Uniqueness in law of parabolic SPDE's and infinite-dimensional SDE's

Richard Bass
University of Connecticut
r.bass@uconn.edu
www.math.uconn.edu/∼bass

This is joint work with Ed Perkins of UBC.

(Please direct any questions about SPDE's to him!)

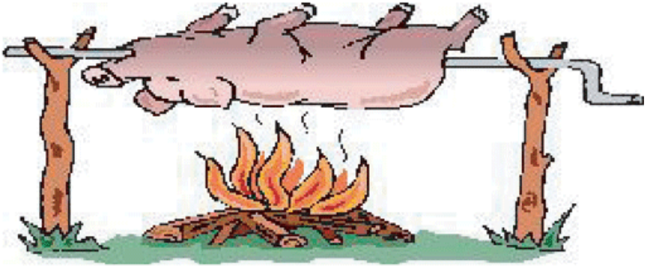
The SPDE

Consider the SPDE

$$\frac{\partial u}{\partial t}(x,t) = \tfrac{1}{2}\frac{\partial^2 u}{\partial x^2}(x,t) + A(u(\cdot,t))(x)\,\dot{W}_{x,t}.$$

We are on the interval $[0,1]$ with reflection on the boundaries. A is an operator mapping $C[0,1]$ into $C[0,1]$, and $\dot{W}_{x,t}$ is a space-time white noise.

Notice this is the heat equation with insulated ends and a random forcing term.



The question to be answered: give reasonable conditions on A, presumably of Hölder type, so that uniqueness in law holds for the SPDE.

Some examples of $A(u)$

We will put conditions on A later on, but here are some possibilities. Not all of these can be handled by our theorem.

A could be localized:

$$A(u)(x) = f(u(x)), \qquad \text{e.g.,} \qquad A(u)(x) = |u(x)|^\alpha.$$

Or A could be non-localized, as in the following:

$$A(u)(x) = f\Big(x, \int_0^1 g(u(y))\, dy\Big),$$

as for example,

$$A(u)(x) = 2 + \sin x + \Big| \int_0^1 u(y)\, dy \Big|^\alpha.$$

The examples we would most like to handle are of the form

$$A(u)(x) = B(u)(x)|u(x)|^{1/2},$$

where B is smooth and bounded above and below by positive constants.

The solution to the SPDE in this case is the density of a superprocess with branching interaction, and uniqueness for the SPDE would imply uniqueness for such superprocesses.

Unfortunately, our conditions on A will not allow the $|u|^{1/2}$ factor.

When does uniqueness hold?

If A is Lipschitz:

$$\|A(u) - A(v)\| \leq c\|u - v\|,$$

then Picard iteration proves uniqueness.

Certain special cases, e.g., $A(u)(x) = |u(x)|^p$ for $1/2 \leq p < 1$, have been proved by Mytnik using the method of duality. This method is not robust, however.

There hasn't been any general theorem that gives uniqueness for a reasonably large class of A's using only a Hölder continuity condition.

Space-time white noise

The theory is due to Walsh. If φ is a reasonable function, then

$$\int_0^t \int \varphi(x,s)\, dW_{x,s}$$

is a martingale with quadratic variation $\int_0^t \int \varphi(x,s)^2\, dx\, ds$.

Solutions

Saying u is a solution to the SPDE means the following. Let $u_t(x) = u(x,t)$. Let $\langle f, g \rangle$ be the usual inner product in $L^2[0,1]$.

Then for all $\varphi \in C^2[0,1]$,

$$\langle u_t, \varphi \rangle = \langle u_0, \varphi \rangle + \int_0^t \langle u_s, \varphi''/2 \rangle \, ds + \int_0^t \int \varphi(x) A(u_s)(x) \, dW_{x,s}.$$

The localized case

We'll impose some conditions on A. It turns out our conditions preclude examples like

$$A(u)(x) = f(u(x)).$$

Mytnik and Perkins have shown pathwise uniqueness for the SPDE if $A(u)(x) = f(u(x))$ is of this form and f is Hölder continuous of order greater than $3/4$. Together with Mueller they showed $3/4$ is sharp.

