STOCHASTIC INTEGRATION AND STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS: A TUTORIAL

A VIGRE MINICOURSE ON STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS HELD BY THE DEPARTMENT OF MATHEMATICS THE UNIVERSITY OF UTAH MAY 8–19, 2006

DAVAR KHOSHNEVISAN

Abstract. These are supplementary notes for three introductory lectures on SPDEs that were held as part of a VIGRE minicourse on SPDEs at the Department of Mathematics at the University of Utah (May 8–19, 2006). The notes are rough, and not meant for publication. So read them at your own risk.

1991 *Mathematics Subject Classification.* Primary: 60H15; 60H20. Secondary: 35R60; 60G60.
Key words and phrases. Stochastic partial differential equations.
†Author's work is supported in part by a grant from the National Science Foundation.
‡This minicourse is made possible by financial support from the Department of Mathematics at the University of Utah and a VIGRE grant from the National Science Foundation.

Contents

1. What is an SPDE? 2
2. Gaussian Random Vectors 2
3. Gaussian Processes 3
4. Regularity of Processes 7
4.1. A Diversion 8
4.2. Kolmogorov's Continuity Theorem 8
5. Martingale Measures 10
5.1. A White Noise Example 10
5.2. More General Martingale Measures 12
6. The Semilinear Heat Equation 19
References 26

1. What is an SPDE?

Consider a perfectly even, infinitesimally-thin wire of length L. Lay it down flat, so that we can think of it as $[0\,,L]$. Now we apply pressure to the string in order to make it vibrate. Let $F(t\,,x)$ denote the amount of pressure per unit length applied in the direction of the y-axis at place $x \in [0\,,L]$. [$F < 0$ means press down toward $y = -\infty$, and $F > 0$ means the opposite.] Then the position $u(t\,,x)$ of the wire solves the partial differential equation,

$$\frac{\partial^2 u(t\,,x)}{\partial t^2} = \kappa \frac{\partial^2 u(t\,,x)}{\partial x^2} + F(t\,,x), \qquad (t \geq 0,\ 0 \leq x \leq L). \tag{1.1}$$

κ is a physical constant that depends on the linear mass density and the tension of the wire. Equation (1.1) is the so-called *one-dimensional wave equation* whose physical derivation goes back to De l'Ambert (Laplace's mentor). Its solution—via separation of variables and superposition—is likewise old. What if F is "random noise"? Here is an amusing interpretation (Walsh, 1986): If a guitar string is bombarded by particles of sand then its vibrations are determined by a suitable version of (1.1). It turns out that for most interesting random noises F, (1.1) no longer has a classical meaning. But it can be interpreted as an infinite-dimensional integral equation. These notes are a way to get you started thinking in this direction. They are based on the Saint-Flour lecture notes of John B. Walsh (1986, Chapters 1–3), which remains as one of the best introductions to this subject to date.

2. Gaussian Random Vectors

Let $\boldsymbol{g} := (g_1\,, \ldots\,, g_n)$ be an n-dimensional random vector. We say that the distribution of $\boldsymbol{g}$ is *Gaussian* if $\boldsymbol{t}'\boldsymbol{g} := \sum_{j=1}^n t_j g_j$ is a Gaussian random variable for all $\boldsymbol{t} := (t_1, \ldots, t_n) \in \mathbf{R}^n$.

That is, there must exist $\boldsymbol{\mu} \in \mathbf{R}^n$ and an $n \times n$, symmetric, nonnegative-definite matrix $\boldsymbol{C}$ such that

$$\mathrm{E}\left[\exp\left(i\boldsymbol{t}'\boldsymbol{g}\right)\right] = \exp\left(i\boldsymbol{t}\cdot\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}'\boldsymbol{C}\boldsymbol{t}\right). \tag{2.1}$$

Exercise 2.1. Prove this assertion. [Recall that $\boldsymbol{C}$ is *nonnegative definite* if and only if $\boldsymbol{t}'\boldsymbol{C}\boldsymbol{t} \geq 0$ for all $\boldsymbol{t} \in \mathbf{R}^n$. Equivalently, if all eigenvalues of $\boldsymbol{C}$ are nonnegative.]

3. Gaussian Processes

Let T be a set, and $G = \{G(t)\}_{t\in T}$ a collection of random variables indexed by T. We say that G is a *Gaussian process* if $(G(t_1), \ldots, G(t_k))$ is a k-dimensional Gaussian random vector for all $t_1, \ldots, t_k \in T$. The *finite-dimensional distributions* of the process G are the collection of all probabilities obtained as follows:

$$\mu_{t_1,\ldots,t_k}(A_1, \ldots, A_k) := \mathrm{P}\left\{G(t_1) \in A_1, \ldots, G(t_k) \in A_k\right\}, \tag{3.1}$$

as $A_1, \ldots, A_k$ range over Borel subsets of $\mathbf{R}$, and k over all positive integers. In principle, these are the only pieces of information that one has about the random process G. All properties of G are supposed to follow from properties of these distributions.

A theorem of Kolmogorov (1933) states that the finite-dimensional distributions of G are uniquely determined by two functions:

1. The *mean function*: $\mu(t) := \mathrm{E}[G(t)]$; and
2. the covariance function $C(s, t) := \mathrm{Cov}(G(s), G(t))$.

Of course, μ is a real-valued function on T, whereas C is a real-valued function on $T \times T$.

Exercise 3.1. Prove that C is *nonnegative definite*. That is, prove that for all $t_1, \ldots, t_k \in T$ and all $z_1, \ldots, z_k \in \mathbf{C}$,

$$\sum_{j=1}^{k}\sum_{l=1}^{k} C(t_j, t_l) z_j \overline{z_l} \geq 0. \tag{3.2}$$

Exercise 3.2. Prove that whenever $C : T \times T \to \mathbf{R}$ is nonnegative definite,

$$|C(s, t)|^2 \leq C(s, s)C(t, t), \qquad \text{for all } s, t \in T. \tag{3.3}$$

This is the *Cauchy–Schwarz inequality*. In particular, $C(t, t) \geq 0$ for all $t \in T$.

Exercise 3.3. Prove that if G is a Gaussian process with mean function μ and covariance function C then $\{G(t) - \mu(t)\}_{t\in T}$ is a Gaussian process with mean function zero and covariance function C.

Exercise 3.4. Suppose there exist $E, F \subset T$ such that $C(s\,,t) = 0$ for all $s \in E$ and $t \in F$. Then prove that $\{G(s)\}_{s\in E}$ and $\{G(t)\}_{t\in F}$ are *independent* Gaussian processes. That is, prove that for all $s_1, \ldots, s_n \in E$ and all $t_1, \ldots, t_m \in F$, $(G(s_1)\,, \ldots, G(s_n))$ and $(G(t_1)\,, \ldots, G(t_m))$ are independent Gaussian random vectors.

A theorem of Bochner states that the collection of all nonnegative definite functions f on $T \times T$ matches all covariance functions, as long as f is symmetric. [*Symmetry* means that $f(s\,,t) = f(t\,,s)$.] This, and the aforementioned theorem of Kolmogorov, together imply that given a function $\mu : T \to \mathbf{R}$ and a nonnegative-definite function $C : T \times T \to \mathbf{R}$ there exists a Gaussian process $\{G(t)\}_{t\in T}$ whose mean function is μ and covariance function is C.

Example 3.5 (Brownian Motion)**.** Let $T = \mathbf{R}_+ := [0\,,\infty)$, $\mu(t) := 0$, and $C(s\,,t) := \min(s\,,t)$ for all $s, t \in \mathbf{R}_+$. I claim that C is nonnegative definite. Indeed, for all $z_1, \ldots, z_k \in \mathbf{C}$ and $t_1, \ldots, t_k \geq 0$,

$$\begin{aligned}\sum_{j=1}^k \sum_{l=1}^k \min(t_j\,,t_l) z_j \overline{z_l} &= \sum_{j=1}^k \sum_{l=1}^k z_j \overline{z_l} \int_0^\infty \mathbf{1}_{[0,t_j]}(x) \mathbf{1}_{[0,t_l]}(x)\,dx \\ &= \int_0^\infty \left| \sum_{j=1}^k \mathbf{1}_{[0,t_j]}(x) z_j \right|^2 dx,\end{aligned} \tag{3.4}$$

which is manifestly greater than or equal to zero. Because C is also symmetric it must be the covariance function of *some* mean-zero Gaussian process $B := \{B(t)\}_{t\geq 0}$. That process B is called *Brownian motion*; it was first invented by L. Bachelier (1900). Brownian motion has the following important additional property: *Let $s > 0$ be fixed. Then the process $\{B(t+s)-B(s)\}_{t\geq 0}$ is independent of $\{B(u)\}_{0\leq u\leq s}$.* Indeed, thanks to Exercise 3.4 it suffices to prove that for all $t \geq 0$ and $0 \leq u \leq s$, $\mathrm{E}[(B(t+s) - B(s))B(u)] = 0$ (why? Hash this out carefully!). But this is easy to see because

$$\begin{aligned}\mathrm{E}[(B(t+s) - B(s))B(u)] &= \mathrm{Cov}(B(t+s)\,, B(u)) - \mathrm{Cov}(B(s)\,, B(u)) \\ &= \min(t+s\,,u) - \min(s\,,u) = u - u = 0.\end{aligned} \tag{3.5}$$

By d-dimensional Brownian motion we mean the d-dimensional Gaussian process $B := \{(B_1(t)\,, \ldots, B_d(t))\}_{t\geq 0}$, where $B_1, \ldots, B_d$ are independent [one-dimensional] Brownian motions.

