Equivalence of stochastic equations and martingale problems

Thomas G. Kurtz *
Departments of Mathematics and Statistics
University of Wisconsin - Madison
480 Lincoln Drive
Madison, WI 53706-1388
kurtz@math.wisc.edu
http://www.math.wisc.edu/∼kurtz/

July 3, 2010

Abstract

The fact that the solution of a martingale problem for a diffusion process gives a weak solution of the corresponding Itô equation is well-known since the original work of Stroock and Varadhan. The result is typically proved by constructing the driving Brownian motion from the solution of the martingale problem and perhaps an auxiliary Brownian motion. This constructive approach is much more challenging for more general Markov processes where one would be required to construct a Poisson random measure from the sample paths of the solution of the martingale problem. A "soft" approach to this equivalence is presented here which begins with a joint martingale problem for the solution of the desired stochastic equation and the driving processes and applies a Markov mapping theorem to show that any solution of the original martingale problem corresponds to a solution of the joint martingale problem. These results coupled with earlier results on the equivalence of forward equations and martingale problems show that the three standard approaches to specifying Markov processes (stochastic equations, martingale problems, and forward equations) are, under very general conditions, equivalent in the sense that existence and/or uniqueness of one implies existence and/or uniqueness for the other two.

MSC 2000 subject classifications: 60J25, 60H10, 60J35

Keywords: Markov processes, stochastic equations, martingale problems, forward equations, Markov mapping theorem, existence, uniqueness

*Research supported in part by NSF grant DMS 08-05793

1 Introduction.

Let X be a solution of an Itô equation

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds, \tag{1.1}$$

where X has values in $\mathbb{R}^d$, W is standard, m-dimensional Brownian motion, σ is a locally bounded $d \times m$-matrix-valued function, and b is a locally bounded $\mathbb{R}^d$-valued function. Let L be the corresponding differential generator

$$Lf(x) = \frac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}f(x) + \sum_i b_i(x)\frac{\partial}{\partial x_i}f(x).$$

If we define $Af = Lf$ for $f \in \mathcal{D}(A) \equiv C_c^2(\mathbb{R}^d)$, the twice continuously differentiable functions with compact support in $\mathbb{R}^d$, then it follows from Itô's formula and the properties of the Itô integral that

$$f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds = \int_0^t \nabla f(X(s))^T \sigma(X(s))dW(s) \tag{1.2}$$

is a martingale and hence that X is a solution of the martingale problem for A (or more precisely, the $C_{\mathbb{R}^d}[0,\infty)$-martingale problem for A). That the converse to this observation is, in a useful sense, true is an important fact observed early in the study of martingale problems for diffusion processes (Stroock and Varadhan (1972)). To state precisely the sense in which the assertion is true, we say that a process X with sample paths in $C_{\mathbb{R}^d}[0,\infty)$ is a *weak solution* of (1.1) if and only if there exists a probability space $(\Omega, \mathcal{F}, P)$ and stochastic processes $\widetilde{X}$ and $\widetilde{W}$ adapted to a filtration $\{\mathcal{F}_t\}$ such that $\widetilde{X}$ has the same distribution as X, $\widetilde{W}$ is an $\{\mathcal{F}_t\}$-Brownian motion, and

$$\widetilde{X}(t) = \widetilde{X}(0) + \int_0^t \sigma(\widetilde{X}(s))d\widetilde{W}(s) + \int_0^t b(\widetilde{X}(s))ds. \tag{1.3}$$

We then have

Theorem 1.1 *X is a solution of the $C_{\mathbb{R}^d}[0,\infty)$-martingale problem for A if and only if X is a weak solution of (1.1).*

Taking expectations in (1.2) we obtain the identity

$$\nu_t f = \nu_{\mathbf{0}} f + \int_0^t \nu_s Af ds, \quad f \in \mathcal{D}(A), \tag{1.4}$$

which is just the weak form of the forward equation for $\{\nu_t\}$, the one-dimensional distributions of X. The converse of the observation that every solution of the martingale problem gives a solution of the forward equation is also true, and we have the following theorem. (See the construction in Ethier and Kurtz (1986), Theorem 4.9.19, or Kurtz (1998), Theorem 2.6.)

Theorem 1.2 *If X is a solution of the martingale problem for A, then $\{\nu_t\}$, the one-dimensional distributions of X, is a solution of the forward equation (1.4). If $\{\nu_t\}$ is a solution of (1.4), then there exists a solution X of the martingale problem for A such that $\{\nu_t\}$ are the one-dimensional distributions of X.*

Note that Theorem 1.1, as stated, applies to solutions of the $C_{\mathbb{R}^d}[0,\infty)$-martingale problem, that is, solutions whose sample paths are in $C_{\mathbb{R}^d}[0,\infty)$, while Theorem 1.2 does not have this restriction. In general, we cannot rule out the possibility that a solution of (1.1) hits infinity in finite time unless we add additional restrictions to the coefficients. One way around this issue is to allow X to take values in $\mathbb{R}^{d\Delta}$, the one-point compactification of $\mathbb{R}^d$, and to allow ν_t to be in $\mathcal{P}(\mathbb{R}^{d\Delta})$. To avoid problems with the definition of the stochastic integral in (1.1), we can replace (1.1) by the requirement that (1.2) hold for all $f \in C_c^2(\mathbb{R}^d)$, extending f to $\mathbb{R}^{d\Delta}$ by defining $f(\Delta) = 0$.

Given an initial distribution $\nu_0 \in \mathcal{P}(\mathbb{R}^{d\Delta})$, we say that uniqueness holds for the martingale problem (or $C_{\mathbb{R}^d}[0,\infty)$-martingale problem) for (A, ν_0) if any two solutions of the martingale problem (resp. $C_{\mathbb{R}^d}[0,\infty)$-martingale problem) for A with initial distribution ν_0 have the same finite dimensional distributions. Similarly, weak uniqueness holds for (1.2) (or (1.1)) with initial distribution ν_0 if any two weak solutions of (1.2) (resp. (1.1)) with initial distribution ν_0 have the same finite dimensional distributions, and uniqueness holds for the forward equation (1.4) if any two solutions with initial distribution ν_0 are the same.

Note that neither Theorem 1.1 nor Theorem 1.2 assumes uniqueness. Consequently, existence and uniqueness for the three problems are equivalent.

Corollary 1.3 *Let $\nu_0 \in \mathcal{P}(\mathbb{R}^d)$. The following are equivalent:*

a) Uniqueness holds for the martingale problem for (A, ν_0).

b) Weak uniqueness holds for (1.2) with initial distribution ν_0.

c) Uniqueness holds for (1.4) with initial distribution ν_0.

