A new technique for proving uniqueness for martingale problems

Richard F. Bass* and Edwin Perkins†

September 7, 2008

Version:Bismut paper; c.tex

Abstract

A new technique for proving uniqueness of martingale problems is introduced. The method is illustrated in the context of elliptic diffusions in $\mathbb{R}^d$.

1 Introduction

When trying to prove uniqueness of a stochastic process corresponding to an operator, one of the most useful approaches is to consider the associated martingale problem. If $\mathcal{L}$ is an operator and w is a point in the state space $\mathcal{S}$, a probability $\mathbb{P}$ on the set of paths $t \to X_t$ taking values in $\mathcal{S}$ is a solution of the martingale problem for $\mathcal{L}$ started at w if $\mathbb{P}(X_0 = w) = 1$ and $f(X_t) - f(X_0) - \int_0^t \mathcal{L} f(X_s)\, ds$ is a martingale with respect to $\mathbb{P}$ for every f in an appropriate class $\mathcal{C}$ of functions.

The archetypical example is to let

$$\mathcal{L} f(x) = \sum_{i,j=1}^{d} a_{ij}(x) D_{ij} f(x). \tag{1.1}$$

*Research partially supported by NSF grant DMS-0244737.
†Research partially supported by NSERC.

Here, and for the rest of this paper, the state space $\mathcal{S}$ is $\mathbb{R}^d$, the probability measure is on the set of functions that are continuous maps from $[0,\infty)$ into $\mathbb{R}^d$ with the σ-field generated by the cylindrical sets, $D_{ij}f = \partial^2 f/\partial x_i \partial x_j$, and the class $\mathcal{C}$ of functions is the collection C_b^2 of C^2 functions which are bounded and whose first and second partial derivatives are bounded.

Stroock and Varadhan introduced the notion of martingale problem and proved in the case above that there was existence and uniqueness of the solution to the martingale problem provided the a_{ij} were bounded and continuous in x and the matrix $a(x)$ was strictly positive definite for each x. See [2] or [5] for an account of this result.

In this paper we present a new method of proving uniqueness for martingale problems. We illustrate it for the operator $\mathcal{L}$ given in (1.1) under the assumption that the a_{ij} are Hölder continuous in x. Our proof does not give as strong a result as that of Stroock and Varadhan in that we require Hölder continuity. (Actually, we only require a Dini-like condition, but this is still more than just requiring continuity.) In fact, when the a_{ij} are Hölder continuous, an older method using Schauder estimates can be applied.

Nevertheless our technique is applicable to situations for which no other known method seems to work. A precursor of our method, much disguised, was used in [1] to prove uniqueness for pure jump processes which were of variable order, i.e., the operator can not be viewed as a perturbation of a symmetric stable process of any fixed order. The result of [1] was improved in [4] to allow more general jump processes. Moreover our technique is useful in problems arising from certain infinite dimensional situations in the theory of stochastic partial differential equations and the theory of superprocesses; see [3]. Finally, even in the elliptic diffusion case considered here, the proof is elementary and short.

Stroock and Varadhan's method was essentially to view $\mathcal{L}$ given in (1.1) as a perturbation of the Laplacian with respect to the space L^p for appropriate p. The method using Schauder estimates views $\mathcal{L}$ as a perturbation of the Laplacian with respect to the Hölder space C^α for appropriate α. We use a quite different approach. We view $\mathcal{L}$ as a mixture of constant coefficient operators and use a mixture of the corresponding semigroups as an approximation of the semigroup for $\mathcal{L}$.

We use our method to prove the following theorem.

Theorem 1.1 *Suppose $\mathcal{L}$ is given by (1.1), the matrices $a(x)$ are bounded and uniformly positive definite, and there exist c_1 and α such that*

$$|a_{ij}(x) - a_{ij}(y)| \le c_1(1 \wedge |x-y|^\alpha) \tag{1.2}$$

for all $i, j = 1, \ldots, d$ and all $x, y \in \mathbb{R}^d$. Then for each $w \in \mathbb{R}^d$ the solution to the martingale problem for $\mathcal{L}$ started at w is unique.

We do not consider existence, since that is much easier, and we have nothing to add to the existing proofs. The same comment applies to the inclusion of drift terms. In Section 2 we give some easy estimates and in Section 3 we prove Theorem 1.1. The letter c denotes constants whose exact value is unimportant and may change from occurrence to occurrence.