Exercise 3.6. Prove that if $s > 0$ is fixed and B is Brownian motion, then the process $\{B(t+s) - B(s)\}_{t\geq 0}$ is a *Brownian motion* independent of $\{B(u)\}_{0\leq u\leq s}$. This and the independent-increment property of B [Example 3.5] together prove that B is a *Markov process*.

Example 3.7 (Brownian Sheet)**.** Let $T := \mathbf{R}_+^N := [0\,,\infty)^N$, $\mu(\boldsymbol{t}) := 0$ for all $\boldsymbol{t} \in \mathbf{R}_+^N$, and define $C(\boldsymbol{s}\,,\boldsymbol{t}) := \prod_{n=1}^N \min(s_n\,,t_n)$ for all $\boldsymbol{s},\boldsymbol{t} \in \mathbf{R}_+^N$. Then C is a nonnegative-definite, symmetric function on $\mathbf{R}_+^N \times \mathbf{R}_+^N$, and the resulting mean-zero Gaussian process $B = \{B(\boldsymbol{t})\}_{\boldsymbol{t}\in\mathbf{R}_+^N}$ is the N-parameter *Brownian sheet*. When $N = 1$, this is just Brownian motion. One can also introduce d-dimensional, N-parameter Brownian sheet as the d-dimensional process whose coordinates are independent, [one-dimensional] N-parameter Brownian sheets.

Exercise 3.8 (Ornstein–Uhlenbeck Sheet)**.** Let $\{B(\boldsymbol{t})\}_{\boldsymbol{t}\in\mathbf{R}_+^N}$ denote N-parameter Brownian sheet, and define a new N-parameter stochastic process X as follows:

$$X(\boldsymbol{t}) := \frac{B\left(e^{t_1},\ldots,e^{t_N}\right)}{e^{(t_1+\cdots+t_N)/2}}, \qquad \text{for all } \boldsymbol{t} \in \mathbf{R}_+^N. \tag{3.6}$$

This is called the N-paramerter *Ornstein–Uhlenbeck sheet*. When $N = 1$, it is called the *Ornstein–Uhlenbeck process*. Prove that X is also a mean-zero, N-parameter Gaussian process and its covariance function $C(\boldsymbol{s}\,,\boldsymbol{t})$ depends on $(\boldsymbol{s}\,,\boldsymbol{t})$ only through $\sum_{i=1}^N |s_i - t_i|$. Such processes are called *stationary Gaussian processes*. This process was predicted in the works of noble laureates Ornstein and Uhlenbeck. Its existence was proved in a landmark paper of Doob (1942).

Example 3.9 (White Noise)**.** Let $T := \mathscr{B}(\mathbf{R}^N)$ denote the collection of all Borel-measurable subsets of $\mathbf{R}^N$, and $\mu(A) := 0$ for all $A \in \mathscr{B}(\mathbf{R}^N)$. Define $C(A\,,B) := |A \cap B|$, where $|\cdots|$ denotes the N-dimensional Lebesgue measure. Clearly, C is symmetric. It turns out that C is also nonnegative definite (Exercise 3.10). The resulting Gaussian process $\dot{W} := \{\dot{W}(A)\}_{A\in\mathscr{B}(\mathbf{R}^N)}$ is called *white noise* on $\mathbf{R}^N$.

Exercise 3.10. Complete the previous example by proving that the covariance of white noise is indeed a nonnegative-definite function on $\mathscr{B}(\mathbf{R}^N) \times \mathscr{B}(\mathbf{R}^N)$.

Exercise 3.11. Prove that if $A, B \in \mathscr{B}(\mathbf{R}^N)$ are disjoint then $\dot{W}(A)$ and $\dot{W}(B)$ are independent random variables. Use this to prove that if $A, B \in \mathscr{B}(\mathbf{R}^N)$ are non-random, then with probability one, $\dot{W}(A \cup B) = \dot{W}(A) + \dot{W}(B) - \dot{W}(A \cap B)$.

Exercise 3.12. Despite what the preceding may seem to imply, $\dot{W}$ is not a random signed measure in the obvious sense. Let $N = 1$ for simplicity. Then, prove that with probability one,

$$\lim_{n\to\infty} \sum_{j=0}^{2^n-1} \left| \dot{W}\left(\left[\frac{j-1}{2^n}\,,\,\frac{j}{2^n} \right] \right) \right|^2 = 1. \tag{3.7}$$

Use this to prove that with probability one,

$$\lim_{n\to\infty}\sum_{j=0}^{2^n-1}\left|\dot W\left(\left[\frac{j-1}{2^n}\,,\,\frac{j}{2^n}\right]\right)\right|=\infty. \tag{3.8}$$

Conclude that if $\dot W$ *were* a random measure then with probability one $\dot W$ is not sigma-finite. Nevertheless, the following example shows that one can integrate some things against $\dot W$.

Example 3.13 (The Isonormal Process)**.** Let $\dot W$ denote white noise on $\mathbf{R}^N$. We wish to define $\dot W(h)$ where h is a nice function. First, we identify $\dot W(A)$ with $\dot W(\mathbf{1}_A)$. More generally, we define for all disjoint $A_1,\ldots,A_k\in\mathscr{B}(\mathbf{R}^N)$ and $c_1,\ldots,c_k\in\mathbf{R}$,

$$\dot W\left(\sum_{j=1}^k c_j\mathbf{1}_{A_j}\right):=\sum_{j=1}^k c_j\dot W(A_j). \tag{3.9}$$

Thanks to Exercise 3.11, $\dot W(A_1),\ldots,\dot W(A_k)$ are independent. Therefore,

$$\left\|\dot W\left(\sum_{j=1}^k c_j\mathbf{1}_{A_j}\right)\right\|_{L^2(\mathrm{P})}^2=\sum_{j=1}^k c_j^2|A_j|=\left\|\sum_{j=1}^k c_j\mathbf{1}_{A_j}\right\|_{L^2(\mathbf{R}^N)}^2. \tag{3.10}$$

Classical integration theory tells us that for all $h\in L^2(\mathbf{R}^N)$ we can find h_n of the form $\sum_{j=1}^{k(n)}c_{jn}\mathbf{1}_{A_{j,n}}$ such that $A_{1,n},\ldots,A_{k(n),n}\in\mathscr{B}(\mathbf{R}^N)$ are disjoint and $\|h-h_n\|_{L^2(\mathbf{R}^N)}\to 0$ as $n\to\infty$. This, and (3.10) tell us that $\{\dot W(h_n)\}_{n=1}^\infty$ is a Cauchy sequence in $L^2(\mathrm{P})$. Denote their limit by $\dot W(h)$. This is the *Wiener integral of* $h\in L^2(\mathbf{R}^N)$, and is sometimes written as $\int h\,dW$ [no dot!]. Its key feature is that

$$\left\|\dot W(h)\right\|_{L^2(\mathrm{P})}=\|h\|_{L^2(\mathbf{R}^N)}. \tag{3.11}$$

That is, $\dot W: L^2(\mathbf{R}^N)\to L^2(\mathrm{P})$ is an isometry; (3.11) is called *Wiener's isometry* (Wiener, 1923a). [Note that we now know how to construct the stochastic integral $\int h\,dW$ only if $h\in L^2(\mathbf{R}^N)$ is *non-random*.] The process $\{\dot W(h)\}_{h\in L^2(\mathbf{R}^N)}$ is called the *isonormal process* (Dudley, 1967). It is a Gaussian process; its mean function is zero; and its covariance function is $C(h\,,g)=\int_{\mathbf{R}^N}h(x)g(x)\,dx$—the $L^2(\mathbf{R}^N)$ inner product—for all $h,g\in L^2(\mathbf{R}^N)$.

Exercise 3.14. Prove that for all [non-random] $h,g\in L^2(\mathbf{R}^N)$ and $a,b\in\mathbf{R}$,

$$\int(ah+bg)\,dW=a\int h\,dW+b\int h\,dW, \tag{3.12}$$

almost surely.