The usual proof of Theorem 1.1 involves the construction of W in terms of the given solution X of the martingale problem. If $d = m$ and σ is nonsingular, this construction is simple. In particular, if we define

$$M(t) = X(t) - \int_0^t b(X(s))ds,$$

then

$$W(t) = \int_0^t \sigma^{-1}(X(s))dM(s),$$

where σ^{-1} denotes the inverse of σ. If σ is singular, the construction involves an auxiliary Brownian motion independent of X. (See, for example, Stroock and Varadhan (1979), Theorems 4.5.1 and 4.5.2, or Ethier and Kurtz (1986), Theorem 5.3.3.)

A possible alternative approach is to consider the process $Z = (X, Y) = (X, Y(0) + W)$. Of course

$$dZ(t) = d\begin{pmatrix} X(t) \\ Y(0) + W(t) \end{pmatrix} = \begin{pmatrix} \sigma(X(t)) \\ I \end{pmatrix} dW(t) + \begin{pmatrix} b(X(t) \\ 0 \end{pmatrix} dt. \qquad (1.5)$$

Note that each weak solution of (1.1) gives a weak solution of (1.5) and each weak solution of (1.5) gives a weak solution of (1.1). As before, using Itô's formula, it is simple to compute the generator $\widehat{A}$ corresponding to (1.5) (take the domain to be $C_c^2(\mathbb{R}^{d+m})$). Furthermore, since if one knows Z one knows W, it follows immediately that every solution of the martingale problem for $\widehat{A}$ is a weak solution of the stochastic differential equation. In particular, weak uniqueness for (1.5) implies uniqueness for the martingale problem for $\widehat{A}$. Note, however, that the assertion that every solution of the martingale problem for $\widehat{A}$ is a weak solution of (1.5) (and hence gives a weak solution of (1.1)) does not immediately imply that every solution of the martingale problem for A is a weak solution of (1.1) since we must obtain the driving Brownian motion. In particular, we cannot immediately conclude that uniqueness for (1.1) implies uniqueness for the martingale problem for A.

In fact, however, an argument along the lines described can be used to show that each solution of the martingale problem for A is a weak solution of (1.1). For simplicity, assume $d = m = 1$. Instead of augmenting the state by $Y(0) + W$, augment the state by

$$Y(t) = Y(0) + W(t) \bmod 2\pi.$$

We can still recover W from observations of the increments of Y. For example, if we set

$$\zeta(t) = \begin{pmatrix} \cos(Y(t)) + \int_0^t \frac{1}{2}\cos(Y(s))ds \\ \sin(Y(t)) + \int_0^t \frac{1}{2}\sin(Y(s))ds \end{pmatrix}, \tag{1.6}$$

and

$$W(t) = \int_0^t (-\sin(Y(s)), \cos(Y(s)))d\zeta(s), \tag{1.7}$$

then W is a standard Brownian motion and ζ satisfies

$$d\zeta(t) = \begin{pmatrix} -\sin(Y(t) \\ \cos(Y(t)) \end{pmatrix} dW(t). \tag{1.8}$$

The introduction of Y may look strange, but the heart of our argument depends on being able to compute the conditional distribution of $Y(t)$ given $\mathcal{F}_t^X \equiv \sigma(X(s) : s \leq t)$. If $Y(0)$ is uniformly distributed on $[0, 2\pi]$ and is independent of W, then the conditional distribution of $Y(t)$ given $\mathcal{F}_t^X$ is uniform on $[0, 2\pi]$. In fact, that is the conditional distribution even if we condition on both X and W.

Let $\mathcal{D}(\widehat{A})$ be the collection of $f \in C_c^2(\mathbb{R}\times[0, 2\pi))$ such that $f(x, 0) = f(x, 2\pi-)$, $f_y(x, 0) = f_y(x, 2\pi-)$, and $f_{yy}(x, 0) = f_{yy}(x, 2\pi-)$. Applying Itô's formula, for $f \in \mathcal{D}(\widehat{A})$, we have

$$\widehat{A}f = \frac{1}{2}\sigma^2 f_{xx} + \sigma f_{xy} + \frac{1}{2}f_{yy} + bf_x.$$

Suppose $Z = (X, Y)$ is a solution of the martingale problem for $\widehat{A}$, and define ζ by (1.6) and W by (1.7). Applying Lemma A.1 with $f_1(x, y) = f(x)$, $f_2(x, y) = \cos(y)$, $f_3(x, y) = \sin(y)$, $g_1(x, y) = 1$, $g_2(x, y) = f'(x)\sigma(x)\sin(y)$, and $g_3(x, y) = -f'(x)\sigma(x)\cos(y)$ implies

$$\begin{aligned} M(t) &= f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds \\ &\quad + \int_0^t f'(X(s))\sigma(X(s))\sin(Y(s))d\zeta_1(s) - \int_0^t f'(X(s))\sigma(X(s))\cos(Y(s))d\zeta_2(s) \end{aligned}$$

satisfies $\langle M\rangle \equiv 0$ so $M \equiv 0$ and hence

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s))\sigma(X(s))dW(s) + \int_0^t Af(X(s))ds. \tag{1.9}$$

It follows that any solution of the martingale problem for $\widehat{A}$ satisfying $\sup_{s\leq t}|X(s)| < \infty$ a.s. for each t is a weak solution of (1.1).

Of course, this last observation does not prove Theorem 1.1. We still have the question of whether or not every solution of the martingale problem for A corresponds to a solution of the martingale problem for $\widehat{A}$. The following result from Kurtz (1998) provides the tools needed to answer this question affirmatively. Let (E, r) be a complete, separable metric space, $B(E)$, the bounded, measurable functions on E, and $\overline{C}(E)$, the bounded continuous functions on E. If E is locally compact, then $\widehat{C}(E)$ will denote the continuous functions vanishing at infinity. We say that an operator $B \subset B(E) \times B(E)$ is *separable* if there exists a countable subset $\{g_k\} \subset \mathcal{D}(B)$ such that B is contained in the bounded, pointwise closure of the linear span of $\{(g_k, Bg_k)\}$. B is a pre-generator if it is dissipative and there are sequences of functions $\mu_n : E \to \mathcal{P}(E)$ and $\lambda_n : E \to [0,\infty)$ such that for each $(f, g) \in B$

$$g(x) = \lim_{n\to\infty} \lambda_n(x) \int_{\mathbb{S}} (f(y) - f(x))\mu_n(x, dy), \tag{1.10}$$

for each $x \in E$.