2 Some estimates

All the matrices we consider will be d by d, bounded, symmetric, and uniformly elliptic, that is, there exist constants Λ_m and Λ_M such that

$$\Lambda_m \sum_{i=1}^d z_i^2 \le \sum_{i,j=1}^d a_{ij} z_i z_j \le \Lambda_M \sum_{i=1}^d z_i^2, \qquad (z_1, \ldots, z_d) \in \mathbb{R}^d. \tag{2.1}$$

Given any such matrix a, we use A for a^{-1}. It follows easily that

$$\sup_j \Big(\sum_{i=1}^d a_{ij}^2 \Big)^{1/2} \le \Lambda_M, \quad \sup_j \Big(\sum_{i=1}^d A_{ij}^2 \Big)^{1/2} \le \Lambda_m^{-1} \tag{2.2}$$

Define

$$p^a(t,x,y) = (2\pi t)^{-d/2} (\det a)^{-1/2} e^{-(y-x)^T A (y-x)/(2t)}, \tag{2.3}$$

and let

$$P_t^a f(x) = \int p^a(t,x,y) f(y)\, dy \tag{2.4}$$

be the corresponding transition operator. We assume throughout that the matrix valued function $a(y)$ satisfies the hypotheses of Theorem 1.1 and (2.1). Note that for a fixed, $p^a(t,x,y)\, dy$ is a Gaussian distribution for each

x, but that $p^{a(y)}(t,x,y)\,dy$ need not be a probability measure. All numbered constants will depend only Λ_m, Λ_M and d.

We have the following.

Proposition 2.1 *There exist c_1, c_2 and a function $c_3(p), p > 0$, depending only on Λ_M and Λ_m, such that for all $t, N, p > 0$ and $x \in \mathbb{R}^d$,*

(a) $\int p^{a(y)}(t,x,y)\,dy \le c_1$.

(b)

$$\int_{|y-x|>N/\sqrt{t}} p^{a(y)}(t,x,y)\,dy \le c_1 e^{-c_2 N^2}.$$

(c) For each $i \le d$,

$$\int \Big(\frac{|x_i - y_i|^2}{t}\Big)^p p^{a(y)}(t,x,y)\,dy \le c_3(p).$$

Proof. For (a), after a change of variables $z = (y-x)/\sqrt{t}$, we need to bound

$$\begin{aligned}\int (2\pi)^{-d/2}&(\det a(x+z\sqrt{t}))^{-1/2} e^{-z^T A(x+z\sqrt{t})z/2}\,dz\\ &\le \Big(\frac{\Lambda_M}{\Lambda_m}\Big)^{d/2} \int (2\pi\Lambda_M)^{-d/2} e^{-z^T z/2\Lambda_M}\,dz \le \Big(\frac{\Lambda_M}{\Lambda_m}\Big)^{d/2}.\end{aligned}$$

(b) and (c) are similar. □

Let $\|f\|$ be the C_0 norm of f.

Proposition 2.2 *Let $g \in C^2$ with compact support and let*

$$F_\varepsilon(x) = \int g(y) p^{a(y)}(\varepsilon^2, x, y)\,dy.$$

Then $F_\varepsilon(x)$ converges to $g(x)$ boundedly and pointwise as $\varepsilon \to 0$.

Proof. Because g is bounded, using Proposition 2.1(a) we see that the quantity $\sup_{\varepsilon>0} \|F_\varepsilon\|$ is finite. We next consider pointwise convergence. After a change of variables, we have

$$F_\varepsilon(x) = \int g(x+\varepsilon z)(2\pi)^{-d/2}(\det a(x+\varepsilon z))^{-1/2} e^{-z^T A(x+\varepsilon z)z/2}\,dz.$$

Since $|g(x+\varepsilon z)-g(x)|\le \varepsilon|z|\,\|\nabla g\|$, F_ε differs from

$$g(x)\int (2\pi)^{-d/2}(\det(a(x+\varepsilon z))^{-1/2}e^{-z^TA(x+\varepsilon z)z/2}\,dz$$

by at most

$$\|\nabla g\|\int (2\pi)^{-d/2}(\det(a(x+\varepsilon z)))^{-1/2}\varepsilon|z|e^{-z^TA(x+\varepsilon z)z/2}\,dz,$$

and this goes to 0 as $\varepsilon\to 0$ by a change of variables and Proposition 2.1(c) with $p=1/2$. Let

$$V(\varepsilon,x,z)=(2\pi)^{-d/2}(\det(a(x+\varepsilon z)))^{-1/2}e^{-z^TA(x+\varepsilon z)z/2}.$$

It therefore suffices to show

$$\int V(\varepsilon,x,z)\,dz\to\int V(0,x,z)\,dz,$$

where we note this right-hand side is 1. Using Proposition 2.1(b) and the same change of variables, it suffices to show

$$\int_{|z|\le N}V(\varepsilon,x,z)\,dz\to\int_{|z|\le N}V(0,x,z)\,dz.$$

But this last follows by dominated convergence. □

Proposition 2.3 *There exists a constant c_4 such that*

$$\int |a_{ij}(y)-a_{ij}(x)|\,|D_{ij}p^{a(y)}(t,x,y)|\,dy\le\begin{cases}c_4t^{\frac{\alpha}{2}-1}, & t\le 1,\\ c_4t^{-1}, & t\ge 1.\end{cases}$$

