problem if (1.1) holds and then deduce its consequences when the martingale problem is well-posed.

Under the assumptions mentioned above, we also prove that the evolution equation (1.2) admits a unique solution. Uniqueness of solutions to a perturbed form of the evolution equation is also proved.

In Section 5, we consider time dependent operators (A_t) which give rise to a time inhomogeneous Markov process. Uniqueness of solutions to the evolution equation for this process [see (5.5)] and its perturbed form is deduced from the time homogeneous case.

It may be noted that the uniqueness of the solution to (5.5) is used to prove uniqueness of the solution to the Boltzman equation in the spatially homogeneous noncutoff case [see Horowitz and Karandikar (1990)].

In the last section we apply the results on uniqueness of solutions to the perturbed evolution equation to nonlinear filtering theory.

2. Preliminaries. Throughout this article, (E, d) will denote a complete separable metric space, $B(E)$ will denote the space of real-valued bounded measurable functions on E, $C_b(E)$ will denote the space of bounded continuous real-valued functions on E, $\mathscr{B}(E)$ will denote the Borel σ-field on E, $\mathscr{P}(E)$ will denote the space of probability measures on $(E, \mathscr{B}(E))$, $\mathscr{M}(E)$ will denote the space of finite positive measures on $(E, \mathscr{B}(E))$, $D([0,\infty), E)$ will denote the set of all r.c.l.l. functions from $[0,\infty)$ into E and I_B will denote the indicator function of the set B. For $g \in B(E)$, $\|g\|$ will denote its sup norm.

For f_k, f in $B(E)$, we say that $f_k \to_{\text{bp}} f$ (where bp stands for bounded pointwise) if $\|f_k\| \le M$ and $f_k(x) \to f(x)$ for all $x \in E$. A class of functions $\mathscr{U} \subset B(E)$ is said to be bp-closed if $f_k \in \mathscr{U}$, $f_k \to_{\text{bp}} f$ implies $f \in \mathscr{U}$. For $\mathscr{V} \subseteq B(E)$, bp-closure $(\mathscr{V})$ is defined to be the smallest class of functions in $B(E)$ which contains $\mathscr{V}$ and is bp-closed. It may be noted that bp-closure is not closure in any topology.

For a measurable process $(X(t))$, defined on some probability space $(\Omega, \mathscr{F}, P)$,

$$\mathscr{F}_t^X := \sigma(X(s): s \le t),$$

$${}^*\mathscr{F}_t^X := \sigma\left(X(s), \int_0^s f(X(u))\,du: s \le t,\ f \in C_b(E)\right).$$

It can be shown that $\mathscr{F}_t^X$ and ${}^*\mathscr{F}_t^X$ differ only by P-null sets.

In this article, A will stand for an operator on $C_b(E)$ with domain $\mathscr{D}(A)$.

A measurable process $X(t)$ adapted to $(\mathscr{G}_t)$ is said to be a solution to the martingale problem for (A, μ) with respect to $(\mathscr{G}_t)$ if for all $f \in \mathscr{D}(A)$,

$$M^f(t) := f(X(t)) - \int_0^t Af(X(s))\,ds$$

is a $(\mathscr{G}_t)$-martingale and $P \circ X_0^{-1} = \mu$. When $\mathscr{G}_t =^* \mathscr{F}_t^X$, the σ-fields are dropped from the statement.

The martingale problem for (A, μ) is said to be *well-posed* if there exists a solution X and for any two solutions $X, \tilde{X}$, the finite dimensional distributions of $X, \tilde{X}$ are the same.

Let $(\theta(t))$ denote the coordinate mappings on $D([0,\infty), E)$. A measure P on $D([0,\infty), E)$ is said to be a solution to the martingale problem for (A, μ) if, under P, the coordinate process $(\theta(t))$ is a solution to the martingale problem for (A, μ).

The $D([0,\infty), E)$-martingale problem for (A, μ) is said to be well-posed if there is a unique measure P_μ on $D([0,\infty), E)$, which is a solution to the martingale problem for (A, μ).

When the martingale problem for (A, μ) is well-posed, the solution X is a Markov process, that is,

$$E\big[f(X(t))|\mathscr{F}_s^X\big] = E\big[f(X(t))|\sigma(X(s))\big]$$

for all $s < t$, $f \in C_b(E)$. See Ethier and Kurtz (1986), page 184. Under some conditions it can be shown that this Markov process admits a transition function $P(s, x, B)$. The next result gives one such set of conditions.

THEOREM 2.1. *Suppose the $D([0,\infty), E)$-martingale problem for (A, δ_x) is well-posed for each $x \in E$. Let P_x denote the solution. Suppose further that A satisfies the following separability condition:*

HYPOTHESIS I. *There exists a countable subset $\{g_k\} \subset \mathscr{D}(A)$ such that*

$$bp\text{-}closure\big(\{(g_k, Ag_k): k \geq 1\}\big) \supset \{(g, Ag): g \in \mathscr{D}(A)\}.$$

Then

(i) *$x \mapsto P_x(C)$ is measurable for all Borel sets C in $D([0,\infty), E)$.*

(ii) *For all $\mu \in \mathscr{P}(E)$, the $D([0,\infty), E)$-martingale problem for (A, μ) is well-posed, with the solution P_μ given by*

$$P_\mu(B) = \int_E P_x(B)\mu(dx). \tag{2.1}$$

(iii) *Under P_μ, $\theta(t)$ is a Markov process with transition probability function P given by*

$$P(s, x, B) = P_x(\theta_s \in B). \tag{2.2}$$

PROOF. (i) is essentially Theorem IV.4.6 in Ethier and Kurtz (1986). See Remark 2.1 in Horowitz and Karandikar (1990). For part (ii), it is easy to see that P_μ defined by (2.1) is a solution to the martingale problem for (A, μ). To see that the martingale problem is well-posed, note that in view of Hypothesis I, $(X(t))$ is a solution to the martingale problem for (A, μ) if and only if

$$g_k(X(t)) - \int_0^t Ag_k(X(s))\,ds$$

is a martingale for $k \geq 1$. If Q is a solution to the $D([0,\infty), E)$ martingale problem for (A, μ) and Q_ω is the regular conditional probability of Q given $\theta(0)$, it can be shown that Q_ω is a solution to the martingale problem for A for Q-a.e. ω [see Theorem 6.1.3 in Stroock and Varadhan (1978)]. Hence

$$Q_\omega = P_{\theta(0,\,\omega)}.$$

Thus it follows that $Q(C) = \int P_x(C)\,d\mu(x) = P_\mu(C)$. Part (iii) follows easily, adapting arguments in the result cited above. □

REMARK 2.1. When A satisfies the conditions of the preceding theorem, we associate the following Markov semigroup (T_t) with A:

$$T_t f(x) = \int f(y) P(t, x, dy) \tag{2.3}$$

for $f \in B(E)$, where $P(\cdot, \cdot, \cdot)$ is given by (2.2).

We will require the following lemma in the next section.

LEMMA 2.2. *Let $\{g_k\} \subset C_b(E)$ be a countable subset that separates points in E and vanishes nowhere. Let U_n, U be E-valued random variables defined on $(\Omega_0, \mathscr{F}_0, P_0)$ such that $P_0 \circ U_n^{-1} = P_0 \circ U^{-1}$ for all n. Suppose $g_k(U_n) \to g_k(U)$ in probability as $n \to \infty\ \forall\, k$. Then $U_n \to U$ in probability as $n \to \infty$.*

PROOF. Note that $g_j(U_n)g_k(U) \to g_j(U)g_k(U)$ in probability $\forall\, j, k$. Let

$$\begin{aligned}\mathscr{D}_0 = \big\{h\colon E \times E \to \mathbb{R};\ & h(x_1, x_2)\\ &= g_j(x_1)g_k(x_2)\ \forall\, x_1, x_2 \text{ for some } j \geq 1,\, k \geq 1\big\}\end{aligned}$$

and let $\mathscr{U}$ be the algebra generated by $\mathscr{D}_0$. Since $\{g_k\}$ separates points in E, $\mathscr{U}$ separates points in $E \times E$. Also for $h \in \mathscr{U}$,

$$h(U_n, U) \to h(U, U) \quad \text{in probability}.$$

Also (U_n, U) is relatively compact by hypothesis. Let $\varepsilon > 0$, $\delta > 0$ be arbitrary but fixed. Choose a compact subset K_δ of $E \times E$ with

$$P_0\{(U_n, U) \in K_\delta\} \geq 1 - \delta \qquad \forall\, n.$$

Now the metric d restricted to K_δ, which we continue to denote by d, is continuous and $\mathscr{U}' = \mathscr{U}|_{K_\delta}$ is an algebra that separates points and vanishes nowhere. Hence by the Stone–Weierstrass theorem $\mathscr{U}'$ is dense in $C(K_\delta)$ in the uniform topology. Choose $h_k \in \mathscr{U}'$ such that $\|h_k - d\| \to 0$ as $k \to \infty$. Then remembering that $d(U,U) = 0$, we have

$$\begin{aligned} P_0\{d(U_n,U) > \varepsilon\} &\le P_0\{d(U_n,U) > \varepsilon, (U_n,U) \in K_\delta\} + P_0\{(U_n,U) \notin K_\delta\} \\ &\le P_0\{|d(U_n,U) - h_k(U_n,U)| > \varepsilon/3, (U_n,U) \in K_\delta\} \\ &\quad + P_0\{|h_k(U_n,U) - h_k(U,U)| > \varepsilon/3, (U_n,U) \in K_\delta\} \\ &\quad + P_0\{|d(U,U) - h_k(U,U)| > \varepsilon/3, (U_n,U) \in K_\delta\} + \delta. \end{aligned}$$