For a measurable, E_0-valued process U, $\widehat{\mathcal{F}}_t^U$ is the completion of $\sigma(\int_0^r h(U(s))ds : r \leq t, h \in B(E_0)) \vee \sigma(U(0))$. Let $\mathbf{T}^U = \{t : U(t) \text{ is } \widehat{\mathcal{F}}_t^U\text{-measurable}\}$. ($\mathbf{T}^U$ has full Lebesgue measure, and if U is cadlag with no fixed points of discontinuity, then $\mathbf{T}^U = [0,\infty)$. See Appendix A.2 of Kurtz and Nappo (2009).) Let $M_{E_0}[0,\infty)$ be the space of measurable functions from $[0,\infty)$ to E_0 topologized by convergence in Lebesgue measure.

Theorem 1.4 *Suppose that $B \subset \overline{C}(E) \times \overline{C}(E)$ is separable and a pre-generator and that $\mathcal{D}(B)$ is closed under multiplication and separates points in E. Let (E_0, r_0) be a complete, separable metric space, $\gamma : E \to E_0$ be Borel measurable, and α be a transition function from E_0 into E ($y \in E_0 \to \alpha(y, \cdot) \in \mathcal{P}(E)$ is Borel measurable) satisfying $\alpha(y, \gamma^{-1}(y)) = 1$. Define*

$$C = \{(\int_E f(z)\alpha(\cdot, dz), \int_E Bf(z)\alpha(\cdot, dz)) : f \in \mathcal{D}(B)\}.$$

Let $\mu_0 \in \mathcal{P}(E_0)$, and define $\nu_0 = \int \alpha(y, \cdot)\mu_0(dy)$. If $\widetilde{U}$ is a solution of the martingale problem for (C, μ_0), then there exists a solution V of the martingale problem for (B, ν_0) such that $\widetilde{U}$ has the same distribution on $M_{E_0}[0,\infty)$ as $U = \gamma \circ V$ and

$$P\{V(t) \in \Gamma | \widehat{\mathcal{F}}_t^U\} = \alpha(U(t), \Gamma), \quad \Gamma \in \mathcal{B}(E), t \in \mathbf{T}^U. \tag{1.11}$$

If $\widetilde{U}$ (and hence U) has a modification with sample paths in $D_E[0,\infty)$, then the modified $\widetilde{U}$ and U have the same distribution on $D_E[0,\infty)$.

Assume that σ and b in (1.1) are continuous. (This assumption can be removed with the application of more complicated technology. See Section 4.) Let B in the statement of

Theorem 1.4 be $\widehat{A}$, $E = \mathbb{R} \times [0, 2\pi)$, $E_0 = \mathbb{R}$, $\gamma(x, y) = x$, and for $f \in B(\mathbb{R} \times [0, 2\pi))$, define $\alpha f(x) = \frac{1}{2\pi} \int_0^{2\pi} f(x, y)dy$. For $f \in \mathcal{D}(\widehat{A})$, a straight forward calculation gives

$$\alpha \widehat{A} f(x) = A\alpha f(x),$$

so $A = C$. It follows that if X is a solution of the martingale problem for A, then there exists a solution $(\widetilde{X}, \widetilde{Y})$ of the martingale problem for $\widehat{A}$ such that X and $\widetilde{X}$ have the same distribution. Consequently, if X has sample paths in $C_{\mathbb{R}}[0, \infty)$, then X is a weak solution for (1.1), and Theorem 1.1 follows. Every solution of the martingale problem for A will have a modification with sample paths in $D_{\mathbb{R}^\Delta}[0, \infty)$, where $\mathbb{R}^\Delta$ denotes the one-point compactification of $\mathbb{R}$, and any solution with sample paths in $D_{\mathbb{R}}[0, \infty)$ will, in fact, have sample paths in $C_{\mathbb{R}}[0, \infty)$.

Invoking Theorem 1.4 is obviously a much less straight forward approach to Theorem 1.1 than the usual argument; however, the state augmentation approach extends easily to much more general settings in which the constructive argument becomes technically very complicated if not impossible.

2 Stochastic differential equations for Markov processes.

Typically, a Markov process X in $\mathbb{R}^d$ has a generator of the form

$$Af(x) = \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + \widehat{b}(x) \cdot \nabla f(x) + \int_{\mathbb{R}^d} (f(x+y) - f(x) - \mathbf{1}_{B_1}(y) y \cdot \nabla f(x)) \eta(x, dy)$$

where B_1 is the ball of radius 1 centered at the origin and η satisfies

$$\int 1 \wedge |y^2| \eta(x, dy) < \infty \tag{2.1}$$

for each x. (See, for example, Stroock (1975), Çinlar, Jacod, Protter, and Sharpe (1980).) The three terms are, respectively, the diffusion term, the drift term, and the jump term. In particular, $\eta(x, \Gamma)$ gives the "rate" at which jumps satisfying $X(s) - X(s-) \in \Gamma$ occur. Note that B_1 can be replaced by any set C containing an open neighborhood of the origin provided that the drift term is replaced by

$$b_C(x) \cdot \nabla f(x) = \left(b(x) + \int_{\mathbb{R}^d} y(\mathbf{1}_C(y) - \mathbf{1}_{B_1}(y)) \eta(x, dy) \right) \cdot \nabla f(x).$$

Suppose that there exist $\lambda : \mathbb{R}^d \times S \to [0, 1]$, $\gamma : \mathbb{R}^d \times S \to \mathbb{R}^d$, and a σ-finite measure ν on a measurable space $(S, \mathcal{S})$ such that

$$\eta(x, \Gamma) = \int_S \lambda(x, u) \mathbf{1}_\Gamma(\gamma(x, u)) \nu(du).$$

This representation is always possible. In fact, there are many such representations. For example, we can rewrite

$$\begin{aligned}\eta(x,\Gamma) &= \int_S \lambda(x,u)\mathbf{1}_{[0,1]}(|\gamma(x,u|))\mathbf{1}_\Gamma(\gamma(x,u))\nu(du) \\ &\qquad + \int_S \lambda(x,u)\mathbf{1}_{(1,\infty)}(|\gamma(x,u|))\mathbf{1}_\Gamma(\gamma(x,u))\nu(du) \\ &= \int_S \lambda_1(x,u)\mathbf{1}_\Gamma(\gamma(x,u))\nu(du) + \int_S \lambda_2(x,u)\mathbf{1}_\Gamma(\gamma(x,u))\nu(du) \\ &= \int_{S_1} \lambda(x,u)\mathbf{1}_\Gamma(\gamma(x,u))\nu(du) + \int_{S_2} \lambda(x,u)\mathbf{1}_\Gamma(\gamma(x,u))\nu(du),\end{aligned}$$

where S_1 and S_2 are copies of S and λ on S_1 is given by λ_1 and λ on S_2 is given by λ_2. Noting that $\mathbf{1}_{S_1}(u) = \mathbf{1}_{B_1}(\gamma(x,u))$, we can replace S by $S_1 \cup S_2$, and assuming

$$\int_S \lambda(x,u)(\mathbf{1}_{S_1}(u)|\gamma(x,u)|^2 + \mathbf{1}_{S_2}(u))\nu(du) < \infty,$$

$$\begin{aligned}Af(x) &= \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} f(x) + b(x)\cdot\nabla f(x) \\ &\qquad + \int_S \lambda(x,u)(f(x+\gamma(x,u)) - f(x) - \mathbf{1}_{S_1}(u)\gamma(x,u)\cdot\nabla f(x))\nu(du).\end{aligned} \tag{2.2}$$

We will take $\mathcal{D}(A) = C_c^2(\mathbb{R}^d)$ and assume that for $f \in \mathcal{D}(A)$, $Af \in \overline{C}(\mathbb{R}^d)$. Removal of the continuity assumption will be discussed in Section 4. The assumption that Af is bounded can also be relaxed, but that issue is not addressed here.