Proof. A computation shows that

$$\begin{aligned}&D_{ij}p^{a(y)}(t,x,y)\\&\quad=t^{-1}p^{a(y)}(t,x,y)\Big[\sum_k\sum_l\frac{(y_k-x_k)A_{ki}(y)A_{lj}(y)(y_l-x_l)}{t}-A_{ij}(y)\Big].\end{aligned}\tag{2.5}$$

By (2.2) and Cauchy-Schwarz we have

$$\int |a_{ij}(y) - a_{ij}(x)|\, |D_{ij} p^{a(y)}(t,x,y)|\, dy$$
$$\leq \Big[\int |a_{ij}(y) - a_{ij}(x)| t^{-1} p^{a(y)}(t,x,y)[|x-y|^2 t^{-1}\Lambda_m^{-2} + \Lambda_m^{-1}]\, dy. \qquad (2.6)$$

Suppose first that $t \leq 1$. By the Hölder condition on a the above is at most

$$c\int \frac{|y-x|^\alpha}{t^{\alpha/2}}\Big[\frac{|x-y|^2}{t} + 1\Big] p^{a(y)}(t,x,y)\, dy\, t^{\alpha/2-1}$$
$$\leq ct^{\alpha/2-1},$$

where we have used Proposition 2.1(c) in the last inequality.

For the case $t > 1$ simply use the boundedness of a in (2.6) and Proposition 2.1 again to bound it by ct^{-1}.

3 Proof of Theorem 1.1

For $f \in C_b^2$ and a a matrix with constant coefficients define

$$\mathcal{M}^a f(x) = \sum_{i,j=1}^d a_{ij} D_{ij} f(x).$$

Define the corresponding semigroup by (2.4), and let $R_\lambda^a f = \int_0^\infty e^{-\lambda t} P_t^a f\, dt$. For $f \in C_b^2$ we have

$$\mathcal{L} f(x) = \mathcal{M}^{a(x)} f(x).$$

Note that

$$(\lambda - \mathcal{M}^{a(y)}) R_\lambda^{a(y)} P_\varepsilon^{a(y)} f(x) = P_\varepsilon^{a(y)} f(x). \qquad (3.1)$$

One way to verify that the superscript $a(y)$ does not cause any difficulty here is to check that

$$\sum_{i,j=1}^d a_{ij}(y) \frac{\partial^2}{\partial x_i \partial x_j} p^{a(y)}(s,x,y) = \frac{\partial}{\partial s} p^{a(y)}(s,x,y),$$

and then in the definition of $R_\lambda^{a(y)}$ use integration by parts in the time variable. By replacing ε with $\varepsilon/2$, setting $f(z) = p^{a(y)}(\varepsilon/2, z, y)$ and using Chapman-Kolmogorov, we see that (3.1) implies

$$(\lambda - \mathcal{M}^{a(y)})(R_\lambda^{a(y)} p^{a(y)}(\varepsilon, \cdot, y))(x) = p^{a(y)}(\varepsilon, x, y). \tag{3.2}$$

We are now ready to prove Theorem 1.1.

Proof. Suppose $\mathbb{P}_1, \mathbb{P}_2$ are two solutions to the martingale problem for $\mathcal{L}$ started at a point w. Define

$$S_\lambda^i f = \mathbb{E}_i \int_0^\infty e^{-\lambda t} f(X_t)\, dt, \qquad i = 1, 2,$$

and

$$S_\lambda^\Delta f = S_\lambda^1 f - S_\lambda^2.$$

We make two observations. First, because $\mathbb{P}_i$ need not come from a Markov process, $S_\lambda^i f$ is not a function, and so S_λ^Δ is a linear functional. Second, if

$$\Theta = \sup_{\|f\| \le 1} |S_\lambda^\Delta f|,$$

then $\Theta < \infty$.