Exercise 3.15. Let $\{h_j\}_{j=1}^\infty$ be a complete orthonormal system [c.o.n.s.] in $L^2(\mathbf{R}^N)$. Then prove that $\{\dot W(h_j)\}_{j=1}^\infty$ is a c.o.n.s. in $L^2(\mathrm{P})$. In particular, for all random variables $Z\in$

$L^2(\mathrm{P})$,

$$Z = \sum_{j=1}^{\infty} a_j \dot{W}(h_j) \quad \text{almost surely, where} \qquad a_j := \mathrm{Cov}\left(Z\,, \dot{W}(h_j)\right), \tag{3.13}$$

and the infinite sum converges in $L^2(\mathrm{P})$. This permits one possible entry into the "Malliavin calculus."

Exercise 3.16. Prove that (3.9) is legitimate. That is, if we also have disjoint sets $B_1, \ldots, B_\ell \in \mathscr{B}(\mathbf{R}^N)$ and $d_1, \ldots, d_\ell \in \mathbf{R}$ such that $\sum_{j=1}^k c_j \mathbf{1}_{A_j} = \sum_{l=1}^\ell d_l \mathbf{1}_{B_l}$, then prove that

$$\dot{W}\left(\sum_{j=1}^k c_j \mathbf{1}_{A_j}\right) = \dot{W}\left(\sum_{l=1}^\ell d_l \mathbf{1}_{B_l}\right) \quad \text{almost surely.} \tag{3.14}$$

4. Regularity of Processes

Our construction of Gaussian processes is very general. That makes our construction both useful, as well as useless. It is useful because we can make sense of objects such as Brownian motion, Brownian sheet, white noise, etc. It is useless because our "random functions" [namely, the Brownian motion and more generally sheet] are not yet nice random functions. This has to do with the structure of Kolmogorov's existence theorem. But rather than discuss this technical subject let us consider a simple example.

Let $\{B(t)\}_{t\geq 0}$ denote the Brownian motion, and suppose U is an independent positive random variable with an absolutely continuous distribution. Define $B'(t) := B(t)$ if $t \neq U$, and $B'(t) = 5000$ if $t = U$. Then B' and B have the same finite-dimensional distributions. Therefore, B' is also a Brownian motion. This little example shows that there is no hope of proving that a given Brownian motion is (say) a continuous random function. [Sort the logic out!] Therefore, the best one can hope to do is to produce a *modification* of Brownian motion that is continuous. A remarkable theorem of Wiener (1923b) states that this can always be done. Thus, a *Wiener process* is a Brownian motion B such that the random function $t \mapsto B(t)$ is continuous. It suffices to define a "modification."

Definition 4.1. Let X and X' be two stochastic processes indexed by some set T. We say that X' is a *modification* of X if $\mathrm{P}\{X'(t) = X(t)\} = 1$ for all $t \in T$.

Exercise 4.2. Prove that any modification of a stochastic process X is a process with the same finite-dimensional distributions as X. Construct an example where X' is a modification of X, but $\mathrm{P}\{X' = X\} = 0$.

4.1. **A Diversion.** In order to gel the ideas we consider first a simple finite-dimensional example. Let $f \in L^1(\mathbf{R})$ and denote its Fourier transform by $\mathscr{F}f$, viz.,

$$(\mathscr{F}f)(z) := \int_{-\infty}^{\infty} e^{izx} f(x)\, dx. \tag{4.1}$$

Suppose, in addition, that $\mathscr{F}f \in L^1(\mathbf{R})$. We can proceed [intentionally] carelessly, and use the inversion formula,

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-izx} (\mathscr{F}f)(z)\, dz. \tag{4.2}$$

It follows readily from this and the dominated convergence theorem that f is uniformly continuous. But this cannot be so! For example, set $g(x) = f(x)$ for all $x \neq 0$, and $g(x) = f(x) + 1$. If f were continuous then g is not. But because $\mathscr{F}f = \mathscr{F}g$ the preceding argument would "show" that g is continuous too, which is a contradiction. The technical detail that we overlooked is that (4.2) is true only for almost all $x \in \mathbf{R}$. Therefore, the formula $(2\pi)^{-1} \int_{-\infty}^{\infty} e^{-izx} (\mathscr{F}f)(z)\, dz$ *defines* a "modification" of f which happens to be uniformly continuous.

4.2. **Kolmogorov's Continuity Theorem.** Now we come to the question, "when does a stochastic process X have a continuous modification?" If X is a Gaussian process then the answer is completely known, but is very complicated (Dudley, 1967; Preston, 1972; Fernique, 1975; Talagrand, 1985; 1987). When X is a fairly general process, there are also complicated sufficient conditions for the existence of a continuous modification. In the special case that X is a process indexed by $\mathbf{R}^N$ however there is a very useful theorem of Kolmogorov which gives a sufficient condition as well.

Theorem 4.3 (Kolmogorov's Continuity Theorem)**.** *Suppose $\{X(\boldsymbol{t})\}_{\boldsymbol{t}\in T}$ is a stochastic process indexed by a compact set $T \subset \mathbf{R}^N$. Suppose also that there exist constants $C > 0$, $p > 0$, and $\gamma > N$ such that uniformly for all $\boldsymbol{s}, \boldsymbol{t} \in T$,*

$$\mathrm{E}\left[|X(\boldsymbol{t}) - X(\boldsymbol{s})|^p\right] \leq C|\boldsymbol{t} - \boldsymbol{s}|^\gamma. \tag{4.3}$$

Then X has a continuous modification $\bar{X}$. Moreover, whenever $0 < \theta < (\gamma - N)$,

$$\left\| \sup_{\boldsymbol{s}\neq\boldsymbol{t}} \frac{|\bar{X}(\boldsymbol{s}) - \bar{X}(\boldsymbol{t})|}{|\boldsymbol{s} - \boldsymbol{t}|^\theta} \right\|_{L^p(\mathrm{P})} < \infty. \tag{4.4}$$

Remark 4.4. Here, $|\boldsymbol{x}|$ could be any Euclidean norm for $x \in \mathbf{R}^k$. Some examples are: $|\boldsymbol{x}| := \max_{1\leq j\leq k} |x_j|$; $|\boldsymbol{x}| := (|x_1|^p + \cdots + |x_k|^p)^{1/p}$ for $p \geq 1$; $|\boldsymbol{x}| := |x_1|^p + \cdots + |x_k|^p$ for $0 < p < 1$; and even an inhomogeneous norm such as $|\boldsymbol{x}| := (|x_1|^{p_1} + \cdots + |x_k|^{p_k})^{1/p}$ works where $1 \leq p_1, \ldots, p_k$ with $p := p_1 + \cdots + p_k$.

Definition 4.5 (Hölder Continuity)**.** A function $f : \mathbf{R}^N \to \mathbf{R}$ is said to be *globally Hölder continuous* with index α if there exists a constant A such that $|f(x) - f(y)| \le A|x-y|^\alpha$. It is said to be [locally] *Hölder continuous* with index α if for all compact sets $K \subset \mathbf{R}^N$ there exists a constant A_K such that $|f(x) - f(y)| \le A_K|x-y|^\alpha$ for all $x, y \in K$.

Exercise 4.6. Suppose $\{X(\boldsymbol{t})\}_{\boldsymbol{t}\in T}$ is a process indexed by a compact set $T \subset \mathbf{R}^N$ that satisfies (4.3) for some $C, p > 0$ and $\gamma > N$. Choose and fix $\alpha \in (0\,, (\gamma - N)/p)$. Prove that with probability one, X has a modification which is Hölder continuous with index α.

Exercise 4.7. Suppose $\{X(\boldsymbol{t})\}_{\boldsymbol{t}\in R^N}$ is a process indexed by $\mathbf{R}^N$. Suppose for all compact $T \subset \mathbf{R}^N$ there exist constants $C_T, p_T > 0$ and $\gamma := \gamma_T > N$ such that

$$\mathrm{E}\left[|X(\boldsymbol{s}) - X(\boldsymbol{t})|^{p_T}\right] \le C_T|\boldsymbol{s} - \boldsymbol{t}|^\gamma, \qquad \text{for all } \boldsymbol{s}, \boldsymbol{t} \in T. \tag{4.5}$$

Then, prove that X has a modification $\bar{X}$ which is [locally] Hölder continuous with some index ε_T. Warning: Mind your null sets!