Let ξ be a Poisson random measure on $[0,1]\times S\times[0,\infty)$ with mean measure $m\times\nu\times m$, and let $\widetilde{\xi}(A) = \xi(A) - m\times\nu\times m(A)$. Let $(S_0, \mathcal{S}_0)$ be a measurable space, μ a σ-finite measure on $(S_0, \mathcal{S}_0)$, W a Gaussian white noise on $S_0\times[0,\infty)$ satisfying $E[W(A,s)W(B,t)] = s\wedge t\mu(A\cap B)$, and $\sigma : \mathbb{R}^d\times S_0 \to \mathbb{R}^d$ satisfying $\int_{S_0}|\sigma(x,u)|^2\mu(du) < \infty$ and

$$a(x) = \int_{S_0}\sigma(x,u)\sigma^T(x,u)\mu(du).$$

Again, there are many possible choices for μ and σ. The usual form for an Itô equation corresponds to taking μ to be counting measure on a finite set S_0.

Assume that for each compact $K \subset \mathbb{R}^d$

$$\begin{aligned}\sup_{x\in K}\Big(|b(x)| + \int_{S_0}|\sigma(x,u)|^2\mu(du) + \int_{S_1}\lambda(x,u)|\gamma(x,u)|^2\nu(du) \\ + \int_{S_2}\lambda(x,u)|\gamma(x,u)|\wedge 1\nu(du)\Big) < \infty.\end{aligned} \tag{2.3}$$

Then X should satisfy a stochastic differential equation of the form

$$
\begin{aligned}
X(t) &= X(0) + \int_{S_0\times[0,t]} \sigma(X(s),u)W(du\times ds) + \int_0^t b(X(s))ds \\
&\quad + \int_{[0,1]\times S_1\times[0,t]} \mathbf{1}_{[0,\lambda(X(s-),u)]}(v)\gamma(X(s-),u)\widetilde{\xi}(dv\times du\times ds) \\
&\quad + \int_{[0,1]\times S_2\times[0,t]} \mathbf{1}_{[0,\lambda(X(s-),u)]}(v)\gamma(X(s-),u)\xi(dv\times du\times ds),
\end{aligned} \tag{2.4}
$$

for $t < \tau_\infty \equiv \lim_{k\to\infty} \inf\{t : |X(t-)| \text{ or } |X(t)| \geq k\}$. Stochastic equations of this form appeared first in Itô (1951).

An application of Itô's formula again shows that any solution of (2.4) gives a solution of the martingale problem for A. We will apply an extension of Theorem 1.4 to show that every solution of the $D_{\mathbb{R}^d}[0,\infty)$-martingale problem for A is a weak solution of (2.4), or more generally, we can replace (2.4) by the analog of (1.2) and drop the requirement that the solution have sample paths in $D_{\mathbb{R}^d}[0,\infty)$. (In any case, the solution will have a modification with sample paths in $D_{\mathbb{R}^{d\Delta}}[0,\infty)$.)

As in the introduction, we will need to represent the driving processes W and ξ in terms of processes whose conditional distributions given X are their stationary distributions. To avoid the danger of measure-theoretic or functional-analytic *faux-pas*, we will assume that S and S_0 are complete, separable metric spaces and that ν and μ are σ-finite Borel measures.

2.1 Representation of W by stationary processes.

Let $\varphi_1, \varphi_2, \ldots$ be a complete, orthonormal basis for $L_2(\mu)$. Then W is completely determined by

$$
W(\varphi_i, t) = \int_{S_0\times[0,t]} \varphi_i(u)W(du\times ds), \quad i = 1,2,\ldots.
$$

In particular, if H is an $\{\mathcal{F}_t\}$-adapted process with sample paths in $D_{L_2(\mu)}[0,\infty)$, then

$$
\int_{S_0\times[0,t]} H(s-,u)W(du\times ds) = \sum_{i=1}^\infty \int_0^t \langle H(s-,\cdot),\varphi_i\rangle dW(\varphi_i,s).
$$

In turn, if we define $Y_i(t) = Y_i(0) + W(\varphi_i,t) \bmod 2\pi$ and

$$
\zeta_i(t) = \begin{pmatrix} \cos(Y_i(t)) + \int_0^t \frac{1}{2}\cos(Y_i(s))ds \\ \sin(Y_i(t)) + \int_0^t \frac{1}{2}\sin(Y_i(s))ds \end{pmatrix} = \begin{pmatrix} -\int_0^t \sin(Y_i(s))dW(\varphi_i,s) \\ \int_0^t \cos(Y_i(s))dW(\varphi_i,s) \end{pmatrix},
$$

then

$$
W(\varphi_i,t) = \int_0^t \left(-\sin(Y_i(s)),\, \cos(Y_i(s))\right) d\zeta_i(s), \tag{2.5}
$$

and hence,

$$
\int_{S_0\times[0,t]} H(s,u)W(du\times ds) = \sum_{i=1}^\infty \int_0^t \langle H(s,\cdot),\varphi_i\rangle(-\sin(Y_i(s)),\cos(Y_i(s)))d\zeta_i(s).
$$

Note that if $Y_i(0)$ is uniformly distributed on $[0, 2\pi)$ and independent of W, then Y_i is a stationary process and for each t, the $Y_i(t)$ are independent and independent of $\sigma(W(\varphi_j, s) : s \leq t, j = 1, 2, \ldots)$. Identifying 2π with 0, $[0, 2\pi)$ is compact and $Y = \{Y_i\}$ is a Markov process with compact state space $[0, 2\pi)^\infty$