Conditions on $A(u)$

There are two types of smoothness involved with A. One is: how small must $A(u) - A(v)$ be if $u - v$ is small? (In most examples, A is not a linear operator.) The other is, given u, $A(u)$ is a function of x. How smooth must this function be?

The conditions we impose on A are the following:

(1) $A(u)$ is bounded above and below by positive constants independent of u.
(2) A maps $C[0,1]$ into a bounded subset of C^γ for some γ sufficiently large.
(3) A is Hölder continuous in u of order α bigger than $1/2$ with respect to a certain norm of Wasserstein type.

(1) is clear, although it precludes examples like $A(u)(x) = |u(x)|^{1/2}$.

On (2), one can ask how large must γ be (for our proof it has to be bigger than 1 and must get larger as α gets closer to $1/2$).

Also, there is the question of precisely what reflection means for C^γ functions when $\gamma \geq 1$.

We require the odd derivatives of $A(u)$ of order up to γ must be 0 at 0 and 1. This allows us to reflect $A(u)$ across the origin to be an even function and still be in C^γ, and then allows us to extend $A(u)$ to be a function on all of $\mathbb{R}$ of period 2.

The most questions arise from (3). What we require is that

$$\|A(u) - A(v)\|_2 \leq c \sup_{\|\varphi\|_{C^\beta} \leq 1} |\langle u - v, \varphi \rangle|^\alpha.$$

This is a Wasserstein type norm, and is one of the classical ones if $\beta = 1$.

The fact that we take the supremum over $\varphi \in C^\beta$ eliminates consideration of localized A. (If we took the supremum over $\varphi \in L^2$, we would get a Hölder condition on A that would be more reasonable.)

We need

$$\alpha \in (\tfrac{1}{2}, 1], \qquad \gamma > \frac{2\alpha}{2\alpha - 1},$$

and

$$\beta > \Big(\frac{9}{2\alpha} - 1\Big) \vee \Big(\frac{\gamma}{\alpha(2-\gamma)}\Big).$$

Examples

Let $\varphi_1, \dots, \varphi_n$ be smooth functions, let $f(x, y_1, \dots, y_n)$ be C^γ in x and Hölder continuous of order α in $y_1, \dots, y_n$. Then take

$$A(u)(x) = f(x, \langle u, \varphi_1 \rangle, \dots, \langle u, \varphi_n \rangle).$$

Our previous example

$$A(u)(x) = 2 + \sin x + \Big| \int_0^1 u(y)\, dy \Big|^\alpha$$

fits into this framework.

For another class of examples, let $\varphi_1, \ldots, \varphi_n$ be smooth, ψ be smooth, and $f(y_1, \ldots, y_n)$ Hölder continuous of order α in each variable. Let

$$A(u)(x) = \psi * f((\varphi_1 * u), \ldots, (\varphi_n * u))(x).$$

One might give the following interpretation of this. One can't measure the heat at a point exactly, only an average near the point. This is the $\varphi * u$. One applies the Hölder function f to a finite number of measurements. One wants to modulate the random noise by $f(\cdots)$, but the controls are not that sensitive and one can only get a smoothed version of the control one wants.

Connection with SDE's

Recall that saying u satisfied the SPDE means that

$$\langle u_t, \varphi \rangle = \langle u_0, \varphi \rangle + \int_0^t \langle u_t, \varphi''/2 \rangle + \text{ martingale } .$$

If we take $\varphi(x) = e_n(x) = e^{2\pi i n x}$ and let $X^n(t) = \langle u_t, e_n \rangle$ and then sort out the quadratic variations, we see that we get the equations

$$dX_t^j = \sum_k \sigma_{jk}(X_t)\, dW_t^k - \lambda_j X_t^j\, dt,$$

where the W^k are independent Brownian motions, λ_j are the eigenvalues for the operator $(1/2)f''$, hence are equal to cj^2, and the σ turn out to depend on the operator A in a very specific way: $\sigma(x) = \sqrt{a(x)}$, where

$$a_{jk}(x) = \int_0^1 A(u(x))^2(y) e_j(y) \overline{e}_k(y)\, dy.$$

We identify u with its Fourier coefficients $\{x^n\}$, $n \in \mathbb{Z}$. One approach to proving uniqueness in law for the SPDE is to prove uniqueness in law for the system of SDEs. To do that we look at the corresponding martingale problem:

Let

$$\mathcal{L}f(x) = \frac{1}{2}\sum_{j,k} a_{jk}(x) D_{jk} f(x) - \sum_j \lambda_j x_j D_j f(x)$$

for functions that depend only on finitely many coordinates and are C_b^2 in those coordinates. A probability measure $\mathbb{P}$ solves the martingale problem started at a point $v \in \ell^2$ if $\mathbb{P}(X_0 = v) = 1$ and

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\, ds$$

is a $\mathbb{P}$-martingale for all such f.

Related work

This last is related to situations that have been much studied by Röckner and his co-authors and students and by Da Prato and his co-authors and students. When the a_{jk} are only Hölder continuous, the relevant papers are by Cannarsa and Da Prato (1996), by Zambotti (2000), and by Athreya, Bass, Gordina, and Perkins (2006). These papers develop Schauder estimates and then apply a standard perturbation argument to get uniqueness.

None of these papers are of any use here. They are all require the (infinite dimensional) matrix a to be equal to a constant matrix plus something relatively small, specifically, a perturbation by a trace class operator.

In our context,

$$a_{jk}(x) = \int_0^1 A(u)(x))^2(y)e_j(y)\overline{e}_k(y)\,dy$$
$$= \int_0^1 A(u)(x))^2(y)e^{2\pi i(j-k)y}\,dy.$$

Thus the a's are in Toeplitz form, i.e., a_{jk} depends only on $j-k$, or phrased another way, the a matrix is constant (for each fixed x) along the diagonals. Even if a were a diagonal matrix, there is no way it could be constant (in x) plus trace class, unless it were constant.

On the other hand, we need some way to handle the infinite dimensionality. The $-\lambda X^j \, dt$ guarantees that the process doesn't explode instantaneously, but we still need a condition that says that higher dimensions are more nicely behaved.

Conditions on the a_{jk}

The conditions are what allows the proof to work, and when they are translated to conditions on the $A(u)$, they become the conditions of the theorem on SPDEs. On the a_{jk} we require Toeplitz form, symmetric,

(1) boundedness and positive definite:

$$\Lambda_0 |w|^2 \le \langle a(x)w, w \rangle \le \Lambda_1 |w|^2$$

for $w \in \ell^2$,
(2) off-diagonal decay:

$$|a_{ij}(x)| \le \frac{c}{1 + |i-j|^\gamma},$$

(3) Hölder continuity that gets better in higher dimensions:

$$|a_{ij}(y + he_k) - a_{ij}(y)| \le c|h|^\alpha k^{-\beta'}.$$

We impose more or less the same constraints on α, γ, and β, and set $\beta' = \alpha\beta$.

Under these conditions there is a unique solution to the martingale problem (existence is known), and hence uniqueness in law to the infinite-dimensional SDE.

Ideas of the proof

Here are some ideas on the proof. First of all, we let $X_t^n = \langle u(\cdot, t), e_n \rangle$, but instead of a Fourier series, we use a Fourier cosine series.

This avoids our X's being complex-valued. Our boundary conditions were set up so that the collection of Fourier cosine series is dense in $L^2[0, 1]$.

As a result, our

$$a_{ij}(x) = \int_0^1 A(u)(y)^2 e_i(y) e_j(y)\, dy$$

are no longer of Toeplitz form. But they are equal to something of Toeplitz form plus something small, so that's OK.

We couldn't get standard perturbation techniques to work, so we used a technique I used in the late 1980's for jump processes and that Ed and I refined in a recent paper. Standard perturbation means you start with an operator you understand well, e.g., the Laplacian, and look at $\Delta + B$, where B is smaller than Δ in some sense. More specifically, look at BR_θ, where $R_\theta = (\theta - \Delta)^{-1}$. In the technique that Ed and I use, we perturb off a mixture of operators.