Exercise 4.8 (Regularity of Gaussian Processes)**.** Suppose $\{X(\boldsymbol{t})\}_{\boldsymbol{t}\in T}$ is a Gaussian *random field*; i.e., a Gaussian process where $T \subseteq \mathbf{R}^N$ for some $N \ge 1$. Then, Check that for all $p > 0$,

$$\mathrm{E}\left(|X(\boldsymbol{t}) - X(\boldsymbol{s})|^p\right) = c_p\left[\mathrm{E}\left(|X(\boldsymbol{t}) - X(\boldsymbol{s})|^2\right)\right]^{p/2}, \tag{4.6}$$

where

$$c_p := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|x|^p e^{-x^2/2}\,dx = \frac{2^{p/2}}{\sqrt{\pi}}\Gamma\left(\frac{p+1}{2}\right). \tag{4.7}$$

Suppose we can find $\varepsilon > 0$ with the following property: For all compact sets $K \subset T$ there exists a positive and finite constant $A(K)$ such that

$$\mathrm{E}\left(|X(\boldsymbol{t}) - X(\boldsymbol{s})|^2\right) \le A(K)|\boldsymbol{t} - \boldsymbol{s}|^\varepsilon \qquad \text{for all } \boldsymbol{t}, \boldsymbol{s} \in K. \tag{4.8}$$

Then prove that X has a modification that is [locally] Hölder continuous of any given order $< \varepsilon/2$.

Example 4.9 (Brownian Motion)**.** Let $B := \{B(t)\}_{t\ge 0}$ denote a Brownian motion. Note that for all $s, t \ge 0$, $X(t) - X(s)$ is normally distributed with mean zero and variance $|t-s|$. Therefore, $\mathrm{E}(|X(t) - X(s)|^2) = |t-s|$ for all $s, t \ge 0$. It follows that X has a modification that is Hölder of any given order $\alpha < 1/2$. This is due to Wiener (1923b). **Warning:** This is not true for $\alpha = 1/2$. Let B denote the modification as well. [This should not be confusing.]

Then, a theorem of Khintchine (1933) asserts that

$$\mathrm{P}\left\{\limsup_{t\downarrow s}\frac{|B(t)-B(s)|}{\sqrt{2(t-s)\ln\ln\left(\frac{1}{t-s}\right)}}=1\right\}=1 \qquad \text{for all } s>0. \tag{4.9}$$

In particular, for all $s>0$,

$$\mathrm{P}\left\{\limsup_{t\downarrow s}\frac{|B(t)-B(s)|}{\sqrt{(t-s)}}=\infty\right\}=1. \tag{4.10}$$

Thus, B is not Hölder continuous of order $1/2$ at $s=0$, for instance.

Exercise 4.10. Let B denote N-parameter Brownian sheet. Prove that B has a modification which is [locally] Hölder continuous with any non-random index $\alpha\in(0\,,1/2)$. This generalized Wiener's theorem on Brownian motion.

Exercise 4.11. Let B be a continuous Brownian motion. Then prove that the event in (4.9) whose probability is one is measurable. Do the same for the event in (4.10).

5. Martingale Measures

5.1. **A White Noise Example.** Let $\dot W$ be white noise on $\mathbf{R}^N$. We have seen already that $\dot W$ is *not* a signed sigma-finite measure with any positive probability. However, it is not hard to deduce that it has the following properties:

(1) $\dot W(\varnothing)=0$ a.s.

(2) For all disjoint [non-random] sets $A_1,A_2,\ldots\in\mathscr{B}(\mathbf{R}^N)$,

$$\mathrm{P}\left\{\dot W\left(\bigcup_{i=1}^\infty A_i\right)=\sum_{i=1}^\infty \dot W(A_i)\right\}=1, \tag{5.1}$$

where the infinite sum converges in $L^2(\mathrm{P})$.

That is,

Proposition 5.1. *White noise is an $L^2(\mathrm{P})$-valued, sigma-finite, signed measure.*

Proof. In light of Exercise 3.11 it suffices to prove two things: (a) If $A_1\supset A_2\supset\cdots$ are all in $\mathscr{B}(\mathbf{R}^N)$ and $\cap A_n=\varnothing$, then $\dot W(A_n)\to 0$ in $L^2(\mathrm{P})$ as $n\to\infty$; and (b) For all compact sets K, $\mathrm{E}[(\dot W(K))^2]<\infty$.

It is easy to prove (a) because $\mathrm{E}[(\dot W(A_n))^2]$ is just the Lebesgue measure of A_n, and $|A_n|\to 0$ because Lebesgue measure is a measure. (b) is even easier to prove because $\mathrm{E}[(\dot W(K))^2]=|K|<\infty$ because Lebesgue measure is sigma-finite. □

Oftentimes in SPDEs one studies the "white-noise process" $\{W_t\}_{t\ge 0}$ defined by $W_t(A) := \dot{W}([0\,,t]\times A)$, where $A\in\mathscr{B}(\mathbf{R}^{N-1})$. This is a proper stochastic process as t varies, but an $L^2(\mathrm{P})$-type noise in A.

Let $\mathscr{F}$ be the filtration of the process $\{W_t\}_{t\ge 0}$. By this I mean the following: For all $t\ge 0$, we define $\mathscr{F}_t$ to be the sigma-algebra generated by $\{W_s(A);\, 0\le s\le t,\, A\in\mathscr{B}(\mathbf{R}^{N-1})\}$.

Exercise 5.2. Check that $\mathscr{F} := \{\mathscr{F}_t\}_{t\ge 0}$ is a *filtration* in the sense that $\mathscr{F}_s\subseteq\mathscr{F}_t$ whenever $s\le t$.

Lemma 5.3. $\{W_t(A)\}_{t\ge 0, A\in\mathscr{B}(\mathbf{R}^{N-1})}$ *is a "martingale measure" in the sense that:*

1. *For all* $A\in\mathscr{B}(\mathbf{R}^{N-1})$, $W_0(A)=0$ *a.s.;*
2. *If* $t>0$ *then* W_t *is a sigma-finite,* $L^2(\mathrm{P})$*-valued signed measure; and*
3. *For all* $A\in\mathscr{B}(\mathbf{R}^{N-1})$, $\{W_t(A)\}_{t\ge 0}$ *is a mean-zero martingale.*

Proof. Note that $\mathrm{E}[(W_t(A))^2] = t|A|$ where $|A|$ denotes the $(N-1)$-dimensional Lebesgue measure of A. Therefore, $W_0(A)=0$ a.s. This proves (1). (2) is proved in almost exactly the same way that Proposition 5.1 was. [Check the details!] Finally, choose and fix $A\in\mathscr{B}(\mathbf{R}^{N-1})$. Then, whenever $t\ge s\ge u\ge 0$,

$$\begin{aligned}\mathrm{E}\left[(W_t(A)-W_s(A))\,W_u(A)\right] &= \mathrm{E}\left[\left(\dot{W}\left([0\,,t]\times A\right)-\dot{W}\left([0\,,s]\times A\right)\right)\dot{W}\left([0\,,u]\times A\right)\right]\\ &= \min(t\,,u)|A| - \min(s\,,u)|A| = 0.\end{aligned}\tag{5.2}$$

Therefore, $W_t(A)-W_s(A)$ is independent of $\mathscr{F}_s$ (Exercise 3.4, page 4). As a result, with probability one,

$$\begin{aligned}\mathrm{E}\left[W_t(A)\mid\mathscr{F}_s\right] &= \mathrm{E}\left[W_t(A)-W_s(A)\mid\mathscr{F}_s\right] + W_s(A)\\ &= \mathrm{E}\left[W_t(A)-W_s(A)\right]+W_s(A)\\ &= W_s(A).\end{aligned}\tag{5.3}$$

This is the desired martingale property. ☐

Exercise 5.4. Choose and fix $A\in\mathscr{B}(\mathbf{R}^{N-1})$ such that $1/c := |A|^{1/2}>0$. Then prove that $\{cW_t(A)\}_{t\ge 0}$ is a Brownian motion.

Exercise 5.5 (Important)**.** Suppose $h\in L^2(\mathbf{R}^{N-1})$. Note that $t^{-1/2}W_t$ is white noise on $\mathbf{R}^{N-1}$. Therefore, we can define $W_t(h) := \int h(x)W_t(dx)$ for all $h\in L^2(\mathbf{R}^{N-1})$. Prove that $\{W_t(h)\}_{t\ge 0}$ is a continuous martingale with quadratic variation

$$\langle W_\bullet(h)\ ,\ W_\bullet(h)\rangle_t = t\int_{\mathbf{R}^{N-1}} h^2(x)\,dx.\tag{5.4}$$

It might help to recall that if $\{Z_t\}_{t\ge0}$ is a continuous $L^2(\mathrm{P})$-martingale, then its quadratic variation is uniquely defined as the continuous increasing process $\{\langle Z\,,Z\rangle_t\}_{t\ge0}$ such that $\langle Z\,,Z\rangle_0=0$ and $t\mapsto Z_t^2-\langle Z\,,Z\rangle_t$ is a continuous martingale. More generally, if Z and Y are two continuous $L^2(\mathrm{P})$-martingales then $Z_tY_t-\langle Z\,,Y\rangle_t$ is a continuous $L^2(\mathrm{P})$-martingale, and $\langle Z\,,Y\rangle_t$ is the only such "compensator." In fact prove that for all $t\ge0$ and $h,g\in L^2(\mathbf{R}^{N-1})$, $\langle W_\bullet(h)\,,W_\bullet(g)\rangle_t = t\int_{\mathbf{R}^{N-1}} h(x)g(x)\,dx$.