Choosing k such that $\|h_k - d\| < \varepsilon/3$, we get

$$P_0\{d(U_n,U) > \varepsilon\} < P_0\{|h_k(U_n,U) - h_k(U,U)| > \varepsilon/3, (U_n,U) \in K_\delta\} + \delta.$$

Taking lim sup over n, we conclude $\limsup P_0(d(U_n,U) > \varepsilon) \le \delta$. Hence $U_n \to U$ in probability as $n \to \infty$. □

The next result is a key step in the proof of our main result. This is a generalization of the Riesz representation theorem.

THEOREM 2.3. *Let E be a complete separable metric space and let Λ be a positive linear functional on $C_b(E \times E)$ with $\Lambda 1 = 1$. Suppose that there exist (countably additive) probability measures μ_1, μ_2 on E such that*

$$\Lambda(F_f) = \int f(x)\, d\mu_1(x),$$

$$\Lambda(G^g) = \int g(y)\, d\mu_2(y)$$

for $f, g \in C_b(E)$, where $F_f(x,y) = f(x)$, $G^g(x,y) = g(y)$.

Then there exists a countably additive probability measure ν on $E \times E$ such that

$$\Lambda(F) = \int_{E\times E} F\, d\nu. \tag{2.4}$$

PROOF. First, note that there exists a unique finitely additive measure ν on the Borel field of $E \times E$ satisfying (2.4). [See Parthasarathy (1967), Theorem II.5.7.] Fix $\varepsilon > 0$. Since E is a complete separable metric space, we can choose K, a compact subset of E with $\mu_i(K) \ge 1 - \varepsilon$, $i = 1,2$. Then

$$\begin{aligned} \nu\big((K \times K)^c\big) &\le \nu(K^c \times E) + \nu(E \times K^c) \\ &= \mu_1(K^c) + \mu_2(K^c) \le 2\varepsilon. \end{aligned}$$

Let

$$\tilde{K} = K \times K \subset E \times E$$

and $\{F_m\} \subset C_b(E \times E)$ be decreasing to zero. Then $\|F_m\| \le \|F_1\|$. For $\delta > 0$, let B_m^δ be defined by $B_m^\delta = \{F_m \ge \delta\}$. Then B_m^δ is a decreasing sequence of closed sets and $B_m^\delta \cap \tilde{K}$ is compact in $E \times E$. Since $\bigcap_{m=1}^\infty \{B_m^\delta \cap \tilde{K}\} = \varnothing$, $\exists\ m_0$ such that $B_m^\delta \cap \tilde{K} = \varnothing$ whenever $m \geq m_0$. Hence for $m \geq m_0$, we have

$$\begin{aligned}\Lambda F_m &= \int_{E\times E} F_m \, d\nu \\ &\leq \int_{E\times E} I_{\{\|F_m\| < \delta\}} F_m \, d\nu + \int_{E\times E} I_{\{\tilde{K}^c\}} F_m \, d\nu \\ &\leq \delta + 2\|F_1\|\varepsilon.\end{aligned}$$

Since this holds for all δ and ε, we get,

$$\Lambda F_m \to 0 \quad \text{as } m \to \infty.$$

Hence by Daniell's theorem [see Neveu (1965), Proposition II.7.1] there exists a unique σ-additive probability measure, again denoted by ν, defined on the Borel σ-field of $E \times E$, satisfying (2.4). □

3. A criterion for invariant measures of a Markov process. The following is the main result of this article. This is an extension of Theorem IV.9.17 in Ethier and Kurtz (1986).

THEOREM 3.1. *Let $\mathscr{D}(A)$ be an algebra that separates points and vanishes nowhere. Suppose A satisfies Hypothesis* I *and that for all $\nu \in \mathscr{P}(E)$, there exists a solution to the $D([0,\infty))$-martingale problem for (A, μ). Suppose that $\mu \in \mathscr{P}(E)$ satisfies*

$$(3.1) \qquad \int_E Af \, d\mu = 0 \qquad \forall\, f \in \mathscr{D}(A).$$

Then on some probability space, there exists a $(\mathscr{G}_t)$-progressively measurable process X such that X is a stationary process and that X is a solution to the (A, μ)-martingale problem w.r.t. $(\mathscr{G}_t)$.

PROOF. For $n \geq 1$ define A_n on $\mathscr{R}(I - n^{-1}A)$ by

$$A_n f = n\left[(I - n^{-1}A)^{-1} - I\right] f.$$

Note that since the martingale problem for (A, δ_x) admits a solution for all $x \in E$, A is dissipative [see Ethier and Kurtz (1986), page 178] and hence A_n is well defined. Also for $f \in \mathscr{D}(A)$, $f_n := (I - n^{-1}A)f$ satisfy

$$(3.2) \qquad A_n f_n = Af \quad \forall\, n; \qquad \|f_n - f\| \to 0 \quad \text{as } n \to \infty.$$

Note that for all $g \in D(A_n)$, we have

$$(3.3) \qquad \int_E A_n g \, d\mu = 0.$$

We divide the rest of the proof into steps.

STEP 1. Construction of stationary solution to the martingale problem for A_n:

Fix n. Let $M \subset C_b(E \times E)$ be the linear space of functions of the form

$$F(x,y) = \sum_{i=1}^{m} f_i(x) g_i(y) + g(y), \tag{3.4}$$

$f_1, \dots, f_m, g \in C_b(E)$; $g_1, \dots g_m \in \mathscr{R}(I - n^{-1}A)$; $m \geq 1$. Define Λ on M by

$$\Lambda F = \int_E \sum_{i=1}^{m} \Big[f_i(x)(I - n^{-1}A)^{-1} g_i(x) + g(x) \Big] \mu(dx) \tag{3.5}$$

for F as in (3.4). Then $\Lambda 1 = 1$. Using (3.1) and proceeding exactly as in Theorem 9.17 of Ethier and Kurtz (1986), we get that $|\Lambda F| \leq \|F\|\ \forall\ F \in M$ and $\Lambda F \geq 0$ whenever $F \geq 0$. Using the Hahn–Banach theorem we extend Λ to a bounded, positive linear functional on $C_b(E \times E)$. From the construction it is clear that

$$\Lambda(F_f) = \int_E f(x)\, d\mu(x), \tag{3.6}$$

$$\Lambda(G^g) = \int_E g(y)\, d\mu(y), \tag{3.7}$$

where $F_f(x,y) = f(x)$ and $G^g(x,y) = g(y)$. Using Theorem 2.3, we get a probability measure ν on $E \times E$ satisfying

$$\Lambda F = \int_{E \times E} F\, d\nu \qquad \forall\ F \in M. \tag{3.8}$$

Since E is a complete separable metric space, there exists a transition probability function [see e.g., Ethier and Kurtz (1986), Appendix] $\eta\colon E \times \mathscr{B}(E) \to [0,1]$ satisfying

$$\nu(B_1 \times B_2) = \int_{B_1} \eta(x, B_2)\mu(dx) \qquad \forall\ B_1, B_2 \in \mathscr{B}(E).$$

Note that (3.6) and (3.7) give

$$\int_E \eta(x, B)\mu(dx) = \nu(E \times B) = \mu(B) \qquad \forall\ B \in \mathscr{B}(E). \tag{3.9}$$

Now, (3.5) and (3.8) imply that for all $f \in C_b(E)$, $g \in \mathscr{R}(I - n^{-1}A)$

$$\int_{E \times E} f(x) g(y) \nu(dx, dy) = \int_E f(x)(I - n^{-1}A)^{-1} g(x)\mu(dx).$$

Hence we have

$$\int_E g(y)\eta(x, dy) = (I - n^{-1}A)^{-1} g(x) \qquad \mu\text{-a.s.} \tag{3.10}$$

for all $g \in \mathscr{R}(I - n^{-1}A)$.

Let $Y(0), Y(1), \dots, Y(k), \dots$ be an E-valued Markov chain with initial distribution μ and transition function η. Equation (3.9) implies that $(Y(k))$ is stationary and (3.10) gives that

$$g(Y(k)) - \sum_{i=0}^{k-1} n^{-1}A_n g(Y(i)) \tag{3.11}$$

is a $\sigma(Y(1),\dots,Y(k))$-martingale for all $g \in \mathscr{R}(I - n^{-1}A)$.