Let me give a few of the ideas. Recall

$$\mathcal{L}f(x) = \tfrac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) - \sum_i \lambda_i x_i \frac{\partial f}{\partial x_i}.$$

Define for each z

$$\mathcal{M}^z f(x) = \tfrac{1}{2}\sum_{i,j} a_{ij}(z)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) - \sum_i \lambda_i x_i \frac{\partial f}{\partial x_i}.$$

Then

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

Let p_t^z and r_θ^z be the transition densities and resolvent densities corresponding to $\mathcal{M}^z$.

Let

$$S_i f = \mathbb{E}_i \int_0^\infty e^{-\theta t} f(X_t)\, dt,$$

where $\mathbb{P}_1$ and $\mathbb{P}_2$ are two solutions to the martingale problem. Some routine calculations show

$$S_i(\theta - \mathcal{L})f = f,$$

and so

$$S_\Delta(\theta - \mathcal{L})f = 0,$$

where

$$S_\Delta = S_1 - S_2.$$

Given g bounded and smooth, look at

$$f(x) = \int g(y) r_\theta^y(x,y)\, dy.$$

For comparison, the resolvent of a Markov process is

$$R_\theta g(x) = \int g(y) r_\theta(x,y)\, dy.$$

(I switched to finite dimensions so that we can use Lebesgue measure. This is only for the purposes of explanation.)

In general f is not in the domain of $\mathcal{L}$, but we can approximate and this is not a major problem. So pretending it is in the domain, we write

$$\begin{aligned}(\theta - \mathcal{L})f(x) &= \int (\theta - \mathcal{M}^y) r_\theta^y(x,y) g(y)\, dy \\ &\qquad + \int (\mathcal{M}^y - \mathcal{M}^x) r_\theta^y(x,y) g(y)\, dy \\ &= g(y) + \int (\mathcal{M}^y - \mathcal{M}^x) r_\theta^y(x,y) g(y)\, dy.\end{aligned}$$

If we can show

$$\Big|\int (\mathcal{M}^y - \mathcal{M}^x) r_\theta^y(x,y) g(y)\,dy\Big| \le \tfrac{1}{2}\|g\|_\infty$$

when θ is large we would then get

$$|S_\Delta g| \le \tfrac{1}{2}\|S_\Delta\|\,\|g\|_\infty.$$

This would imply the norm of the linear functional S_Δ is zero, so $S_1 = S_2$. Inverting the Laplace transform, this shows equality of the one-dimensional distributions under $\mathbb{P}_i$, and from there it is standard to get uniqueness.

We can get the bound we need from a bound on

$$\int \left| (\mathcal{M}^y - \mathcal{M}^x) r_\theta^y(x,y) \right| dy,$$

and that would come from a suitable bound on

$$\int \left| (\mathcal{M}^y - \mathcal{M}^x) p_t^y(x,y) \right| dy.$$

It turns out we only need to worry about t small.

Why $\alpha > 1/2$?

We calculate

$$D_{ij} p_t^y(x,y),$$

where the derivatives are in the x variable. We get

$$e^{-(\lambda_i+\lambda_j)t} \frac{1}{t} \Big[\frac{(A(y)(y-x))_i}{t} \frac{(A(y)(y-x))_j}{t} - A_{ij}(y) \Big] p_t^y(x,y)\, dy,$$

where $A = a^{-1}$. (Not quite, because we have Ornstein-Uhlenbeck semigroups, not Brownian ones.)

Because of the factor

$$e^{-(\lambda_i+\lambda_j)t}$$

that appears, given t and remembering $\lambda_j \approx cj^2$, we only have to sum over 1 to $J \approx t^{-1/2}$ instead of from 1 to infinity.

We get a factor t^{-1} when computing $D_{ij} p_t(x, y)$. From the terms

$$|a_{ij}(y) - a_{ij}(x)| \approx |y - x|^{\alpha}$$

we get a factor $t^{\alpha/2}$.

If we consider only the main diagonal, we sum over $J \approx t^{-1/2}$ terms, but they behave somewhat like sums of independent random variables, so the contribution is $\sqrt{J} \approx t^{-1/4}$. So altogether we have

$$t^{-1} t^{\alpha/2} t^{-1/4},$$

which is integrable near 0 provided $\alpha > 1/2$.