5.2. **More General Martingale Measures.** Let $\mathscr{F}:=\{\mathscr{F}_t\}_{t\ge0}$ be a filtration of sigma-algebras. We assume that $\mathscr{F}$ is *right-continuous*; i.e.,

$$\mathscr{F}_t=\bigcap_{s>t}\mathscr{F}_s \qquad \text{for all } t\ge0. \tag{5.5}$$

[This ensures that continuous-time martingale theory works.]

Definition 5.6 (Martingale Measures)**.** A process $\{M_t(A)\}_{t\ge0,A\in\mathscr{B}(\mathbf{R}^n)}$ is a *martingale measure* [with respect to $\mathscr{F}$] if:

(1) $M_0(A)=0$ a.s.;
(2) If $t>0$ then M_t is a sigma-finite $L^2(\mathrm{P})$-valued signed measure; and
(3) For all $A\in\mathscr{B}(\mathbf{R}^n)$, $\{M_t(A)\}_{t\ge0}$ is a mean-zero martingale with respect to the filtration $\mathscr{F}$.

Exercise 5.7. Double-check that you understand that if $\dot W$ is white noise on $\mathbf{R}^N$ then $W_t(A)$ defines a martingale measure on $\mathscr{B}(\mathbf{R}^{N-1})$.

Exercise 5.8. Let μ be a sigma-finite $L^2(\mathrm{P})$-valued signed measure on $\mathscr{B}(\mathbf{R}^n)$, and $\mathscr{F}:=\{\mathscr{F}_t\}_{t\ge0}$ a right-continuous filtration. Define $\mu_t(A):=\mathrm{E}[\mu(A)\,|\,\mathscr{F}_t]$ for all $t\ge0$ and $A\in\mathscr{B}(\mathbf{R}^n)$. Then prove that $\{\mu_t(A)\}_{t\ge0,A\in\mathscr{B}(\mathbf{R}^n)}$ is a martingale measure.

Exercise 5.9. Let $\{M_t(A)\}$ be a martingale measure. Prove that for all $T\ge t\ge0$, $M_t(A)=\mathrm{E}[M_T(A)\,|\,\mathscr{F}_t]$ a.s. Thus, every martingale measure locally look like those of the preceding exercise.

It turns out that martingale measures are a good class of integrators. In order to define stochastic integrals we follow Walsh (1986, Chapter 2), and proceed as one does when one constructs ordinary Itô integrals.

Definition 5.10. A function $f:\mathbf{R}^n\times\mathbf{R}_+\times\Omega\to\mathbf{R}$ is *elementary* if it has the form

$$f(x\,,t\,,\omega)=X(\omega)\mathbf{1}_{(a,b]}(t)\mathbf{1}_A(x), \tag{5.6}$$

where: (a) X is bounded and $\mathscr{F}_a$-measurable; and (b) $A \in \mathscr{B}(\mathbf{R}^n)$. Finite [nonrandom] linear combinations of elementary functions are called *simple* functions. Let $\mathscr{S}$ denote the class of all simple functions.

If M is a martingale measure and f is an elementary function of the form (6.2), then we define the stochastic-integral process of f as

$$(f \cdot M)_t(B)(\omega) := X(\omega)\left[M_{t\wedge b}(A\cap B) - M_{t\wedge a}(A \cap B)\right](\omega). \tag{5.7}$$

Exercise 5.11 (Important)**.** Prove that if f is an elementary function then $(f \cdot M)$ is a martingale measure. This constructs new martingale measures from old ones. For instance, if f is elementary and $\dot{W}$ is white noise then $(f \cdot W)$ is a martingale measure.

If $f \in \mathscr{S}$ then we can write f as $f = c_1 f_1 + \cdots + c_k f_k$ where $c_1, \ldots, c_k \in \mathbf{R}$ and $f_1, \ldots, f_k$ are elementary. We can then define

$$(f \cdot M)_t(B) := \sum_{j=1}^{k} c_j (f_j \cdot M)_t(B). \tag{5.8}$$

Exercise 5.12. Prove that the preceding is well defined. That is, prove that the definition of $(f \cdot M)$ does not depend on a particular representation of f in terms of elementary functions.

Exercise 5.13. Prove that if $f \in \mathscr{S}$ then $(f \cdot M)$ is a martingale measure. Thus, if $\dot{W}$ is white noise and $f \in \mathscr{S}$ then $(f \cdot W)$ is a martingale measure.

The right class of integrands are functions f that are "predictable." That is, they are measurable with respect to the "predictable sigma-algebra" $\mathscr{P}$ that is defined next.

Definition 5.14. Let $\mathscr{P}$ denote the sigma-algebra generated by all functions in $\mathscr{S}$. $\mathscr{P}$ is called the *predictable sigma-algebra*.

In order to go beyond stochastic integration of $f \in \mathscr{S}$ we need a technical condition—called "worthiness"—on the martingale measure M. This requires a little background.

Definition 5.15. Let M be a martingale measure. The *covariance functional* of M is defined as

$$\bar{Q}_t(A\,, B) := \langle M_\bullet(A)\,, M_\bullet(B)\rangle_t, \qquad \text{for all } t \geq 0 \text{ and } A, B \in \mathscr{B}(\mathbf{R}^n). \tag{5.9}$$

Exercise 5.16. Prove that:

(1) $\bar{Q}_t(A\,, B) = \bar{Q}_t(B\,, A)$ almost surely;

(2) If $B \cap C = \varnothing$ then $\bar{Q}_t(A\,, B \cup C) = \bar{Q}_t(A\,, B) + \bar{Q}_t(A\,, C)$ almost surely;
(3) $|\bar{Q}_t(A\,, B)|^2 \le \bar{Q}_t(A\,, A)\bar{Q}_t(B\,, B)$ almost surely; and
(4) $t \mapsto \bar{Q}_t(A\,, A)$ is almost surely non-decreasing.

Exercise 5.17. Let $\dot{W}$ be white noise on $\mathbf{R}^N$ and consider the martingale measure defined by $W_t(A) := \dot{W}((0\,,t] \times A)$, where $t \ge 0$ and $A \in \mathscr{B}(\mathbf{R}^{N-1})$. Verify that the quadratic functional of this martingale measure is described by $\bar{Q}_t(A\,, B) := t\mathscr{L}^{N-1}(A \cap B)$, where $\mathscr{L}^k$ denotes the Lebesgue measure on $\mathbf{R}^k$.

Next we define a random set function Q, in steps, as follows: For all $t \ge s \ge 0$ and $A, B \in \mathscr{B}(\mathbf{R}^n)$ define

$$Q\left(A\,, B\,; (s\,,t]\right) := \bar{Q}_t(A\,, B) - \bar{Q}_s(A\,, B). \tag{5.10}$$

If $A_i \times B_i \times (s_i\,, t_i]$ $(1 \le i \le m)$ are disjoint, then we can define

$$Q\left(\bigcup_{i=1}^n (A_i \times B_i \times (s_i\,, t_i])\right) := \sum_{i=1}^n Q\left(A_i\,, B_i\,, (s_i\,, t_i]\right). \tag{5.11}$$

This extends the definition of Q to rectangles. It turns out that, in general, one cannot go beyond this; this will make it impossible to define a completely general theory of stochastic integration in this setting. However, all works fine if M is "worthy" (Walsh, 1986). Before we define worthy martingale measures we point out a result that shows the role of Q.

Proposition 5.18. *Suppose $f \in \mathscr{S}$ and M is a worthy martingale measure. Then,*

$$\mathrm{E}\left[((f \cdot M)_t(B))^2\right] = \mathrm{E}\left[\iiint_{B \times B \times (0,t]} f(x\,,t)f(y\,,t)\,Q(dx\,dy\,dt)\right]. \tag{5.12}$$

Question 5.19. Although Q is not a proper measure, the triple-integral is well-defined. Why?