Let V be a Poisson process with parameter n, which is independent of Y. Define

$$X_n(t) := Y(V(t)).$$

Then X_n is a stationary Markov process with initial distribution μ. Equation (3.11) implies that

$$g(X_n(t)) - \int_0^t A_n g(X_n(s))\,ds$$

is a martingale for all $g \in \mathscr{D}(A_n)$. Thus X_n is a stationary solution to the martingale problem for A_n. This completes Step 1.

STEP 2. Convergence of finite dimensional distributions of (a subsequence of) X_n:

For $f \in \mathscr{D}(A)$, let f_n be as in (3.2). Define

$$\xi_n(t) := f_n(X_n(t)),$$

$$\phi_n(t) := A_n f_n(X_n(t)) = Af(X_n(t)).$$

Then ξ_n and ϕ_n satisfy the conditions of Theorem III.9.4 of Ethier and Kurtz (1986), and since $\mathscr{D}(A)$ is an algebra, the same theorem applies and we get relative compactness of $(f_1 \circ X_n, f_2 \circ X_n, \dots, f_i \circ X_n)$ in $D([0,\infty), \mathbb{R}^i)$, for $f_1, f_2, \dots, f_i \in \mathscr{D}(A)$, $i \ge 1$.

Let $\mathscr{D}_0 = \{g_k\}_{k=1}^{\infty}$ be the countable subset of Hypothesis I. Let $\|g_k\| = a_k$ and $\hat{E} = \prod_{k=1}^{\infty}[-a_k, a_k]$. Since $\mathscr{D}(A)$ separates points and vanishes nowhere, so does $\mathscr{D}_0$. It now follows that $(g_1(X_n(\cdot)), g_2(X_n(\cdot)), \dots, g_k(X_n(\cdot)), \dots)$ is relatively compact in $D([0,\infty), \hat{E})$. Thus we get a subsequence, which we relabel as X_n, such that $(g_1(X_n(\cdot)), g_2(X_n(\cdot)), \dots, g_k(X_n(\cdot)), \dots)$ converges weakly to a $D([0,\infty), \hat{E})$ valued random variable, say, $Z(\cdot) = (Z_1(\cdot), \dots, Z_k(\cdot), \dots)$, that is,

$$(g_1(X_n(\cdot)), \dots g_k(X_n(\cdot)), \dots) \to_{\mathscr{L}} Z(\cdot) \quad \text{as } n \to \infty. \tag{3.12}$$

Define $\underline{g}\colon E \to \hat{E}$ by

$$\underline{g}(x) = (g_1(x), \dots, g_k(x), \dots). \tag{3.13}$$

Then $\underline{g}$ is a one to one, continuous function. This implies that $\underline{g}(E)$ is a Borel subset of $\hat{E}$. Also $\underline{g}^{-1}$ defined on $\underline{g}(E)$ is measurable. [See Parthasarathy (1967), Corollary I.3.3]. We extend the definition of $\underline{g}^{-1}$ to all of $\hat{E}$ by setting $\underline{g}^{-1}(z) = e$ for $z \notin \underline{g}(E)$, where e is a fixed point in E. We now use a Skorohod representation to get a probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ and $D([0,\infty), \hat{E})$ valued random variables $\tilde{\xi}_n$ and $\tilde{Z}$ defined on it satisfying

$$\mathscr{L}(\tilde{\xi}_n) = \mathscr{L}(\underline{g}(X_n)) \qquad \forall\, n, \tag{3.14}$$

$$\mathscr{L}(\tilde{Z}) = \mathscr{L}(Z), \tag{3.15}$$

$$\tilde{\xi}_n \to \tilde{Z} \quad \text{a.s. as } n \to \infty. \tag{3.16}$$

Now,

$$\mathscr{L}(\underline{g}(X_n(t))) = \mu \circ \underline{g}^{-1} := \tilde{\mu} \qquad \forall\, n, t.$$

Hence

$$\mathscr{L}(\tilde{\xi}_n(t)) = \tilde{\mu} \qquad \forall\, n, t, \tag{3.17}$$

which implies $\tilde{\xi}_n(t) \in \underline{g}(E)$ a.s. Then defining

$$\tilde{X}_n(t) = \underline{g}^{-1}(\tilde{\xi}_n(t)), \tag{3.18}$$

it follows that $\tilde{X}_n$ is a measurable process. Since $(X_n(t))$ is a stationary process, it follows that $(\tilde{\xi}_n(t))$ is a stationary process and hence $(\tilde{Z}(t))$ is a $\hat{E}$ valued r.c.l.l. stationary process. Hence $(\tilde{Z}(t))$ does not have any fixed points of discontinuity, that is, $P(\tilde{Z}(t) = \tilde{Z}(t-)) = 1\ \forall\, t$. Thus $\tilde{\xi}_n(t) \to \tilde{Z}(t)$ a.s. for all t. Since $\mathscr{L}(\tilde{\xi}_n(t)) = \tilde{\mu}$ for all n, it follows that $\mathscr{L}(\tilde{Z}(t)) = \tilde{\mu}$. Hence

$$P(\tilde{Z}(t) \in \underline{g}(E)) = 1 \qquad \forall\, t$$

and defining

$$\tilde{X}(t) = \underline{g}^{-1}(\tilde{Z}(t)), \tag{3.19}$$

we get a stationary $(\mathscr{G}_t)$-progressively measurable process $\tilde{X}$, where $\mathscr{G}_t = \mathscr{F}_t^{\tilde{Z}}$. Further,

$$\underline{g}(\tilde{X}_n(t)) \to \underline{g}(\tilde{X}(t)) \quad \text{a.s.}\ \forall\, t. \tag{3.20}$$

This and Lemma 2.2 imply that $\tilde{X}_n(t)$ converges to $\tilde{X}(t)$ as $n \to \infty$ in E in $\tilde{P}$ probability for each t. This completes Step 2.

To complete the proof, we will show that $\tilde{X}$ is a solution to the martingale problem for (A, μ) w.r.t. $(\mathscr{G}_t)$. Recall that we have already proved that $\tilde{X}$ is a stationary process and is $(\mathscr{G}_t)$ progressively measurable. Note that $(\tilde{X}_n(t))$ has the same finite dimensional distributions as $(X_n(t))$, and hence $\tilde{X}_n$ is a solution to the martingale problem for A_n, that is, for all $g \in \mathscr{D}(A_n)$,

$$g(\tilde{X}_n(t)) - \int_0^t A_n g(\tilde{X}_n(s))\, ds \tag{3.21}$$

is a $\tilde{P}$ martingale. Now for $g \in C_b(E)$,

$$g(\tilde{X}_n(t)) \to g(\tilde{X}(t)) \quad \text{in } \tilde{P} \text{ probability} \tag{3.22}$$

as $n \to \infty$. This holds for all t. An application of the dominated convergence theorem gives for $g \in C_b(E)$,

$$\mathbb{E}|g(\tilde{X}_n(s)) - g(\tilde{X}(s))| \to 0 \quad \text{as } n \to \infty,$$

and hence using Fubini's theorem we get

$$\mathbb{E}\int_0^t |g(\tilde{X}_n(s)) - g(\tilde{X}(s))|\, ds \to 0 \quad \text{as } n \to \infty.$$

As a consequence, we have

$$\int_0^t g(\tilde{X}_n(s))\, ds \to \int_0^t g(\tilde{X}(s))\, ds \tag{3.23}$$

in $\tilde{P}$ probability for all t.

Fix $f \in \mathscr{D}_0$ and let f_n be given by (3.2). From (3.21) it follows that for $0 \le t_1 < t_2 < \cdots < t_{m+1}$, $h_k \in C_b(E)$,

$$\mathbb{E}\left[\left(f_n(\tilde{X}_n(t_{m+1})) - f_n(\tilde{X}_n(t_m)) - \int_{t_m}^{t_{m+1}} Af(\tilde{X}_n(s))\, ds\right) \prod_{k=1}^m h_k(\tilde{X}_n(t_k))\right] = 0,$$

and recalling that $\|f_n - f\| \to 0$ as $n \to \infty$, we get

$$\mathbb{E}\left[\left(f(\tilde{X}_n(t_{m+1})) - f(\tilde{X}_n(t_m)) - \int_{t_m}^{t_{m+1}} Af(\tilde{X}_n(s))\, ds\right) \prod_{k=1}^m h_k(\tilde{X}_n(t_k))\right] \to 0$$

as $n \to \infty$. Now (3.22), (3.23) and the dominated convergence theorem give

$$\mathbb{E}\left[\left(f(\tilde{X}(t_{m+1})) - f(\tilde{X}(t_m)) - \int_{t_m}^{t_{m+1}} Af(\tilde{X}(s))\, ds\right) \prod_{k=1}^m h_k(\tilde{X}(t_k))\right] = 0.$$

Since $\mathscr{F}_t^{\tilde{X}}$ and $\mathscr{F}_t^{\tilde{Z}} = \mathscr{G}_t$ differ only by null sets, it now follows that $(\tilde{X}(t))$ is a solution to the martingale problem for (A, μ) with respect to $(\mathscr{G}_t)$, completing the proof. □

It should be noted that the stationary solution constructed above may not have r.c.l.l. paths. Thus even when the $D([0,\infty))$ martingale problem for (A, ν) is well-posed for all $\nu \in \mathscr{P}(E)$, in addition to the conditions in Theorem 3.1, it does not follow that μ is a stationary initial distribution for the Markov process associated with A. This can be deduced if we assume that every solution to the (A, μ)-martingale problem admits an r.c.l.l. modification. This is our next result.