2.2 Representation of ξ by stationary processes.

Let $\{D_i\} \subset \mathcal{B}(S)$ be a partition of S satisfying $\nu(D_i) < \infty$, and define $\xi_i(C_1 \times C_2 \times [0, t]) = \xi(C_1 \times C_2 \cap D_i \times [0, t])$. Then the ξ_i are independent Poisson random measures, and setting $N_i(t) = \xi([0, 1] \times D_i \times [0, t])$, ξ_i can be written as

$$\xi_i(\cdot \times [0, t]) = \sum_{i=0}^{N_i(t)-1} \delta_{(V_{i,k}, U_{ik})},$$

where $\{V_{i,k}, U_{i,k}, i \geq 1, k \geq 0\}$ are independent, $V_{i,k}$ is uniform-$[0, 1]$, and $U_{i,k}$ is D_i-valued with distribution

$$\beta_i \equiv \frac{\nu(\cdot \cap D_i)}{\nu(D_i)}.$$

Define

$$Z_i(t) = (V_{i,N_i(t)}, U_{i,N_i(t)}).$$

Then Z_i is a Markov process with stationary distribution $\ell \times \beta_i$, where ℓ is the uniform distribution on $[0, 1]$, and $Z_i(t)$ is independent of $\sigma(\xi(\cdot \times [0, s]), s \leq t)$.

Since, with probability one, $V_{i,k} \neq V_{i,k+1}$, N_i can be recovered from Z_i, and since

$$\int_{[0,1]\times S\times[0,t]} H(v, u, s-)\xi(dv \times du \times ds) = \sum_i \int_0^t H(Z_i(s-), s-)dN_i(s),$$

ξ can be recovered from $\{Z_i\}$.

2.3 Equivalence to martingale problem

To simplify notation, we will replace $\mathbf{1}_{[0,\lambda(x,u)]}(v)\gamma(x, u)$ by $\gamma(x, u)$. There is no loss of generality since S is arbitrary and we can replace $[0, 1] \times S$ by S. Under the new notation, ξ is a Poisson random measure on $S \times [0, \infty)$ with mean measure $\nu \times m$. We will also assume that ν is nonatomic so it is still the case that, with probability one, N_i can be recovered from observations of Z_i.

Let $\mathcal{D}_0 \subset C^2([0, 2\pi))$ be the collection of functions satisfying $f(0) = f(2\pi-)$, $f'(0) = f'(2\pi-)$, and $f''(0) = f''(2\pi-)$, and let $\mathcal{D}_i = \overline{C}(D_i)$. Define

$$\mathcal{D}(\widehat{A}) = \{f_0(x) \prod_{i=1}^{m_1} f_{1i}(y_i) \prod_{i=1}^{m_2} f_{2i}(z_i) : f_0 \in C_c^2(\mathbb{R}^d), f_{1i} \in \mathcal{D}_0, f_{2i} \in \mathcal{D}_i\},$$

and for $f \in \mathcal{D}(\widehat{A})$, derive $\widehat{A}f$ by applying Itô's formula to

$$f_0(X(t)) \prod_{i=1}^{m_1} f_{1i}(Y_i(t)) \prod_{i=1}^{m_2} f_{2i}(Z_i(t)).$$

Define L_x by

$$L_x f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} f(x) + b(x)\cdot\nabla f(x),$$

and L_y by

$$L_y f(y) = \frac{1}{2}\sum_k \frac{\partial^2}{\partial y_k^2} f(y).$$

Note that L_x would be the generator for X if γ were zero and L_y is the generator for $Y = \{Y_i\}$. The quadratic covariation of X_i and Y_k is

$$[X_i, Y_k] = \int_0^t c_{ik}(X(s))ds,$$

where $c_{ik}(x) = \int_{S_0} \sigma_i(x,u)\varphi_k(u)\mu(du)$, so define L_{xy} by

$$L_{xy} f(x,y) = \sum_{i,k} c_{ik}(x)\partial_{x_i}\partial_{y_k} f(x,y).$$

For $u \in S$ and $z \in \prod_i D_i$, let $\rho(z,u)$ be the element of $\prod_i D_i$ obtained by replacing z_i by u provided $u \in D_i$. Define

$$\begin{aligned}
J_i f(x,y,z) &= \int_{D_i} \Big(f(x+\gamma(x,z_i), y, \rho(z,u)) - f(x,y,z) \\
&\qquad\qquad -\mathbf{1}_{S_1}(u)\gamma(x,u)\cdot\nabla_x f(x,y,z)\Big)\nu(du) \\
&= \int_{D_i} \Big(f(x+\gamma(x,z_i), y, \rho(z,u)) - f(x+\gamma(x,u),y,z)\Big)\nu(du) \\
&\quad + \int \Big(f(x+\gamma(x,u),y,z) - f(x,y,z) - \mathbf{1}_{S_1}(u)\gamma(x,u)\cdot\nabla_x f(x,y,z)\Big)
\end{aligned}$$

Then, at least formally, by Itô's formula,

$$f(X(t),Y(t),Z(t)) - f(X(0),Y(0),Z(0)) - \int_0^t \widehat{A} f(X(s),Y(s),Z(s))ds$$

is a martingale for

$$\begin{aligned}
\widehat{A} f(x,y,z) &= L_x f(x,y,z) + L_y f(x,y,z) + L_{xy} f(x,y,z) + \sum_i J_i f(x,y,z) \\
&= \prod_{i=1}^{m_1} f_{1i}(y_i) \prod_{i=1}^{m_2} f_{2i}(z_i) A f_0(x) + L_y f(x,y,z) + L_{xy} f(x,y,z) \\
&\quad + \sum_i \int_{D_i} \Big(f(x+\gamma(x,z_i), y, \rho(z,u)) - f(x+\gamma(x,u),y,z)\Big)\nu(du).
\end{aligned}$$

Unfortunately, in general, $\sum_i J_i$ may not converge. Consequently, the extension needs to be done one step at a time, so define $Z^n = (Z_1, \ldots, Z_n)$ and observe that the generator for (X, Y, Z^n) is

$$\begin{aligned}\widehat{A}_n f(x,y,z) &= L_x f(x,y,z) + L_y f(x,y,z) + L_{xy} f(x,y,z) + \sum_{i=1}^{n} J_i f(x,y,z) \\ &= \prod_{i=1}^{m_1} f_{1i}(y_i) \prod_{i=1}^{m_2} f_{2i}(z_i) A f_0(x) + L_y f(x,y,z) + L_{xy} f(x,y,z) \\ &\quad + \sum_{i=1}^{n} \int_{D_i} \Big(f(x+\gamma(x,z_i), y, \rho(z,u)) - f(x+\gamma(x,u), y, z) \Big) \nu(du),\end{aligned}$$

where we take $\mathcal{D}(\widehat{A}_n) = \{f \in \mathcal{D}(\widehat{A}) : m_2 \le n\}$. Note that as long as $A f_0 \in B(\mathbb{R}^d)$, $\widehat{A}_n f \in B(\mathbb{R}^d \times [0, 2\pi)^\infty \times \prod_{i=1}^n D_i)$.