The Toeplitz form of a comes in because it allows us to write

$$\int |a_{ii}(y) - a_{ii}(x)| \Big| \sum_i D_{ii} p_t^y(x, y) \Big| \, dy.$$

Finiteness

When working with all this, there are good reasons for working in finite dimensions (dimension K, say), and getting estimates independent of K. But it is not even clear that

$$\int_{\mathbb{R}^K} p_t^y(x,y)\,dy$$

can be bounded independently of K, let alone its derivatives. The expression we are looking at is essentially

$$\int_{\mathbb{R}^K} (\det a(y))^{-1/2} e^{-\langle y-x, a^{-1}(y)(y-x)\rangle/2t}\,dy.$$

(Again not quite, because we have Ornstein-Uhlenbeck semigroups here, not Brownian ones.)

The way we approach this is the following. Let a^m be the matrix consisting of the first m columns and rows of a. Write A^K for the inverse of a^K. Let $\pi_{K-1}(y) = (y_1, \ldots, y_{K-1}, 0)$. Then our assumptions imply that $a^K(y)$ is not that far from $a^K(\pi_{K-1}(y))$. Some linear algebra shows that this implies $A^K(y)$ is not that far from $A^K(\pi_{K-1}(y))$ and $\det a^K(y)$ is not that far from $\det a^K(\pi_{K-1}(y))$.

The first fact follows from the identity

$$B^{-1} - C^{-1} = B^{-1}(C - B)C^{-1}.$$

The second follows from

$$\frac{\det B}{\det C} = \det(BC^{-1}) = \det(I + (B - C)C^{-1})$$

and the fact that the spectral radius is bounded by the operator norm.

If we look at

$$\int_{\mathbb{R}^K} (\det a(\pi_{K-1}(y)))^{-1/2} e^{-\langle y-x, A^K(\pi_{K-1}(y))(y-x)\rangle/2t}\, dy,$$

this does not depend on y_K except in the $y-x$ terms. We can integrate out the y_K variable, and some calculus and linear algebra shows we end up with

$$\int_{\mathbb{R}^{K-1}} (\det a^{K-1}(y))^{-1/2} e^{-\langle y-x, A^{K-1}(y)(y-x)\rangle/2t}\, dy.$$

We keep track of the errors, and use induction.

To handle

$$\int \Big| \sum_{i=1}^{J} D_{ii} p_t^y(x,y) \Big| \, dy,$$

we use Cauchy-Schwarz and multiply out the square. We need to estimate

$$\int_{i,j=1}^{J} D_{ii} p_t^y(x,y) D_{jj} p_t^y(x,y) \, dy.$$

This needs to be bounded independently of K. We do this similarly to bounding $\int p_t^y(x,y)$, although the calculations are more complicated.

Off-diagonals

To handle the other diagonals, we use the fact that the a_{ij} decay as $|i-j|$ gets large. So we have similar calculations, but the terms are summable in $|i-j|$.

Recall that when one calculates

$$D_{ij}p_t^y(x,y),$$

one gets an expression similar to

$$\frac{1}{t}\Big[\frac{(A(y)(y-x))_i}{t}\frac{(A(y)(y-x))_j}{t}-A_{ij}(y)\Big]p_t^y(x,y)\,dy.$$

The problem here is that one needs off-diagonal decay in A_{ij} and we only have it for a_{ij}.

Jaffard's theorem

There is a remarkable theorem of Jaffard that saves the day. Suppose a is a $m \times m$ matrix that is bounded and positive definite and

$$|a_{ij}| \le \frac{c_1}{1 + |i - j|^\gamma}$$

for all i, j, where $\gamma > 1$. Write A for a^{-1}. Then

$$|A_{ij}| \le \frac{c_2}{1 + |i - j|^\gamma},$$

where c_2 depends only on c_1 and γ and the boundedness and positive definiteness bounds of a, but NOT m.

It is clear for diagonal matrices!

The idea behind Jaffard's theorem is to write $a = \|a\|(I - b)$, where $\|b\| < 1$. One thus needs to look at

$$(I - b)^{-1} = I + b + b^2 + \cdots .$$