THEOREM 3.2. *Let $\mathscr{D}(A)$ be an algebra that separates points and vanishes nowhere. Suppose A satisfies Hypothesis* I. *Suppose that the $D([0,\infty))$-martingale problem for (A, δ_x) is well-posed for all $x \in E$. Let (T_t) be the semigroup associated with the A-martingale problem in Theorem* 2.1.

Further suppose that:

HYPOTHESIS II. *Every progressively measurable solution to the martingale problem for (A, μ) admits an r.c.l.l. modification.*

If $\mu \in \mathscr{P}(E)$ satisfies

$$\int_E Af\,d\mu = 0 \qquad \forall f \in \mathscr{D}(A),$$

then μ is an invariant measure for the semigroup (T_t).

PROOF. Let X be the stationary solution constructed in Theorem 3.1. In view of Hypothesis II, we can assume that X is r.c.l.l. If Q is the law of X, then it follows that Q is a solution to the $D([0,\infty), E)$-martingale problem for (A, μ) and that $Q \circ (\theta(t))^{-1} = \mu$ for all t. Using (2.2) and (2.3) it now follows that

$$\int_E (T(t)f)\,d\mu = \int_E f\,d\mu$$

for all t and hence that μ is an invariant measure for $T(t)$. □

REMARK 3.1. When E is a compact metric space, Hypothesis II always holds and when E is a locally compact metric space and A is an operator on $\hat{C}(E)$ (continuous functions vanishing at infinity), then also Hypothesis II holds if A is conservative, that is, $(1,0)$ is in the bp-closure of $\{(f, Af): f \in \mathscr{D}(A)\}$ [see, e.g., Ethier and Kurtz (1986), page 179). Note that when A is a second order differential operator on $\mathbb{R}^m$ with bounded coefficients, it is conservative and hence Hypothesis II is satisfied. It may be noted that Echeverria proved this result without assuming that A is conservative in the locally compact case.

REMARK 3.2. Let $E = H$, a real, separable Hilbert space. Let A be the operator corresponding to a diffusion as in Yor (1974). More precisely, fix a complete orthonormal system $\{\phi_i\}$ in H and let P_n be the orthogonal projection onto the linear span of $\{\phi_1, \dots \phi_n\}$. Let $\mathscr{D}(A) = \{f \circ P_n: f \in C_0^2(\mathbb{R}^n)\}$ and

$$[A(f \circ P_n)](h) = \tfrac{1}{2} \sum_{i,j=1}^{n} \big(\sigma^*(h)\phi_i, \sigma^*(h)\phi_j\big) f_{ij} \circ P_n(h) + \sum_{i=1}^{n} (b(h), \phi_i) f_i \circ P_n(h),$$

where $f_i = (\partial/\partial x_i) f$ and $f_{ij} = (\partial/\partial x_j) f_i$. Here $\sigma: H \to \mathscr{L}_2(H, H)$ and $b: H \to H$ are measurable functions, $\mathscr{L}_2(H, H)$ being the space of Hilbert–Schmidt operators with norm $\|\cdot\|_2$. It is assumed that $\|b(h)\| \le C_1$ and $\|\sigma(h)\|_2 \le C_2$. If $(X(t))$ is a progressively measurable process that is a solution to the martingale problem for (A, μ), then it can be shown that $(X^i(t)) := (X(t), \phi_i)$ admits an r.c.l.l. modification, say $(\tilde{X}^i(t))$. This step is similar to the locally compact case referred to in Remark 3.1. X^i is then an Itô process in the sense of Stroock–Varadhan and the r.c.l.l. version $\tilde{X}^i$ is indeed continuous a.s. [see Stroock and Varadhan (1979), page 111]. Let

$$M^i(t) = \tilde{X}^i(t) - \tilde{X}^i(0) - \int_0^t (b(X(s)), \phi_i)\, ds.$$

Then it follows that M^i is a continuous martingale and

$$\langle M^i, M^j \rangle_t = \int_0^t (\sigma^*(X(s))\phi_i, \sigma^*(X(s))\phi_j)\, ds.$$

Now

$$\begin{aligned} \mathbb{E} \sup_{t \le T} \left\| \sum_{j=m}^n M^j(t)\phi_j \right\|^2 &= \mathbb{E} \sup_{t \le T} \sum_{j=m}^n (M^j(t))^2 \\ &\le 4\mathbb{E} \sum_{j=m}^n (M^j(T))^2 \qquad (3.24) \\ &= 4\mathbb{E} \sum_{j=m}^n \int_0^T (\sigma^*(X(s))\phi_i, \sigma^*(X(s))\phi_j)\, ds. \end{aligned}$$

Since

$$\begin{aligned} \mathbb{E} \sum_{j=1}^\infty \int_0^T (\sigma^*(X(s))\phi_i, \sigma^*(X(s))\phi_j)\, ds &= \mathbb{E} \sum_{j=1}^\infty \int_0^T \|\sigma^*(X(s))\|_2^2\, ds \\ &\le TC_1, \end{aligned}$$

it follows that the R.H.S. in (3.24) goes to zero. Thus $\sum_{j=1}^n M^j(t)\phi_j$ converges uniformly in $[0, T]$ a.s. to, say, $(M(t))$, and hence $(M(t))$ is continuous. It is easy to see that

$$M(t) = X(t) - X(0) - \int_0^t b(X(s))\, ds \quad \text{a.s.}$$

Hence, defining $\tilde{X}(t) := M(t) + X(0) + \int_0^t b(X(s))\, ds$, one gets that $\tilde{X}$ is a continuous modification of X. Thus Hypothesis II is satisfied.

4. Evolution equation for Markov processes: uniqueness. We consider the measure valued equation

$$\int_E f\, d\nu_t = \int_E f\, d\nu + \int_0^t \left(\int_E Af\, d\nu_s \right) ds, \qquad f \in \mathscr{D}(A), \tag{4.1}$$

where $\{\nu_t\}_{t \ge 0} \subset \mathscr{P}(E)$ satisfy for every Borel set U in E:

$$t \mapsto \nu_t(U) \text{ is measurable.} \tag{4.2}$$

Note that if X is a solution to the martingale problem for (A, ν), then $\mu_t = \mathscr{L}(X(t))$ is a solution to (4.1). We will now show uniqueness under the conditions of Theorem 3.2. We require the following *imbedding* of the martingale problem for A on a compact space $\hat{E}$. This imbedding was implicit in the proof of the main theorem.

As in Section 3, let $\{g_k\} \subset \mathscr{D}(A)$ be the countable set such that the bp-closure of $\{(g_k, Ag_k)\}$ contains $\{(f, Af)\colon f \in \mathscr{D}(A)\}$ and let $\|g_k\| = a_k$. Let $\hat{E} = \prod_{k=1}^{\infty}[-a_k, a_k]$ and $\underline{g}\colon E \to \hat{E}$ be defined by

$$\underline{g}(x) = (g_1(x), \dots, g_k(x), \dots). \tag{4.3}$$

Define an operator $\mathscr{A}$ as follows. Let $\mathscr{U}$ be the algebra generated by

$$\{u_k \in C(\hat{E})\colon u_k((z_1, \dots, z_k, \dots)) = z_k\}$$

and

$$\mathscr{A}(cu_{i_1}u_{i_2} \cdots u_{i_k})(z) = \begin{cases} c[A(g_{i_1}g_{i_2} \cdots g_{i_k})](x), & \text{if } z = \underline{g}(x), \\ 0, & \text{otherwise}. \end{cases} \tag{4.4}$$

Note that

$$u_k(\underline{g}(x)) = g_k(x) \quad \text{and} \quad \mathscr{A}u_k(\underline{g}(x)) = Ag_k(x).$$

Let $(X(t))$ be a progressively measurable solution to the martingale problem for A and let

$$Z(t) = \underline{g}(X(t)) \qquad \forall\, t \geq 0.$$

$(Z(t))$ is a $\underline{g}(E)$ valued progressively measurable process. Also for $u \in \mathscr{U}$ with $u(\underline{g}(x)) = \bar{g}(x)$, we have that

$$\begin{aligned} &\mathbb{E}\left[\left(u(Z(t_m)) - u(Z(t_{m-1})) - \int_{t_{m-1}}^{t_m} \mathscr{A}u(Z(s))\, ds\right) \prod_{l=1}^{m-1} \underline{h}_l(Z(t_l))\right] \\ &\quad = \mathbb{E}\left[\left(g(X(t_m)) - g(X(t_{m-1})) - \int_{t_{m-1}}^{t_m} Ag(X(s))\, ds\right) \prod_{l=1}^{m-1} h_l(X(t_l))\right] \\ &\quad = 0, \end{aligned} \tag{4.5}$$

$\forall\, 0 \leq t_1 < t_2 < \cdots < t_m$, $\underline{h}_l \in B(\hat{E})$, $h_l = \underline{h}_l \circ \underline{g}$. Hence Z is a solution to the martingale problem for $\mathscr{A}$.