Instead of requiring (X, Y, Z) to be a solution of the martingale problem for $\widehat{A}$, for each n, we require (X, Y, Z^n) to be a solution of the martingale problem for $\widehat{A}_n$.

Lemma 2.1 *If for each n, (X, Y, Z^n) is a solution of the martingale problem for $\widehat{A}_n$ with sample paths in $D_{\mathbb{R}^{d\Delta} \times [0,2\pi)^\infty \times \prod_{i=1}^n D_i}[0,\infty)$, W is given by (2.5), and ξ is given by*

$$\int_{S\times[0,t]} g(u)\xi(du \times ds) = \sum_{i=1}^{\infty} \int_0^t g(Z_i(s-)) dN_i(s),$$

then (X, W, ξ) satisfies (2.4) for $0 \le t < \tau_\infty$.

Remark 2.2 *Any process (X, Y, Z) such that for each n, (X, Y, Z^n) is solution of the martingale problem for $\widehat{A}_n$ will have a modification with sample paths in $D_{\mathbb{R}^{d\Delta} \times [0,2\pi)^\infty \times \prod_{i=1}^\infty D_i}[0,\infty)$ and the modification will satisfy (2.6) for all $f \in C_c^2(\mathbb{R}^d)$, taking $f(\Delta) = 0$.*

Proof. As in the verification of (1.9) Lemma A.1 can again be used to show that (X, W, ξ) satisfies

$$\begin{aligned}f(X(t)) = f(X(0)) &+ \int_0^t Af(X(s))ds + \int_0^t \nabla f(X(s))^T \sigma(X(s),u) W(du \times ds) \qquad (2.6) \\ &+ \int_{S\times[0,t]} (f(X(s-) + \gamma(X(s-),u)) - f(X(s-))\widetilde{\xi}(du \times ds),\end{aligned}$$

$f \in C_c^2(\mathbb{R}^d)$, $t \ge 0$, and it follows that X satisfies (2.4) for $0 \le t < \tau_\infty$. □

Theorem 2.3 *Let A be given by (2.2), and assume that (2.3) is satisfied and that for $f \in C_c^2(\mathbb{R}^d)$, $Af \in B(\mathbb{R}^d)$. Then any solution of the $D_{\mathbb{R}^d}[0,\infty)$-martingale problem for A is a weak solution of (2.4). More generally, any solution of the martingale problem for A has a modification with sample paths in $D_{\mathbb{R}^{d\Delta}}[0,\infty)$ and is a weak solution of (2.4) on the time interval $[0, \tau_\infty)$.*

Remark 2.4 *We need to relax the requirement in Theorem 1.4 that $\mathcal{R}(B) \subset \overline{C}(E)$. This extension is discussed in Section 4.*

Proof. Let $\beta_y \in \mathcal{P}([0,2\pi)^\infty)$ be the product of uniform distributions on $[0,2\pi)$ and $\beta_z^n \in \prod_{i=1}^n \beta_i$. For $x \in \mathbb{R}^d$, $\alpha_n(x,\cdot) = \delta_x \times \beta_y \times \beta_z^n \in \mathcal{P}(\mathbb{R}^d \times [0,2\pi)^\infty \times \prod_{i=1}^n D_i)$. ($\gamma_n$ is just the projection onto $\mathbb{R}^d$.) Computing $\alpha_n \widehat{A}_n f$, observe that $\alpha_n L_x f = L_x \alpha_n f$, that $\alpha_n L_y f = 0$ since β_y is the stationary distribution for L_y, and that $\alpha_n L_{xy} f = 0$ since $\int_0^{2\pi} \partial_{y_k} f(x,y,z) dy_k = 0$. To see that $\alpha_n J_i f = J_i \alpha_n f$, note that

$$\int_{D_i} \int_{D_i} f(x+\gamma(x,z_i), y, \rho(z,u)) \nu(du) \nu(dz_i) = \int_{D_i} \int_{D_i} f(x+\gamma(x,u), y, z) \nu(du) \nu(dz_i).$$

Taking these observations together, we have $\alpha_n \widehat{A}_n f = A \alpha_n f$.

We apply Theorem 4.1. See Section 4. Note that $\mathcal{D}(\widehat{A}_n)$ is closed under multiplication. The separability condition follows from the separability of $\mathcal{D}(\widehat{A}_n)$ under the norm

$$\|f\|_* = \|f\| + \|\nabla_x f\| + \|\partial_x^2 f\|.$$

The pre-generator condition for $\mathbb{B}$ and $\mathbb{B}_n$ defined in Section 4 follows from existence of solutions of the martingale problem for $\mathbb{B}_n^v f \equiv \mathbb{B}_n f(\cdot, v)$. (See the discussion in Section 2 of Kurtz (1998).) Consequently, taking $C = A$ and $B = \widehat{A}_n$ in Theorem 4.1, any solution $\widetilde{X}$ of the martingale problem for A corresponds to a solution (X, Y, Z^n) of the martingale problem for $\widehat{A}_n$. But note also, that $\beta_n \widehat{A}_{n+1} f = \widehat{A}_n \beta_n f$ for $f \in \mathcal{D}(\widehat{A}_{n+1})$. Consequently, any solution the martingale problem for $\widehat{A}_n$ extends to a solution of the martingale problem for $\widehat{A}_{n+1}$. By induction, we obtain the process (X, Y, Z) so the first part of the theorem follows by Lemma 2.1.

If X is a solution of the martingale problem for A, then by Ethier and Kurtz (1986), Corollary 4.3.7, X has a modification with sample paths in $D_{\mathbb{R}^{d\Delta}}[0,\infty)$. For nonnegative $\kappa \in B(\mathbb{R}^d)$, let

$$\gamma(t) = \inf\{s : \int_0^s \kappa^{-1}(X(r)) dr \geq t\}.$$

Then $\widetilde{X}(t) = X(\gamma(t))$ is a solution of the martingale problem for κA. If $\kappa(x) = 1$ for $|x| \leq k$ and $\kappa(x) = 0$ for $|x| \geq k+1$, then for $\tau_k \equiv \inf\{t : |X(t-)| \text{ or } |X(t)| \geq k\}$, $\widetilde{X}(t) = X(t)$ for $t < \tau_k$ and $\widetilde{X}$ has sample paths in $D_{\mathbb{R}^d}[0,\infty)$. It follows that $\widetilde{X}$ is a weak solution of (2.4) with σ replaced by $\sqrt{\kappa}\sigma$, b replaced by κb and λ replaced by $\kappa\lambda$, and hence X is a weak solution of the original equation (2.4) for $t \in [0, \tau_k)$. Since $\tau_\infty = \lim_{k\to\infty} \tau_k$, the theorem follows. □