Proof. First we do this when f is elementary, and say has form (6.2). Then,

$$\begin{aligned}
&\mathrm{E}\left[(f \cdot M)_t^2(B)\right] \\
&= \mathrm{E}\left[X^2\left(M_{t\wedge b}(A \cap B) - M_{t\wedge a}(A \cap B)\right)^2\right] \\
&= \mathrm{E}\left[X^2 M_{t\wedge b}^2(A \cap B)\right] - 2\mathrm{E}\left[X^2 M_{t\wedge b}(A \cap B) M_{t\wedge a}(A \cap B)\right] + \mathrm{E}\left[X^2 M_{t\wedge a}^2(A \cap B)\right].
\end{aligned} \tag{5.13}$$

Recall that X is $\mathscr{F}_a$-measurable. Therefore, by the definition of quadratic variation,

$$\begin{aligned}
&\mathrm{E}\left[X^2\left(M_{t\wedge b}^2(A \cap B) - \langle M(A \cap B)\,, M(A \cap B)\rangle_{t\wedge b}\right)\right] \\
&\quad = \mathrm{E}\left[X^2\left(M_{t\wedge a}^2(A \cap B) - \langle M(A \cap B)\,, M(A \cap B)\rangle_{t\wedge a}\right)\right].
\end{aligned} \tag{5.14}$$

Similarly,

$$\begin{aligned}(5.15)\qquad & \mathrm{E}\left[X^2\left(M_{t\wedge b}(A\cap B)M_{t\wedge a}(A\cap B)-\langle M(A\cap B)\,,M(A\cap B)\rangle_{t\wedge a}\right)\right]\\ &= \mathrm{E}\left[X^2\left(M_{t\wedge a}^2(A\cap B)-\langle M(A\cap B)\,,M(A\cap B)\rangle_{t\wedge a}\right)\right].\end{aligned}$$

Combine to deduce the result in the case that f has form (6.2).

If $f\in\mathscr{S}$ then we can write $f=c_1f_1+\cdots+c_kf_k$ where $f_1,\ldots,f_k$ are elementary with disjoint support, and $c_1,\ldots,c_k$ are reals. [Why disjoint support?] Because $\mathrm{E}[(f_j\cdot M)_t]=0$, we know that $\mathrm{E}[(f\cdot M)_t^2(B)]=\sum_{j=1}^k c_j^2\mathrm{E}[(f_j\cdot M)_t^2(B)]$. The first part of the proof finishes the derivation. □

Definition 5.20. A martingale measure M is *worthy* if there exists a random sigma-finite measure $K(A\times B\times C\,,\omega)$—where $A,B\in\mathscr{B}(\mathbf{R}^n)$, $C\in\mathscr{B}(\mathbf{R}_+)$, and $\omega\in\Omega$—-such that:

1. $A\times B\mapsto K(A\times B\times C\,,\omega)$ is nonnegative definite and symmetric;
2. $\{K(A\times B\times(0\,,t])\}_{t\geq 0}$ is a predictable process (i.e., $\mathscr{P}$-measurable) for all $A,B\in\mathscr{B}(\mathbf{R}^n)$;
3. For all compact sets $A,B\in\mathscr{B}(\mathbf{R}^n)$ and $t>0$, $\mathrm{E}[K(A\times B\times(0\,,t])]<\infty$;
4. For all $A,B\in\mathscr{B}(\mathbf{R}^n)$ and $t>0$, $|Q(A\times B\times(0\,,t])|\leq K(A\times B\times(0\,,t])$ a.s.

[As usual, we drop the dependence on ω.] If and when such a K exists then it is called a *dominating measure* for M.

Remark 5.21. If M is worthy then Q_M can be extended to a measure on $\mathscr{B}(\mathbf{R}^n)\times\mathscr{B}(\mathbf{R}^n)\times\mathscr{B}(\mathbf{R}_+)$. This follows, basically, from the dominated convergence theorem.

Exercise 5.22 (Important)**.** Suppose $\dot W$ denotes white noise on $\mathbf{R}^N$, and consider the martingale measure on $\mathscr{B}(\mathbf{R}^{N-1})$ defined by $W_t(A)=W((0\,,t]\times A)$. Prove that it is worthy. Hint: Try the dominating measure $K(A\times B\times C):=\mathscr{L}^{N-1}(A\cap B)\mathscr{L}^1(C)$, where $\mathscr{L}^k$ denotes the Lebesgue measure on $\mathbf{R}^k$. Is this different than Q?

Proposition 5.23. *If M is a worthy martingale measure and $f\in\mathscr{S}$, then $(f\cdot M)$ is a worthy martingale measure. If Q_N and K_N respectively define the covariance functional and dominating measure of a worthy martingale measure N, then*

$$\begin{aligned}(5.16)\qquad Q_{f\cdot M}(dx\,,dy\,,dt)&=f(x\,,t)f(y\,,t)Q_M(dx\,,dy\,,dt),\\ K_{f\cdot M}(dx\,,dy\,,dt)&=|f(x\,,t)f(y\,,t)|K_M(dx\,,dy\,,dt).\end{aligned}$$

Proof. We will do this for elementary functions f; the extension to simple functions is routine. In light of Exercise 5.11 it suffices to compute $Q_{f\cdot M}$. The formula for $K_{f\cdot M}$ follows from this immediately as well.

Now, suppose f has the form (6.2), and note that for all $t \geq 0$ and $B, C \in \mathscr{B}(\mathbf{R}^n)$,

$$\begin{aligned}
&(f \cdot M)_t(B)(f \cdot M)_t(C) \\
&\quad = X^2 \left[M_{t\wedge b}(A\cap B) - M_{t\wedge a}(A\cap B)\right]\left[M_{t\wedge b}(A\cap C) - M_{t\wedge a}(A\cap C)\right] \\
&\quad = \text{martingale } + X^2 \left\langle M(A\cap B)\,, M(A\cap C)\right\rangle_{t\wedge b} - X^2 \left\langle M(A\cap B)\,, M(A\cap C)\right\rangle_{t\wedge a} \\
&\quad = \text{martingale } + X^2 Q_M\left((A\cap B)\times(A\cap B)\times(s\,,t]\right) \\
&\quad = \text{martingale } + \iiint_{B\times C\times(0\,,t]} f(x\,,s)f(y\,,s)Q_M(dx\,dy\,ds).
\end{aligned} \tag{5.17}$$

[Check!] This does the job. □

From now on we will be interested only in the case where the time variable t is in some finite interval $(0\,,T]$.

If K_M is the dominating measure for a worthy martingale measure M, then we define for all predictable function f,

$$\|f\|_M := \left(\mathrm{E}\left[\iiint_{\mathbf{R}^n\times\mathbf{R}^n\times(0,T]} |f(x\,,t)f(y\,,t)|\, K_M(dx\,dy\,dt)\right]\right)^{1/2}. \tag{5.18}$$

Let $\mathscr{P}_M$ denote the collection of all predictable functions f such that $\mathrm{E}(\|f\|_M) < \infty$.

Exercise 5.24. $\|\cdot\|_M$ is a norm on $\mathscr{P}$, and $\mathscr{P}_M$ is complete [hence a Banach space] in this norm.

I will not prove the following technical result. For a proof see Proposition 2.3 of Walsh (1986, p. 293).

Theorem 5.25. *$\mathscr{S}$ is dense in $\mathscr{P}_M$.*

Note from Proposition 5.18 that

$$\mathrm{E}\left[(f\cdot M)_t^2(B)\right] \leq \|f\|_M^2 \qquad \text{for all } t\in(0\,,T],\ f\in\mathscr{S},\ \text{and } B\in\mathscr{B}(\mathbf{R}^n). \tag{5.19}$$

Consequently, if $\{f_m\}_{m=1}^\infty$ is a Cauchy sequence in $(\mathscr{S}\,,\|\cdot\|_M)$ then $\{(f_m\cdot M)_t(B)\}_{m=1}^\infty$ is Cauchy in $L^2(\mathrm{P})$. If $f_m \to f$ in $\|\cdot\|_M$ then write the $L^2(\mathrm{P})$-limit of $(f_m\cdot M)_t(B)$ as $(f\cdot M)_t(B)$. A few more lines imply the following.

Theorem 5.26. *Let M be a worthy martingale measure. Then for all $f\in\mathscr{P}_M$, $(f\cdot M)$ is a worthy martingale measure that satisfies* (5.16)*. Moreover, for all $t\in(0\,,T]$ and $A, B\in$*

$\mathscr{B}(\mathbf{R}^n)$,

$$\langle (f\cdot M)(A)\,,(f\cdot M)(B)\rangle_t = \iiint\limits_{A\times B\times(0,t]} f(x\,,s)f(y\,,s)\,Q_M(dx\,dy\,ds),$$
$$\mathrm{E}\left[(f\cdot M)_t^2(B)\right] \le \|f\|_M^2. \tag{5.20}$$

The above $L^2(\mathrm{P})$ bound has an L^p version as well.