Similarly it can be seen that if Z is a progressively measurable solution to the martingale problem for $\mathscr{A}$ with

$$P\big(Z(t) \in \underline{g}(E)\big) = 1 \qquad \forall\, t \geq 0, \tag{4.6}$$

then

$$X(t) = \underline{g}^{-1}(Z(t))$$

defines a progressively measurable solution to the martingale problem for A.

The advantage of going to the martingale problem for $\mathscr{A}$ is that compactness of $\hat{E}$ and denseness of $\mathscr{U}$ implies that any measurable solution to the martingale problem for $\mathscr{A}$ admits an r.c.l.l. version. The preceding discussion implies that if the $D([0,\infty))$-martingale problem for (A, δ_x) is well-posed for all $x \in E$ and if every progressively measurable solution to the martingale problem for (A, μ) admits an r.c.l.l. modification, then the law of any solution Z to the martingale problem for $(\mathscr{A}, \mu)$ satisfying (4.6) is uniquely determined.

Let us note here that g is not a homeomorphism in general. However, if $\{g_k\}$ separates points and closed sets, then $\underline{g}$ is a homeomorphism and then $\hat{E}$ is a compactification of E, similar to the one point compactification of a locally compact space.

Let $E_0 = E \times \{-1, 1\}$, $\lambda > 0$, $\nu \in \mathscr{P}(E)$ and B be the operator on $C_b(E_0)$ with domain $\mathscr{D}(B)$ which is the linear span of

$$\{f_1 f_2 \colon f_1 \in \mathscr{D}(A),\ f_2 \in C(\{-1, 1\})\}$$

and

$$(4.7)\quad Bf_1 f_2(x, v) = f_2(v) Af_1(x) + \lambda\left(f_2(-v) \int_E f_1 \, d\nu - f_1(x) f_2(v)\right).$$

THEOREM 4.1. *Suppose that the operator A on $C_b(E)$ satisfies the conditions of Theorem* 3.2. *If $\{\nu_t\}_{t \geq 0} \subset \mathscr{P}(E)$ and $\{\mu_t\}_{t \geq 0} \subset \mathscr{P}(E)$ satisfy* (4.1) *and* (4.2), *then $\nu_t = \mu_t$ for all $t \geq 0$.*

PROOF. By definition of B, it is clear that $\mathscr{D}(B)$ is an algebra that separates points in E_0 and that B satisfies Hypothesis I.

Existence of $D([0, \infty), E_0)$ valued solutions to the martingale problem for (B, μ), for every $\mu \in \mathscr{P}(E_0)$, follows from Proposition IV.10.2 of Ethier and Kurtz (1986). We will prove that the martingale problem for (B, μ) is well-posed (in the class of progressively measurable solutions). Once this is proved, it would follow that the $D([0, \infty))$-martingale problem for (B, μ) is well-posed and that every measurable solution to the martingale problem admits a r.c.l.l. modification.

Let (Y, V) be a measurable solution to the martingale problem for (B, μ). Let

$$(4.8)\qquad Z(t) = \underline{g}(Y(t)) \qquad \forall\, t \geq 0$$

and $\mathscr{B}$ be an operator defined on

$$(4.9)\qquad \mathscr{D}(\mathscr{B}) = \{u f_2 \colon u \in \mathscr{U},\ f_2 \in C(\{-1, 1\})\}$$

by

$$(4.10)\quad \mathscr{B} u f_2(z, v) = f_2(v) \mathscr{A} u(z) + \lambda\left(f_2(-v) \int_{\hat{E}} u \, d\tilde{\nu} - f_2(v) u(z)\right),$$

where

$$(4.11)\qquad \tilde{\nu}(\Gamma) = \nu\left(\underline{g}^{-1}\left(\Gamma \cap \underline{g}(E)\right)\right).$$

Then arguing as in (4.5), it follows that (Z, V) is a solution to the martingale problem for $\mathscr{B}$. Since $\mathscr{D}(\mathscr{B})$ is an algebra that separates points in $\overline{\underline{g}(E)} \times \{-1, 1\}$, it is a measure determining set. Further, since $\overline{\underline{g}(E)} \times \{-1, 1\}$ is a compact, separable space, (Z, V) has a r.c.l.l. modification, say $(\hat{Z}, \hat{V})$, in $\overline{\underline{g}(E)} \times \{-1, 1\}$. [See, e.g., Theorem IV.3.6 of Ethier and Kurtz (1986).]

Now arguments as in Theorem IV.10.3 in Ethier and Kurtz (1986) imply that the law of $(\hat{Z}, \hat{V})$ is determined by the law of the r.c.l.l. solution to the martingale problem for $\mathscr{A}$ satisfying (4.6), which in turn is determined by $\{P_x\}$, the solution to the martingale problem for A and ν. Thus, the law of Z and hence that of Y is determined by $\{P_x\}$ and ν. It is easy to see that the law of V depends only on λ. This shows that the martingale problem for B is well-posed. Thus B satisfies all the conditions of Theorem 3.2.

The rest of the argument is the same as that in the locally compact case, given in the proof of Proposition IV.9.19 in Ethier and Kurtz (1986). □

We now consider the question of uniqueness for solutions to an evolution equation for perturbations of the operator A.

Let $\lambda \in C_b(E)$. If $(X(t))$ is a solution to the $D([0,\infty), E)$-martingale problem for (A, ν), then $\nu_t(B) := \mathbb{E}(I_B(X(t))\exp\{-\int_0^t \lambda(X(s))\,ds\})$ is a solution to the equation

$$\int_E f\,d\nu_t = \int_E f\,d\nu + \int_0^t \left(\int_E (Af - \lambda(\cdot)f)\,d\nu_s \right) ds, \qquad f \in \mathscr{D}(A). \tag{4.12}$$

The next result gives conditions under which $\{\nu_t\}_{t \geq 0}$ is the only solution.

THEOREM 4.2. *Suppose A satisfies the conditions of Theorem 3.2 and that the constant function 1 belongs to $\mathscr{D}(A)$ with $A1 = 0$. If $\{\mu_t\}_{t \geq 0} \subset \mathscr{M}(E)$ and $\{\nu_t\}_{t \geq 0} \subset \mathscr{M}(E)$ satisfy (4.2) and (4.12), then $\mu_t = \nu_t$ for all $t \geq 0$.*

PROOF. Let $\alpha = \inf_{x \in E} \lambda(x)$. Then if $\{\nu_t\}_{t \geq 0}$ satisfy (4.12), then $\nu'_t = \nu_t e^{\alpha t}$ satisfy

$$\int_E f\,d\nu'_t = \int_E f\,d\nu + \int_0^t \left(\int_E (Af - \lambda(\cdot)f + \alpha f)\,d\nu'_s \right) ds, \qquad f \in \mathscr{D}(A) \tag{4.13}$$

and conversely if $\{\nu'_t\}_{t \geq 0} \subset \mathscr{M}(E)$ satisfy (4.13), then $\nu_t = \nu'_t e^{-\alpha t}$ satisfy (4.12). Hence it suffices to prove uniqueness when $\lambda \geq 0$.

Without loss of generality we can assume that $1 \in \mathscr{D}(A)$ with $A1 = 0$. Let Δ be a discrete point outside E and

$$E^\Delta = E \cup \{\Delta\}. \tag{4.14}$$

Extend λ to E^Δ by defining

$$\tilde{\lambda}(\Delta) = 0. \tag{4.15}$$

Define operators A^Δ and C on $C_b(E^\Delta)$ by

$$\mathscr{D}(A^\Delta) = \{ f \in C_b(E^\Delta) \colon f|_E \in \mathscr{D}(A) \} \tag{4.16}$$

and for $f \in \mathscr{D}(A^\Delta)$,

$$\begin{aligned} A^\Delta f(x) &= A((f(x) - f(\Delta))) \qquad \forall\, x \in E, \\ A^\Delta f(\Delta) &= 0. \end{aligned} \tag{4.17}$$

For $f \in C_b(E^\Delta)$, $x \in E^\Delta$,

$$Cf(x) = \tilde{\lambda}(x)(f(\Delta) - f(x)). \tag{4.18}$$

We will first show that the martingale problem for $A^\Delta + C$ is well-posed. Let $\tilde{\mu} \in \mathscr{P}(E^\Delta)$ be defined by $\tilde{\mu}(U) = \mu(E \cap U)$ for U Borel in E^Δ. Existence of a solution to the $D([0,\infty), \hat{E})$-martingale problem for $(A^\Delta + C, \tilde{\mu})$ follows from Theorem IV.10.2 in Ethier and Kurtz (1986).