If $f \in C_b^2$, then by the definition of the martingale problem

$$\mathbb{E}_i f(X_t) - f(w) = \mathbb{E}_i \int_0^t \mathcal{L} f(X_s)\, ds, \qquad i = 1, 2.$$

Multiply both sides by $\lambda e^{-\lambda t}$, integrate over t from 0 to ∞, and use Fubini to obtain

$$f(w) = S_\lambda^i(\lambda f - \mathcal{L} f), \qquad i = 1, 2,$$

or

$$S_\lambda^\Delta(\lambda f - \mathcal{L} f) = 0. \tag{3.3}$$

Let $g \in C^2$ with compact support and set

$$f_\varepsilon(x) = \int R_\lambda^{a(y)}(p^{a(y)}(\varepsilon, \cdot, y))(x) g(y)\, dy.$$

Since this is the same as

$$e^{-\lambda\varepsilon} \int \int_{\varepsilon}^{\infty} e^{-\lambda t} p^{a(y)}(t,x,y)\,dt\,g(y)\,dy,$$

we see that f_ε is in C_b^2 in x by dominated convergence.

To calculate $(\lambda - \mathcal{L})f_\varepsilon$ it is easy to differentiate under the dy integral and so we may write

$$\begin{aligned}
(\lambda - \mathcal{L})f_\varepsilon(x) &= (\lambda - \mathcal{M}^{a(x)})f_\varepsilon(x) \\
&= \int (\lambda - \mathcal{M}^{a(y)})R_\lambda^{a(y)}(p^{a(y)}(\varepsilon,\cdot,y))(x)g(y)\,dy \\
&\quad + \int (\mathcal{M}^{a(y)} - \mathcal{M}^{a(x)})R_\lambda^{a(y)}(p^{a(y)}(\varepsilon,\cdot,y))(x)g(y)\,dy \\
&:= I_\varepsilon(x) + J_\varepsilon(x).
\end{aligned}$$

By Proposition 2.3,

$$\begin{aligned}
|J_\varepsilon(x)| &\le \sum_{i,j=1}^{d} \int_0^\infty e^{-\lambda t} \int |a_{ij}(y) - a_{ij}(x)| \\
&\qquad\qquad \times |D_{ij}p^{a(y)}(\varepsilon + t, x, y)|\,|g(y)|\,dy\,dt \\
&\le d^2\|g\| \int_0^\infty e^{-\lambda t} c_4 t^{-1}(t^{\alpha/2} \wedge 1)\,dt \\
&\le \frac{1}{2}\|g\|,
\end{aligned}$$

for $\lambda \ge \lambda_0(\alpha, d, c_4)$. By (3.2), $I_\varepsilon(x) = \int p^{a(y)}(\varepsilon, x, y)g(y)\,dy$, and so by Proposition 2.2, $I_\varepsilon(x)$ converges to g boundedly and pointwise. Since $S_\lambda^\Delta(\lambda - \mathcal{L})f_\varepsilon = 0$ by (3.3), we have $|S_\lambda^\Delta I_\varepsilon| = |S_\lambda^\Delta J_\varepsilon|$. Letting $\varepsilon \to 0$,

$$|S_\lambda^\Delta g| = \lim_{\varepsilon\to 0} |S_\lambda^\Delta I_\varepsilon| = \lim_{\varepsilon\to 0} |S_\lambda^\Delta J_\varepsilon| \le \Theta \limsup_{\varepsilon\to 0} \|J_\varepsilon\| \le \tfrac{1}{2}\Theta\|g\|.$$

Using a monotone class argument, the above inequality holds for all bounded g, and then taking the supremum over g such that $\|g\| \le 1$, we have $\Theta \le \frac{1}{2}\Theta$. Since $\Theta < \infty$, this implies that $\Theta = 0$.

From this point on, we use standard arguments. By the uniqueness of the Laplace transform together with continuity in t, $\mathbb{E}_1 f(X_t) = \mathbb{E}_2 f(X_t)$ for all

t if f is continuous and bounded. Using regular conditional probabilities, one shows as usual that the finite dimensional distributions under $\mathbb{P}_1$ and $\mathbb{P}_2$ agree. This suffices to prove uniqueness; see [2] or [5] for details. $\square$

Note that no localization argument is needed in the above proof.

References

[1] R.F. Bass. Uniqueness in law for pure jump processes, *Probab. Th. rel. Fields* **79** (1988) 271–287.

[2] R.F. Bass. *Diffusions and Elliptic Operators.* Springer, New York, 1997.

[3] R.F. Bass and E. Perkins. Uniqueness for stochastic partial differential equations with Hölder continuous coefficients. In preparation.

[4] H. Tang, *Uniqueness for the martingale problem associated with pure jump processes of variable order*, Ph.D. dissertation, University of Connecticut, 2006.

[5] D.W. Stroock and S.R.S. Varadhan, *Multidimensional Diffusion Processes*, Springer, New York, 1979.

Richard F. Bass
Department of Mathematics
University of Connecticut
Storrs, CT 06269-3009, USA
bass@math.uconn.edu

Edwin Perkins
Department of Mathematics
University of British Columbia
Vancouver, B.C. V6T 1Z2, Canada
perkins@math.ubc.ca