Theorem 5.27 (Burkholder's Inequality)**.** *For all $p \ge 2$ there exists $c_p \in (0\,,\infty)$ such that for all predictable f and all $t > 0$,*

$$\mathrm{E}\left[|(f\cdot M)_t(B)|^p\right] \le c_p \mathrm{E}\left[\left(\iiint\limits_{\mathbf{R}^n\times\mathbf{R}^n\times(0,T]} |f(x\,,t)f(y\,,t)|\,K_M(dx\,dy\,dt)\right)^{p/2}\right]. \tag{5.21}$$

Proof in a Special Case. It is enough to prove that if $\{N_t\}_{t\ge0}$ is a martingale with $N_0 := 0$ and quadratic variation $\langle N\,,N\rangle_t$ at time t, then $\|N_t\|_{L^p(\mathrm{P})}^p \le c_p\|\langle N\,,N\rangle_t\|_{L^{p/2}(\mathrm{P})}^{p/2}$, but this is precisely the celebrated Burkholder inequality (Burkholder, 1971). Here is why it is true in the case that N is a bounded *continuous* martingale. Recall Itô's formula (Itô, 1944; 1950; 1951): For all f that is C^2 a.e.,

$$f(N_t) = f(0) + \int_0^t f'(N_s)\,dN_s + \frac{1}{2}\int_0^t f''(N_s)\,d\langle N\,,N\rangle_s. \tag{5.22}$$

Apply this with $f(x) := |x|^p$ for $p > 2$ [$f''(x) = p(p-1)|x|^{p-2}$ a.e.] to find that

$$|N_t|^p = \frac{p(p-1)}{2}\int_0^t |N_s|^{p-2}\,d\langle N\,,N\rangle_s + \text{mean-zero martingale}. \tag{5.23}$$

Take expectations to find that

$$\mathrm{E}\left(|N_t|^p\right) \le \frac{p(p-1)}{2}\mathrm{E}\left(\sup_{0\le u\le t}|N_u|^{p-2}\langle N\,,N\rangle_t\right). \tag{5.24}$$

Because $|N_t|^p$ is a submartingale, Doob's maximal inequality says that $\mathrm{E}(\sup_{0\le u\le t}|N_u|^p) \le (p/(p-1))^p\mathrm{E}(|N_t|^p)$. Therefore, $\phi_p(t) := \mathrm{E}(\sup_{0\le u\le t}|N_u|^p)$ satisfies

$$\begin{aligned}\phi_p(t) &\le \frac{p(p-1)}{2}\left(\frac{p}{p-1}\right)^p \mathrm{E}\left(\sup_{0\le u\le t}|N_u|^{p-2}\langle N\,,N\rangle_t\right)\\ &:= a_p\mathrm{E}\left(\sup_{0\le u\le t}|N_u|^{p-2}\langle N\,,N\rangle_t\right).\end{aligned} \tag{5.25}$$

Apply Hölder's inequality to find that

$$\phi_p(t) \le a_p\left(\phi_p(t)\right)^{(p-2)/p}\left(\mathrm{E}\left[\langle N\,,N\rangle_t^{p/2}\right]\right)^{2/p}. \tag{5.26}$$

Solve to finish. □

Exercise 5.28. In the context of the preceding prove that for all $p \geq 2$ there exists $c_p \in (0\,,\infty)$ such that for all bounded stopping times T,

$$\mathrm{E}\left(\sup_{0\leq u\leq T}|N_u|^p\right) \leq c_p\mathrm{E}\left(\langle N\,,N\rangle_T^{p/2}\right). \tag{5.27}$$

Use this to improve itself in the following way: We do not need N to be a bounded martingale in order for the preceding to hold. [Hint: Localize.]

Exercise 5.29 (Harder)**.** In the context of the preceding prove that for all $p \geq 2$ there exists $c_p' \in (0\,,\infty)$ such that for all bounded stopping times T,

$$\mathrm{E}\left(\langle N\,,N\rangle_T^{p/2}\right) \leq c_p'\mathrm{E}\left(\sup_{0\leq u\leq T}|N_u|^p\right). \tag{5.28}$$

Hint: Start with $\langle N\,,N\rangle_t = N_t^2 - \int_0^t N_s\,dN_s \leq N_t^2 + |\int_0^t N_s\,dN_s|$.

From now on we adopt a more standard stochastic-integral notation:

$$(f\cdot M)_t(A) := \iint_{A\times(0,t]} f\,dM := \iint_{A\times(0,t]} f(x\,,s)\,M(dx\,ds). \tag{5.29}$$

[N.B.: The last $f(x\,,s)$ is actually $f(x\,,s\,,\omega)$, but we have dropped the ω as usual.] These martingale integrals have the Fubini–Tonelli property:

Theorem 5.30. *Suppose M is a worthy martingale measure with dominating measure K. Let $(A\,,\mathscr{A}\,,\mu)$ be a measure space and $f:\mathbf{R}^n\times\mathbf{R}_+\times\Omega\times A\to\mathbf{R}$ measurable such that*

$$\iiiint\!\!\int_{\Omega\times\mathbf{R}^n\times\mathbf{R}^n\times[0,T]\times A} |f(x\,,t\,,\omega\,,u)f(y\,,t\,,\omega\,,u)|\,K(x\,,y\,,dt)\,\mu(du)\,\mathrm{P}(d\omega) < \infty. \tag{5.30}$$

Then almost surely,

$$\begin{aligned}&\int_A\left(\iint_{\mathbf{R}^n\times[0,t]} f(x\,,s\,,\bullet\,,u)\,M(dx\,ds)\right)\mu(du)\\ &\qquad = \iint_{\mathbf{R}^n\times[0,t]}\left(\int_A f(x\,,s\,,\bullet\,,u)\,\mu(du)\right)M(dx\,ds).\end{aligned} \tag{5.31}$$

It suffices to prove this for elementary functions of the form (6.2). You can do this yourself, or consult Walsh (1986, p. 297).

6. The Semilinear Heat Equation

We are ready to try and study a class of SPDEs. Let $L > 0$ be fixed, and consider

$$\begin{aligned} &\partial_t u = \partial_{xx} u + f(u)\dot{W}, && t>0,\ x\in[0\,,L], \\ &\partial_x u(0\,,t) = \partial_x u(L\,,t) = 0, && t>0, \\ &u(x\,,0) = u_0(x), && x\in[0\,,L], \end{aligned} \tag{6.1}$$

where $\dot{W}$ is white noise with respect to some given filtration $\{\mathscr{F}_t\}_{t\geq 0}$, and $u_0 : [0\,,L] \to \mathbf{R}$ is a non-random, measurable, and bounded function. As regards the function $f : \mathbf{R} \to \mathbf{R}$, we assume that

$$K := \sup_{0\leq x\neq y\leq L} \frac{|f(x)-f(y)|}{|y-x|} + \sup_{0\leq x\leq L} |f(x)| < \infty. \tag{6.2}$$

In other words, we assume that f is globally Lipschitz, as well as bounded.

Exercise 6.1. Recall that $f : \mathbf{R} \to \mathbf{R}$ is *globally Lipschitz* if there exists a constant A such that $|f(x) - f(y)| \leq A|x - y|$ for all $x, y \in \mathbf{R}$. Verify that any globally Lipschitz function $f : \mathbf{R} \to \mathbf{R}$ satisfies $|f(x)| = O(|x|)$ as $|x| \to \infty$. That is, prove that f has at most linear growth.

Now we multiply (6.1) by $\phi(x)$ and integrate $[dt\,dx]$ to find (formally, again) that for all $\phi \in C^\infty([0\,,L])$ with $\phi'(0) = \phi'(L) = 0$,

$$\begin{aligned} &\int_0^L u(x\,,t)\phi(x)\,dx - \int_0^L u_0(x)\phi(x)\,dx \\ &= \int_0^t\int_0^L \partial_{xx}u(x\,,s)\phi(x)\,dx\,ds + \int_0^t\int_0^L f\left(u(x\,,s)\right)\phi(x)W(dx\,ds). \end{aligned} \tag{6.3}$$

Certainly we understand the stochastic integral now. But $\partial_{xx}u$ is not well defined. Therefore, we try and integrate by parts (again formally!): Because $\phi'(0) = \phi'(L) = 0$, the boundary-values of $\partial_x u$ [formally speaking] imply that

$$\int_0^t\int_0^L \partial_{xx}u(x\,,s)\phi(x)\,dx\,ds = \int_0^t\int_0^L u(x\,,s)\phi''(x)\,dx\,ds. \tag{6.4}$$

And now we have ourselves a proper stochastic-integral equation: Find u such that for all $\phi \in C^\infty([0\,,L])$ with $\phi'(0) = \phi'(L) = 0$,

$$\begin{aligned} &\int_0^L u(x\,,t)\phi(x)\,dx - \int_0^L u_0(x)\phi(x)\,dx \\ &= \int_0^t\int_0^L u(x\,,s)\phi''(x)\,dx\,ds + \int_0^t\int_0^L f\left(u(x\,,s)\right)\phi(x)W(dx\,ds). \end{aligned} \tag{6.5}$$

Exercise 6.2 (Important)**.** Argue that if u solves (6.5), then for all C^∞ functions $\psi(x\,,t)$ with $\partial_x\psi(0\,,t)=\partial_x\psi(L\,,t)=0$,

$$\begin{aligned}&\int_0^L u(x\,,t)\psi(x\,,t)\,dx-\int_0^L u_0(x)\psi(x\,,0)\,dx\\ &=\int_0^t\int_0^L u(x\,,s)\left[\partial_{xx}\psi(x\,,s)+\partial_t\psi(x\,,s)\right]\,dx\,ds\\ &\quad+\int_0^t\int_0^L f\left(u(x\,,s)\right)\psi(x\,,s)W(dx\,ds).\end{aligned}\tag{6.6}$$

This is formal, but important.