Note that since the martingale problem for A is well-posed, so is the martingale problem for A^Δ. Existence of a $D([0,\infty), E^\Delta)$ solution for the martingale problem for (A^Δ, π) for any $\pi \in \mathscr{P}(E^\delta)$ follows easily. Let $\{T^\Delta(t)\}$ be the associated semigroup.

Let X be a measurable solution to the martingale problem for $A^\Delta + C$. Since $I_E \in \mathscr{D}(A^\Delta + C)$ and $(A^\Delta + C)I_E = -\lambda I_E = -\tilde{\lambda}$, we get that

$$M(t) := I_E(X(t)) + \int_0^t \tilde{\lambda}(X(s))\, ds \tag{4.19}$$

is a martingale. Nonnegativity of $\tilde{\lambda}$ implies that $I_E(X(t))$ is a supermartingale. The filtration may not be right continuous. Hence to get an r.c.l.l. modification of $I_E(X(t))$ we proceed as follows.

Using $(I_E(X(t)))^2 = I_E(X(t))$, a simple calculation gives that

$$(M(t))^2 - \int_0^t \tilde{\lambda}(X(s))\, ds$$

is a martingale. [See, e.g., Ethier and Kurtz (1986), Exercise II.29.] Similarly

$$(M(t) - M(s))^2 - \int_s^t \tilde{\lambda}(X(u))\, du, \qquad t \geq s$$

is a martingale. This implies that the map $t \mapsto M(t)$ is continuous in probability. Hence $t \mapsto I_E(X(t))$ is continuous in probability and thus has an r.c.l.l. modification, say $(N(t))$. N can be taken to be $\{0,1\}$-valued. Further N is a positive supermartingale. Let

$$\tau = \inf\{t > 0: N(t) = 0\}. \tag{4.20}$$

Then $N(u) = 0$ for $u \geq \tau$ a.s. [see, e.g., Ethier and Kurtz (1986), page 62). Thus $N(t) \equiv I_{\{\tau > t\}}$ and

$$I_E(X(t)) = I_{\{\tau > t\}} \quad \text{a.s.} \tag{4.21}$$

Hence using (4.19) and integration by parts we get that

$$I_{\{\tau > t\}} \exp\left\{ \int_0^t \tilde{\lambda}(X(s))\, ds \right\} \tag{4.22}$$

is a martingale. Let $\{g_k\}_{k=1}^\infty$ be the countable set satisfying Hypothesis I with $\|g_k\| = a_k$. We will continue to denote by the same symbol g_k the extension of

g_k with $g_k(\Delta) = a_k + 1$. Then $\{g_k\}_{k=1}^{\infty}$ separates points in E^{Δ}. Let

$$\hat{E} = \prod_{k=1}^{\infty} [-a_k - 1, a_k + 1]$$

and $\underline{g}$: $E^{\Delta} \to \hat{E}$ be defined by

$$\underline{g}(x) = (g_1(x), \dots, g_k(x), \dots). \tag{4.23}$$

Define, for $z \notin \underline{g}(E^{\Delta})$, $\underline{g}^{-1}(z) = e$ for some fixed point e in E^{Δ}. Let $\hat{\lambda} \in B(\hat{E})$ be defined by $\hat{\lambda} = \lambda \circ \underline{g}^{-1}$. Then

$$\hat{\lambda}(\underline{g}(x)) = \lambda(x) \qquad \forall\, x \in E^{\Delta}. \tag{4.24}$$

Define the operator $\mathscr{A}$ as in (4.4) with A replaced by A^{Δ} in the definition. Now, on $\mathscr{D}(\mathscr{A})$ define the operator $\mathscr{C}$ by

$$\mathscr{C}u(z) = \mathscr{A}u(z) + \hat{\lambda}(z)\big(u(\underline{g}(\Delta)) - u(z)\big). \tag{4.25}$$

Note that $Z(t) = \underline{g}(X(t))$ is a solution to the martingale problem for $\mathscr{C}$. Let $\hat{Z}$ be the r.c.l.l. modification of Z. Arguing as in (4.5) and using (4.22), we get that

$$I_{\{\tau > t\}} \exp\left\{ \int_0^t \hat{\lambda}(\hat{Z}(s))\, ds \right\} \tag{4.26}$$

is a nonnegative mean one martingale.

Fix $T > 0$. Define Q on $D([0,\infty), \hat{E})$ by

$$\begin{aligned} &Q(\theta(t_1) \in \Gamma_1, \dots, \theta(t_m) \in \Gamma_m) \\ &\quad = \mathbb{E}\left[\prod_{i=1}^{m} I_{\Gamma_i}(\hat{Z}(t_i)) I_{\{\tau > t_m\}} \exp\left\{ \int_0^{t_m} \hat{\lambda}(\hat{Z}(s))\, ds \right\} \right] \end{aligned} \tag{4.27}$$

for all $0 \le t_1 < \cdots < t_m \le T$ and all choices of Borel sets $\Gamma_1, \dots, \Gamma_m$. Here θ is the coordinate process on $D([0,\infty), \hat{E})$. Equation (4.27) defines a probability measure on $D([0,\infty), \hat{E})$, since $\hat{Z}$ is an r.c.l.l. process. Since X is a solution to the martingale problem for $A^{\Delta} + C$, we get, for $f \in \mathscr{D}(A^{\Delta} + C)$ with $f(\Delta) = 0$, that

$$f(X(t)) - \int_0^t (Af(X(s)) - \lambda(X(s)) f(X(s)))\, ds$$

is a martingale and hence using integration by parts, we get that

$$f(X(t)) \exp\left\{ \int_0^t \tilde{\lambda}(X(s))\, ds \right\} - \int_0^t Af(X(s)) \exp\left\{ \int_0^s \tilde{\lambda}(X(u))\, du \right\} ds$$

is a martingale. Since $f(\Delta) = 0$, using (4.21), we get that

$$\begin{aligned} &f(X(t)) I_{\{\tau > t\}} \exp\left\{ \int_0^t \lambda(X(s))\, ds \right\} \\ &\quad - \int_0^t Af(X(s)) I_{\{\tau > s\}} \exp\left\{ \int_0^s \lambda(X(u))\, du \right\} ds \end{aligned}$$

is a martingale. Or, arguing as in (4.5), for $u \in \mathscr{D}(\mathscr{C})$,

$$u(\hat{Z}(t))I_{\{\tau > t\}} \exp\left\{\int_0^t \hat{\lambda}(\hat{Z}(s))\,ds\right\} \tag{4.28}$$
$$- \int_0^t \mathscr{A}u(\hat{Z}(s))I_{\{\tau > s\}} \exp\left\{\int_0^s \hat{\lambda}(\hat{Z}(r))\,dr\right\} ds$$

is a martingale. Hence for $0 \le t_1 < \cdots < t_{m+1} \le T$, $h_1, \dots, h_m \in C(\hat{E})$,

$$\mathbb{E}^Q\left[\left(u(\theta(t_{m+1})) - u(\theta(t_m)) - \int_{t_m}^{t_{m+1}} \mathscr{A}u(\theta(s))\,ds\right) \prod_{k=1}^m h_k(\theta(t_k))\right]$$
$$= \mathbb{E}^P\left[\left(u(\hat{Z}(t_{m+1}))I_{\{\tau > t_{m+1}\}} \exp\left\{\int_0^{t_{m+1}} \hat{\lambda}(\hat{Z}(r))\,dr\right\}\right.\right.$$
$$- u(\hat{Z}(t_m))I_{\{\tau > t_m\}} \exp\left\{\int_0^{t_m} \hat{\lambda}(\hat{Z}(r))\,dr\right\} \tag{4.29}$$
$$\left.\left. - \int_{t_m}^{t_{m+1}} \mathscr{A}u(\hat{Z}(s))I_{\{\tau > s\}} \exp\left\{\int_0^s \hat{\lambda}(\hat{Z}(r))\,dr\right\}\right) ds \prod_{k=1}^m h_k(\hat{Z}(t_k))\right]$$
$$= 0.$$

It follows that under Q, θ is a solution to the martingale problem for $\mathscr{A}$ satisfying

$$Q\big(\theta(t) \in \underline{g}(E^\Delta)\big) = 1 \qquad \forall\, t. \tag{4.30}$$