Corollary 2.5 *Uniqueness holds for the $D_{\mathbb{R}^d}[0,\infty)$-martingale problem for (A, ν_0) if and only if weak uniqueness holds for (2.4) with initial distribution ν_0.*

3 Conditions for uniqueness.

In Itô (1951) as well as in later presentations (for example, Skorokhod (1965) and Ikeda and Watanabe (1989)), L_2-estimates are used to prove uniqueness for (2.4). Graham (1992) points out the possibility and desirability of using L_1-estimates. (In fact, for equations controlling jump rates with factors like $\mathbf{1}_{[0,\lambda(X(t),u]}(v)$, L_1-estimates are essential.) Kurtz and Protter (1996) develop methods that allow a mixing of L_1, L_2, and other estimates.

Theorem 3.1 *Suppose there exists a constant M such that*

$$|b(x)| + \int_{S_0} |\sigma(x,u)|^2 \mu(du) + \int_{S_1} |\gamma(x,u)|^2 \lambda(x,u)\nu(du) + \int_{S_2} \lambda(x,u)|\gamma(x,u)|\nu(du) < M, \tag{3.1}$$

and

$$\sqrt{\int_{S_0} |\sigma(x,u) - \sigma(y,u)|^2 \mu(du)} \le M|x-y| \tag{3.2}$$

$$|b(x) - b(y)| \le M|x-y| \tag{3.3}$$

$$\int_{S_1} (\gamma(x,u) - \gamma(y,u))^2 \lambda(x,u) \wedge \lambda(y,u)\nu(du) \le M|x-y|^2 \tag{3.4}$$

$$\int_{S_1} |\lambda(x,u) - \lambda(y,u)||\gamma(x,u) - \gamma(y,u)|\nu(du) \le M|x-y| \tag{3.5}$$

$$\int_{S_2} \lambda(x,u)||\gamma(x,u) - \gamma(y,u)|\nu(du) \le M|x-y| \tag{3.6}$$

$$\int_{S} |\lambda(x,u) - \lambda(y,u)||\gamma(y,u)|\nu(du) \le M|x-y|. \tag{3.7}$$

Then there exists a unique solution of (2.4).

Proof. Suppose X and Y are solutions of (2.4). Then

$$\begin{aligned} X(t) &= X(0) + \int_{S_0\times[0,t]} \sigma(X(s),u)W(du\times ds) + \int_0^t b(X(s))ds \\ &\quad + \int_{[0,\infty)\times S_1\times[0,t]} \mathbf{1}_{[0,\lambda(X(s),u)\wedge\lambda(Y(s),u)]}(v)\gamma(X(s-),u)\widetilde{\xi}(dv\times du\times ds) \\ &\quad + \int_{[0,\infty)\times S_1\times[0,t]} \mathbf{1}_{(\lambda(Y(s-),u)\wedge\lambda(X(s-),u),\lambda(X(s-),u)]}(v) \\ &\qquad\qquad \gamma(X(s-),u)\widetilde{\xi}(dv\times du\times ds) \\ &\quad + \int_{[0,\infty)\times S_2\times[0,t]} \mathbf{1}_{[0,\lambda(X(s-),u)]}(v)\gamma(X(s-),u)\xi(dv\times du\times ds), \end{aligned} \tag{3.8}$$

and similarly with the roles of X and Y interchanged. Then (3.2) and (3.3) give the necessary Lipschitz conditions for the coefficient functions in the first two integrals on the right, (3.4) gives an L_2-Lipschitz condition for the third integral term, and (3.5), (3.6), and (3.7) give L_1-Lipschitz conditions for the fourth and fifth integral terms on the right. Theorem 7.1 of Kurtz and Protter (1996) gives uniqueness, and Corollary 7.7 of that paper gives existence. □

Corollary 3.2 *Suppose that there exists a function $M(r)$ defined for $r > 0$, such that (3.1) through (3.7) hold with M replaced by $M(|x| \vee |y|)$. Then there exists a stopping time τ_∞ and a process $X(t)$ defined for $t \in [0, \tau_\infty)$ such that (2.4) is satisfied on $[0, \tau_\infty)$ and $\tau_\infty = \lim_{k\to\infty} \inf\{t : |X(t)| \text{ or } |X(t-)| \geq k\}$. If $(\widetilde{X}, \widetilde{\tau})$ also has this property, then $\widetilde{\tau} = \tau_\infty$ and $\widetilde{X}(t) = X(t)$, $t < \tau_\infty$.*

Proof. The corollary follows by a standard localization argument. □

4 Equations with measurable coefficients.

Let E and F be complete, separable metric spaces, and let $\mathbb{B} \subset \overline{C}(E) \times \overline{C}(E \times F)$. Then Theorem 1.4 can be extended to generators of the form

$$Bf(x) = \int_F \mathbb{B}f(x, v)\eta(x, dv), \tag{4.1}$$

where η is a transition function from E to F, that is, $x \in E \to \eta(x, \cdot) \in \mathcal{P}(F)$ is measurable. Note that $B \subset \overline{C}(E) \times B(E)$ but that B may not have range in $\overline{C}(E)$. (The boundedness assumption can also be relaxed with the addition of moment conditions.) Theorem 1.4 extends to operators of this form.

Theorem 4.1 *Suppose that B given by (4.1) is separable, that for each $v \in F$, $\mathbb{B}^v f \equiv \mathbb{B}f(\cdot, v)$ is a pre-generator, and that $\mathcal{D}(B)$ is closed under multiplication and separates points in E. Let (E_0, r_0) be a complete, separable metric space, $\gamma : E \to E_0$ be Borel measurable, and α be a transition function from E_0 into E ($y \in E_0 \to \alpha(y, \cdot) \in \mathcal{P}(E)$ is Borel measurable) satisfying $\alpha(y, \gamma^{-1}(y)) = 1$. Define*

$$C = \{(\int_E f(z)\alpha(\cdot, dz), \int_E Bf(z)\alpha(\cdot, dz)) : f \in \mathcal{D}(B)\}\ .$$

Let $\mu_0 \in \mathcal{P}(E_0)$, and define $\nu_0 = \int \alpha(y, \cdot)\mu_0(dy)$. If $\widetilde{U}$ is a solution of the martingale problem for (C, μ_0), then there exists a solution V of the martingale problem for (B, ν_0) such that $\widetilde{U}$ has the same distribution on $M_{E_0}[0, \infty)$ as $U = \gamma \circ V$ and

$$P\{V(t) \in \Gamma | \widehat{\mathcal{F}}_t^U\} = \alpha(U(t), \Gamma), \quad \Gamma \in \mathcal{B}(E), t \in \mathbf{T}^U. \tag{4.2}$$

If $\widetilde{U}$ (and hence U) has a modification with sample paths in $D_E[0, \infty)$, then the modified $\widetilde{U}$ and U have the same distribution on $D_E[0, \infty)$.