Hence, $X'(t) := \underline{g}^{-1}(\theta(t))$ is a solution to the martingale problem for A^Δ. This step is similar to the arguments given at the beginning of this section. Thus we get for $u \in B(\hat{E})$:

$$\begin{aligned}\mathbb{E}^Q[u(\theta(t))] &= \mathbb{E}^Q\big[u \circ \underline{g}(X'(t))\big]\\ &= \mathbb{E}^Q\big[\big[T^\Delta(t)(u \circ \underline{g})\big](X'(0))\big]\\ &= \mathbb{E}^P\big[\big[T^\Delta(t)(u \circ \underline{g})\big](X(0))\big].\end{aligned}$$

This can be rephrased as

$$\mathbb{E}^P\left[u(\hat{Z}(t))\exp\left\{\int_0^t \hat{\lambda}(\hat{Z}(r))\,dr\right\}I_{\{\tau > t\}}\right] = \mathbb{E}^P\big[T^\Delta(t)(u \circ \underline{g})(X(0))\big] \tag{4.31}$$

for all $0 \le t \le T$. Similarly, if for $s > 0$, fixed, we define $\tilde{Q}$ on $D([0,\infty), \hat{E})$ by

$$\tilde{Q}(\theta(t_1)) \in \Gamma_1, \dots, \theta(t_m) \in \Gamma_m)$$
$$= \frac{\mathbb{E}\left[\prod_{i=1}^m I_{\Gamma_i}(\hat{Z}(s+t_i))I_{\{\tau > s+t_m\}} \exp\left\{\int_s^{s+t_m} \hat{\lambda}(\hat{Z}(r))\,dr\right\}\right]}{P(\tau > s)} \tag{4.32}$$

for all $0 \le t_1 < \cdots < t_m \le T$, for all choices of Borel sets $\Gamma_1, \dots, \Gamma_m$, then $\tilde{Q}$

is a solution to the martingale problem for $\mathscr{A}$ and we get

$$\mathbb{E}^P\left[u(\hat{Z}(t))\exp\left\{\int_s^t \hat{\lambda}(\hat{Z}(r))\,dr\right\}I_{\{\tau > t\}}\right] = \mathbb{E}^P\left[\left[T^\Delta(t-s)(u\circ \underline{g})\right](X(s))\right] \tag{4.33}$$

for all $s \le t \le s + T$. Since $T > 0$ was arbitrary, (4.31) and (4.33) hold for all $t \ge 0$ and $t \ge s$, respectively. Let $f \in B(E^\Delta)$ with $f(\Delta) \equiv 0$ and $u := f \circ \underline{g}^{-1}$. Then note that $u(\underline{g}(x)) = f(x)$ for all $x \in E^\Delta$. Then using $u(\hat{Z}(t)) \equiv 0$ if $\tau \le t$, we get

$$\begin{aligned}
&\mathbb{E}^P\left[u(\hat{Z}(t))\right] - \mathbb{E}^P\left[T^\Delta(t)f(X(0))\right]\\
&\quad= \mathbb{E}^P\left[u(\hat{Z}(t))\left(1-\exp\left\{\int_0^t \hat{\lambda}(\hat{Z}(r))\,dr\right\}I_{\{\tau > t\}}\right)\right]\\
&\quad= \mathbb{E}^P\left[u(\hat{Z}(t))\left(1-\exp\left\{\int_0^t \hat{\lambda}(\hat{Z}(r))\,dr\right\}\right)I_{\{\tau > t\}}\right]\\
&\quad= -\mathbb{E}^P\left[\int_0^t u(\hat{Z}(t))\hat{\lambda}(\hat{Z}(s))\exp\left\{\int_s^t \hat{\lambda}(\hat{Z}(r))\,dr\right\}I_{\{\tau > t\}}\,ds\right]\\
&\quad= -\mathbb{E}^P\left[\int_0^t \mathbb{E}^P\left[u(\hat{Z}(t))\exp\left\{\int_s^t \hat{\lambda}(\hat{Z}(r))\,dr\right\}I_{\{\tau > t\}}\Big|\mathscr{F}_s^{\hat{Z}}\right]\hat{\lambda}(\hat{Z}(s))\,ds\right]\\
&\quad= -\mathbb{E}^P\left[\int_0^t T^\Delta(t-s)f(X(s))\hat{\lambda}(\hat{Z}(s))\,ds\right],
\end{aligned} \tag{4.34}$$

or

$$\begin{aligned}
\mathbb{E}^P[f(X(t))] - \mathbb{E}^P\left[T^\Delta(t)f(X(0))\right]
&= -\mathbb{E}^P\left[\int_0^t T^\Delta(t-s)f(X(s))\lambda(X(s))\,ds\right]\\
&= \int_0^t \mathbb{E}^P\left[CT^\Delta(t-s)f(X(s))\right]ds.
\end{aligned} \tag{4.35}$$

Hence iterating, we get

$$\begin{aligned}
\mathbb{E}^P[f(X(t))] &= \mathbb{E}^P\left[T^\Delta(t)f(X(0))\right]\\
&\quad+ \int_0^t \mathbb{E}^P\left[T^\Delta(s)CT^\Delta(t-s)f(X(0))\right]ds\\
&\quad+ \int_0^t\int_0^s \mathbb{E}^P\left[CT^\Delta(s-r)CT^\Delta(t-s))f(X(r))\right]dr\,ds
\end{aligned}$$

and so on. Thus the distribution of $X(t)$ is determined by C, $\{T(s)\}_{s\ge 0}$ and $X(0)$. Hence the distribution of $X(t)$ is determined for every $t \ge 0$. But this implies that the finite dimensional distributions of X are determined. [See, e.g., Theorem IV.4.2 in Ethier and Kurtz (1986).] Hence we have well-posedness of the martingale problem for $A^\Delta + C$.

It now follows that $A^\Delta + C$ satisfies all the conditions of Theorem 3.2. Hence, applying Theorem 4.1 for the operator $A^\Delta + C$, we get uniqueness of solutions to the measure valued equation

$$\int_{E^\Delta} f\,d\rho_t = \int_{E^\Delta} f\,d\tilde{\nu} + \int_0^t \left(\int_{E^\Delta} (A^\Delta + C) f\,d\rho_s \right) ds. \tag{4.36}$$

Let $\{\nu_t\}$ be a solution to (4.12) [satisfying (4.2)]. Since $1 \in \mathscr{D}(A)$ with $A1 = 0$, we get

$$\nu_t(E) = \nu(E) - \int_0^t \lambda\,d\nu_s\,ds$$

and hence $\nu_t(E) \le 1$. Set $\tilde{\nu}_t(U) = \nu_t(U \cap E) + (1 - \nu_t(E)) I_U(\Delta)$ for U Borel in E^Δ. Then it can be checked that $\tilde{\nu}_t$ is a solution to (4.36). Thus, uniqueness of solutions to (4.36) implies the required uniqueness of solutions to (4.12). □

5. Evolution equation for time inhomogeneous Markov processes. The results of Section 4 can easily be extended to the case where the operator A depends on time. We define the time dependent martingale problem as follows.

Let (E, d) be a complete, separable metric space. For $t \ge 0$, let A_t be linear operators on $C_b(E)$ with a common domain $\mathscr{D} \subset C_b(E)$.

A measurable process X defined on some probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ is said to be a solution to the martingale problem for $(A_t)_{t \ge 0}$ with respect to a filtration $(\mathscr{G}_t)_{t \ge 0}$ if for any $f \in \mathscr{D}$,

$$f(X(t)) - \int_0^t A_s f(X(s))\,ds \tag{5.1}$$

is a $(\mathscr{G}_t)$-martingale.

For $\mu \in \mathscr{P}(E)$, we say that the martingale problem for $((A_t), \mu)$ is well-posed if, whenever, X and Y are two solutions to the martingale problem for (A_t), defined, respectively, on $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ and $(\tilde{\Gamma}, \tilde{\mathscr{G}}, \tilde{Q})$ with respect to some filtrations and satisfying $\tilde{P} \circ X(0)^{-1} = \tilde{Q} \circ Y(0)^{-1} = \mu$, we have

$$\tilde{P}(X(t) \in U) = \tilde{Q}(Y(t) \in U) \qquad \forall\, t > 0; \qquad \forall\, U, \text{Borel in } E.$$

Most of the results on martingale problems can be extended to the time dependent case by considering the space-time process

$$X^0(t) = (t, X(t)). \tag{5.2}$$

Let $E^0 = [0, \infty) \times E$. We state the following theorem. For its proof see, for example, Ethier and Kurtz (1986), page 221.

THEOREM 5.1. *Let $\mathscr{D}' \subset C_b(E^0)$ consist of functions of the form*

$$g(t, x) = \sum_{i=1}^k h_i(t) f_i(x) \qquad h_i \in C_0^1([0, \infty)),\ f_i \in \mathscr{D}. \tag{5.3}$$

Define an operator A^0 on $C_b(E^0)$ with domain $\mathscr{D}'$ by

$$A^0 g(t,x) = \sum_{i=1}^{k}\left[f_i(x)\frac{\partial}{\partial t}h_i(t) + h_i(t)A_t f_i(x)\right]. \tag{5.4}$$

Then X is a solution to the martingale problem for (A_t) if and only if X^0 is a solution to the martingale problem for A^0.

We have the following version of Theorem 4.1 in the time inhomogeneous case. This can be deduced from Theorem 4.1. See, for example, Horowitz and Karandikar (1990), where this is done for the locally compact case.

THEOREM 5.2. *Suppose that the operator A^0 defined by* (5.4) *with domain $\mathscr{D}'$ satisfies the conditions of Theorem* 3.2. *If $\{\nu_t\}_{t\geq 0} \subset \mathscr{P}(E)$ and $\{\mu_t\}_{t\geq 0} \subset \mathscr{P}(E)$ satisfy* (4.2) *and the equation*

$$\int_E f\,d\nu_t = \int_E f\,d\nu + \int_0^t\left(\int_E A_s f\,d\nu_s\right)ds, \qquad f\in\mathscr{D}, \tag{5.5}$$

then $\nu_t = \mu_t$ for all $t \geq 0$.

Now we consider the time dependent version of Theorem 4.2.

THEOREM 5.3. *Suppose that the operator A^0 defined by* (5.4) *with domain $\mathscr{D}'$ satisfies the conditions of Theorem* 4.2. *Let $(s,x)\mapsto \lambda_s(x)$ be a bounded continuous function on* $([0,T]\times E)$. *If $\{\nu_t\}_{t\geq 0} \subset \mathscr{M}(E)$ and $\{\mu_t\}_{t\geq 0} \subset \mathscr{M}(E)$ satisfy* (4.2) *and the equation*

$$\int_E f\,d\nu_t = \int_E f\,d\nu + \int_0^t\left(\int_E (A_s f + \lambda_s(\cdot)f)\,d\nu_s\right)ds, \qquad f\in\mathscr{D}, \tag{5.6}$$

then $\nu_t = \mu_t$ for all $t \geq 0$.

PROOF. Define E^Δ as in (4.14), $\lambda_t \in C_b(E^\Delta)$ and operators A_t^Δ and C_t for each $t \geq 0$ by (4.15)–(4.18), using λ_t and A_t in place of λ and A, respectively. On $\mathscr{D}'$ define operators C^0 and A^0 as in (5.4). We can verify that A^0 satisfies the conditions of Theorem 4.2. The result now follows from Theorem 4.2. □

6. Application to filtering theory. In this section, we will give an application of the results in the previous section to filtering theory. We recall here briefly the white noise model of filtering.

Suppose that the signal process (i.e., the process of interest) $(X(t))$ is a Markov process and that $(X(t))$ is not directly observable. Instead, one can observe a function $h_t(X(t))$ of the signal corrupted by additive noise (e_t)-assumed to be *white noise*. In other words the observation process y_t is

$$y_t = h_t(X(t)) + e_t, \tag{6.1}$$

where $\mathscr{H}_1$ is a separable Hilbert space, $h\colon [0,T]\times E \to \mathscr{H}_1$ is a measurable function such that $\int_0^T \|h_s(X(s))\|_1^2\,ds < \infty$ and (e_t) is $\mathscr{H}_1$ valued white noise.

The norm in $\mathscr{H}_1$ is denoted by $\|\cdot\|_1$ and the inner product by $(\cdot, \cdot)_1$. In the framework of countably additive probability theory, white noise (e_t) does not exist as a process and to formalize this model one has to proceed differently. [See Kallianpur and Karandikar (1988), Appendix and references therein.]

However on a finitely additive probability space, one can construct white noise (e_t) and then the model (6.1) can be given a formal meaning. The sample space for (e_t) and (y_t) is $L^2([0,T], \mathscr{H}_1)$. The quantity of interest in the filtering theory is the conditional distribution $F_t(y)$ of $(X(t))$ given $(y_s\colon 0 \le s \le t)$:

$$F_t(y)(B) = E[I_B(X(t))|y_s\colon 0 \le t]$$

for B Borel in E. We now state a result from Kallianpur and Karandikar [(1988), page 363–366]. For the meaning of conditional expectation in this setup and related matters, we refer the reader to Chapter 6 in the reference cited above. This result is also given in Kallianpur and Karandikar (1985).

Let $c_s^y(x) := (h_s(x), y_s)_1 - 1/2\|h_s(x)\|_1$. Then

$$F_t(y)(B) = \Gamma_t(y)(B) \cdot [\Gamma_t(y)(E)]^{-1},$$

where

$$\Gamma_t(y)(B) = E\left[I_B(X(t))\exp\left(\int_0^t c_s^y(X(s))\, ds\right)\right]. \tag{6.2}$$

$\Gamma_t(y)$ is called the unnormalized conditional distribution of $(X(t))$ given $(y_s\colon 0 \le s \le t)$. We can now deduce the following result from Theorem 5.3.

THEOREM 6.1. *Suppose that the signal process $(X(t))$ is the unique solution to the martingale problem for $((A_t), \nu)$ where (A_t) is as in Section 5. Suppose that the operator A^0 defined by (5.5) with domain $\mathscr{D}'$ satisfies the conditions of Theorem 4.2.*

Also suppose that h is a continuous function with $\|h.(\cdot)\|_1 \le C$ for some constant C. Then for all $y \in C([0, \Gamma], \mathscr{H}_1)$ the unnormalized conditional distribution $\Gamma_t(y)$ is the unique solution to the equation

$$\langle g, \Gamma_t(y)\rangle = \langle g, \nu\rangle + \int_0^T \langle A_s g + c_s^y g, \Gamma_s(y)\rangle\, ds, \qquad g \in \mathscr{D}. \tag{6.3}$$

We can equivalently state the above conclusion as: $\Gamma_t(y)$ is the unique solution to the equation

$$\begin{aligned}\langle f(t,\cdot), \Gamma_t(y)\rangle &= \langle f(0,\cdot), \nu\rangle \\ &\quad + \int_0^T \langle (A^0 f)(s,\cdot) + c_s^y(\cdot) f(s,\cdot), \Gamma_s(y)\rangle\, ds,\end{aligned} \tag{6.4}$$

$$f \in \mathscr{D}'.$$

It may be noted that in Kallianpur and Karandikar (1988), $\Gamma_t(y)$ has been characterized as the unique solution to (6.4), with A^0 replaced by the generator L of the Markov process $(t, X(t))$ and $\mathscr{D}'$ replaced by the domain $\mathscr{D}_L$ of L. In that case, h is not required to be bounded.

Though Theorem 6.1 requires h to be bounded, for $y \in C([0,T], \mathscr{H}_1)$, it yields $\Gamma_t(y)$ as the unique solution to (6.3) or equivalently (6.4). This is a significant improvement, since $\mathscr{D}_L$ can be very large and we have no control over it, whereas we can choose $\mathscr{D}'$ and in most cases we can choose it to be much smaller. When the signal process is an infinite dimensional diffusion (as in Remark 3.2), $\mathscr{D}$ can be taken to consist of cylinder functions, that is, functions depending upon finitely many coordinates, but $\mathscr{D}_L$ will contain functions which are not cylinder functions.

Even though Theorem 6.1 gives a characterization of $\Gamma_t(y)$ for $y \in C([0,T], \mathscr{H}_1)$, it is enough because it is known that $y \to \Gamma_t(y)$ is Lipschitz continuous [see Kallianpur and Karandikar (1988), page 479] and $C([0,T], \mathscr{H}_1)$ is dense in $L^2([0,T], \mathscr{H}_1)$.

REFERENCES

ECHEVERRIA, P. E. (1982). A criterion for invariant measures of Markov processes. *Z. Wahrsch. Verw. Gebiete* **61** 1–16.

ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

HOROWITZ, J. and KARANDIKAR, R. L. (1990). Martingale problems associated with the Boltzman equation. In *Seminar on Stochastic Processes, 1989* (E. Çinlar, K. L. Chung and R. K. Getoor, eds.). Birkhäuser, Boston.

KALLIANPUR, G. and KARANDIKAR, R. L. (1985). White noise calculus and non-linear filtering theory. *Ann. Probab.* **13** 1033–1107.

KALLIANPUR, G. and KARANDIKAR, R. L. (1988). *White Noise Theory Of Prediction Filtering And Smoothing.* Gordon and Breach, New York.

NEVEU, J. (1965). *Mathematical Foundations Of The Calculus Of Probability.* Holden-Day, San Francisco.

PARTHASARATHY, K. R. (1967). *Probability Measures On Metric Spaces.* Academic, New York.

STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes.* Springer, Berlin.

YOR, M. (1974). Existence et unicité de diffusions à valeurs dans un espace de Hilbert. *Ann. Inst. H. Poincaré Sec. B* **10** 55–88.

INDIAN STATISTICAL INSTITUTE
7 S.J.S. SANSANWAL MARG
NEW DELHI, 110016
INDIA