Proof. See Corollary 3.5, Theorem 2.7, and Theorem 2.9d of Kurtz (1998). □

To apply this result in the proof of Theorem 2.3, we must show that $\widehat{A}_n$ can be written in the form (4.1). Suppose that for each compact $K \subset \mathbb{R}^d$,

$$\sup_{x\in K}(|a(x)| + |b(x)| + \int_{S_1} |\gamma(x, u)|^2\nu(du) + \int_{S_2} |\gamma(x, u)| \wedge 1\nu(du) < \infty.$$

Let F_1 be the space of $d \times d$ nonnegative definite matrices with the usual matrix norm, $F_2 = \mathbb{R}^d$, and F_3 the space of $\mathbb{R}^d$-valued functions on S such that

$$\int_{S_1} |\gamma(u)|^2 \nu(du) + \int_{S_2} |\gamma(u)| \wedge 1\nu(du)) < \infty.$$

We can define a metric on F_3 by

$$d_4(\gamma_1, \gamma_2) = \sqrt{\int_{S_1} |\gamma_1(u) - \gamma_2(u)|^2 \nu(du)} + \int_{S_2} |\gamma_1(u) - \gamma_2(u)| \wedge 1\nu(du).$$

Then $F = F_1 \times F_2 \times F_3$ is a complete, separable metric space, and for $v = (v^1, v^2, v^3) \in F$,

$$\begin{aligned} \mathbb{B}f(x,v) &= \frac{1}{2}\sum_{i,j=1}^{d} v^1_{ij} \frac{\partial^2}{\partial x_i \partial x_j} f(x) + v^2 \cdot \nabla f(x) \\ &\quad + \int_S (f(x + v^3(u)) - f(x) - \mathbf{1}_{S_1}(u) v^3(u) \cdot \nabla f(x)) \nu(du) \end{aligned} \tag{4.3}$$

is the generator of a Levy process in $\mathbb{R}^d$. Let

$$\eta(x, \cdot) = \delta_{(a(x), b(x), \gamma(x, \cdot))}.$$

Then

$$Af(x) = \int \mathbb{B}f(x,v)\eta(x,dv).$$

Similarly, we can define $\mathbb{B}_n$ to include Y and Z^n so that

$$\widehat{A}_n f(x,y,z) = \int \mathbb{B}_n f(x,y,z,v)\eta(x,dv).$$

A Appendix.

Lemma A.1 *Let $A \subset B(E) \times B(E)$, and let X be a cadlag solution of the martingale problem for A. For each $f \in \mathcal{D}(A)$, define*

$$M_f(t) = f(X(t)) - \int_0^t Af(X(s))ds.$$

Suppose $\mathcal{D}(A)$ is an algebra and that $f \circ X$ is cadlag for each $f \in \mathcal{D}(A)$. Let $f_1, \ldots, f_m \in \mathcal{D}(A)$ and $g_1, \ldots, g_m \in B(E)$. Then

$$M(t) = \sum_{i=1}^m \int_0^t g_i(X(s-))dM_{f_i}(s)$$

is a square integrable martingale with Meyer process

$$\langle M \rangle_t = \sum_{1 \le i,j \le m} \int_0^t g_i(X(s))g_j(X(s))(Af_i f_j(X(s)) - f_i(X(s))Af_j(X(s)) - f_j(X(s))Af_i(X(s)))ds.$$

Proof. The lemma follows by standard properties of stochastic integrals and the fact that

$$\langle M_{f_1}, M_{f_2}\rangle_t = \int_0^t (Af_1\mathbf{f_2}(X(s)) - f_1(X(s))Af_2(X(s)) - f_2(X(s))Af_1(X(s)))ds.$$

This identity can be obtained by applying Itô's formula to $f_1(X(t))f_2(X(t))$ and the fact that $[f_1 \circ X, f_2 \circ X]_t = [M_{f_1}, M_{f_2}]_t$ to obtain

$$\begin{aligned}[M_{f_1}, M_{f_2}]_t &= f_1(X(t))f_2(X(t)) - f_1(X(0))f_2(X(0)) - \int_0^t Af_1f_2(X(s))ds \\ &\quad - \int_0^t f_1(X(s-))dM_{f_2}(s) - \int_0^t f_2(X(s-))dM_{f_1}(s) \\ &\quad + \int_0^t (Af_1\mathbf{f_2}(X(s)) - f_1(X(s))Af_2(X(s)) - f_2(X(s))Af_1(X(s)))ds\ .\end{aligned}$$

Since the first five terms on the right give a martingale and the last term is predictable, the last term must be $\langle M_{f_1}, M_{f_2}\rangle_t$. □

References

E. Çinlar, J. Jacod, P. Protter, and M. J. Sharpe. Semimartingales and Markov processes. *Z. Wahrsch. Verw. Gebiete*, 54(2):161–219, 1980. ISSN 0044-3719. 2

Stewart N. Ethier and Thomas G. Kurtz. *Markov processes: Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. ISBN 0-471-08186-8. 1, 1, 2.3

Carl Graham. McKean-Vlasov Itô-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Process. Appl.*, 40(1):69–82, 1992. ISSN 0304-4149. 3

Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1989. ISBN 0-444-87378-3. 3

Kiyosi Itô. On stochastic differential equations. *Mem. Amer. Math. Soc.*, 1951(4):51, 1951. ISSN 0065-9266. 2, 3

Thomas G. Kurtz. Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.*, 3:no. 9, 29 pp. (electronic), 1998. ISSN 1083-6489. 1, 1, 2.3, 4

Thomas G. Kurtz and Giovanna Nappo. The filtered martingale problem. To appear in Handbook, 2009. 1

Thomas G. Kurtz and Philip E. Protter. Weak convergence of stochastic integrals and differential equations. II. Infinite-dimensional case. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 197–285. Springer, Berlin, 1996. 3, 3

A. V. Skorokhod. *Studies in the theory of random processes*. Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965. 3

Daniel W. Stroock. Diffusion processes associated with Lévy generators. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 32(3):209–244, 1975. 2

Daniel W. Stroock and S. R. S. Varadhan. On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory*, pages 333–359, Berkeley, Calif., 1972. Univ. California Press. 1

Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979. ISBN 3-540-90353-4. 1