Let $G_t(x\,,y)$ denote the Green's function for the semilinear heat equation. [The subscript t is *not* a derivative, but a variable.] Then there exists a constant $c>0$ such that

$$G_t(x\,,y)=c\sum_{n=-\infty}^{\infty}[\Gamma(t\,;x-y-2nL)+\Gamma(t\,;x+y+2nL)],\tag{6.7}$$

where Γ is the fundamental solution to the linear heat equation (6.1); i.e.,

$$\Gamma(t\,;a)=\frac{1}{(4\pi t)^{1/2}}\exp\left(-\frac{a^2}{4t}\right).\tag{6.8}$$

This follows from the method of images.

Define for all smooth $\phi:[0\,,L]\to\mathbf{R}$,

$$G_t(\phi\,,y):=\int_0^L G_t(x\,,y)\phi(x)\,dx,\tag{6.9}$$

if $t>0$, and $G_0(\phi\,,y):=\phi(y)$. We can integrate (6.1)—with $f(u\,,t)\equiv 0$—by parts for all C^∞ functions $\phi:[0\,,L]\to\mathbf{R}$ such that $\phi'(0)=\phi'(L)=0$, and obtain the following:

$$G_t(\phi\,,y)=\phi(y)+\int_0^t G_s\left(\phi''-\phi\,,y\right)\,ds.\tag{6.10}$$

Fix $t>0$ and define $\psi(x\,,s):=G_{t-s}(\phi\,,x)$ to find that ψ solves

$$\partial_{xx}\psi(x\,,s)+\partial_s\psi(x\,,s)=0,\quad \psi(x\,,t)=\phi(x),\ \psi(x\,,0)=G_t(\phi\,,x).\tag{6.11}$$

Use this ψ in Exercise 6.2 to find that any solution to (6.1) must satisfy

$$\int_0^L u(x\,,t)\phi(x)\,dx-\int_0^L u_0(y)G_t(\phi\,,y)\,dy=\int_0^t\int_0^L f\left(u(y\,,s)\right)G_{t-s}(\phi\,,y)W(dy\,ds).\tag{6.12}$$

This must hold for all smooth ϕ with $\phi'(0)=\phi'(L)=0$. Therefore, we would expect that for Lebesgue-almost all $(x\,,t)$,

$$u(x\,,t)-\int_0^L u_0(y)G_t(x\,,y)\,dy=\int_0^t\int_0^L f\left(u(y\,,s)\right)G_{t-s}(x\,,y)W(dy\,ds).\tag{6.13}$$

If $\dot{W}$ were smooth then this reasoning would be rigorous and honest. As things are, it is still merely a formality. However, we are naturally led to a place where we have an honest stochastic-integral equation.

Definition 6.3. By a "solution" to the formal stochastic heat equation (6.1) we mean a solution u to (6.13) that is adapted. Sometimes this is called a *mild solution*.

With this nomenclature in mind, let us finally prove something.

Theorem 6.4. *The stochastic heat equation* (6.5) *subject to* (6.2) *has an a.s.-unique solution* u *that satisfies the following for all* $T > 0$*:*

$$\sup_{0\le x\le L}\sup_{0\le t\le T} \mathrm{E}\left[u^2(x\,,t)\right] < \infty. \tag{6.14}$$

For its proof we will need the following well-known result whose proof follows from a direct application of induction.

Lemma 6.5 (Gronwall's lemma)**.** *Suppose* $\phi_1, \phi_2, \ldots : [0\,,T] \to \mathbf{R}_+$ *are measurable and non-decreasing. Suppose also that there exist a constant* A *such that for all integers* $n \ge 1$*, and all* $t \in [0\,,T]$*,*

$$\phi_{n+1}(t) \le A\int_0^t \phi_n(s)\,ds. \tag{6.15}$$

Then, $\phi_n(t) \le \phi_1(T)\,(At)^{n-1}/(n-1)!$ *for all* $n \ge 1$ *and* $t \in [0\,,T]$*.*

Remark 6.6. Thanks to the DeMoivre–Stirling formula, $n! \sim \sqrt{2\pi} n^{n+(1/2)} e^{-n}$ as $n \to \infty$. Therefore, there exists a constant B such that $n! \ge B(2AT)^n$ for all integers $n \ge 1$. Thus, it follows that $\phi_n(t) \le \phi_1(T) B^{-1} 2^{-n-1}$. In particular, any positive power of $\phi_n(t)$ is summable in n. Also, if ϕ_n does not depend on n, then it follows that $\phi_n \equiv 0$.

Proof of Theorem 6.4: Uniqueness. Suppose u and v both solve (6.13), and both satisfy the integrability condition (6.14). We wish to prove that u and v are modifications of one another. Let $d(x\,,t) := u(x\,,t) - v(x\,,t)$. Then,

$$d(x\,,t) = \int_0^t\int_0^L \Big[f(u(y\,,s)) - f(v(y\,,s))\Big] G_{t-s}(x\,,y)\,W(dy\,ds). \tag{6.16}$$

According to Theorem 5.26 and (6.2),

$$\mathrm{E}\left[d^2(x\,,t)\right] \le K^2\int_0^t\int_0^L \mathrm{E}\left[d^2(y\,,s)\right] G^2_{t-s}(x\,,s)\,dy\,ds. \tag{6.17}$$

Let $H(t) := \sup_{0\le x\le L}\sup_{0\le s\le t} \mathrm{E}[d^2(x\,,s)]$. The preceding implies that

$$H(t) \le K^2 \int_0^t H(s) \left(\int_0^L G_{t-s}^2(x\,,y)\,dy \right) ds. \tag{6.18}$$

Now from (6.7) and the semigroup properties of Γ it follows that

$$\int_0^L G_t(x\,,y) G_s(y\,,z)\,dy = G_{t+s}(x\,,z), \text{ and } G_t(x\,,y) = G_t(y\,,x). \tag{6.19}$$

Consequently, $\int_0^L G_t^2(x\,,y)\,dy = G_{2t}(x\,,x) = Ct^{-1/2}$. Hence,

$$H(t) \le CK^2 \int_0^t \frac{H(s)}{\sqrt{t-s}}\,ds. \tag{6.20}$$

Now choose and fix some $p \in (1\,,2)$, let q be the conjugate to p [i.e., $p^{-1}+q^{-1} = 1$], and apply Hölder's inequality to find that there exists $A = A_T$ such that uniformly for all $t \in [0\,,T]$,

$$H(t) \le A \left(\int_0^t H(s)^q\,ds \right)^{1/q}. \tag{6.21}$$

Apply Gronwall's Lemma 6.5 with $\phi_1 = \phi_2 = \cdots = H^q$ to find that $H(t) \equiv 0$. □

Proof of Theorem 6.4: Existence. Note from (6.7) that $0 \le \int_0^L G_t(x\,,y)\,dy \le 1$. Because u_0 is assumed to be bounded this proves that $\int_0^L u_0(y) G_t(x\,,y)\,dy$ is bounded; this is the first term in (6.13). Now we proceed with a Picard-type iteration scheme. Let $u_0(x\,,t) := u_0(x)$, and then iteratively define

$$u_{n+1}(x\,,t) = \int_0^L u_0(y) G_t(x\,,y)\,dy + \int_0^t \int_0^L f\left(u_n(y\,,s)\right) G_{t-s}(x\,,y) W(dy\,ds). \tag{6.22}$$

Define $d_n(x\,,t) := u_{n+1}(x\,,t) - u_n(x\,,t)$ to find that

$$d_n(x\,,t) = \int_0^t \int_0^L \left[f\left(u_{n+1}(y\,,s)\right) - f\left(u_n(y\,,s)\right)\right] G_{t-s}(x\,,y)\,W(dy\,ds). \tag{6.23}$$

Consequently, by (6.2),

$$\mathrm{E}\left[d_n^2(x\,,t)\right] \le K^2 \int_0^t \int_0^L \mathrm{E}\left[d_{n-1}^2(y\,,s)\right] G_{t-s}^2(x\,,y)\,dy\,ds. \tag{6.24}$$

Let $H_n^2(t) := \sup_{0\le x\le L}\sup_{0\le s\le t} \mathrm{E}[d_n^2(x\,,s)]$ to find that

$$H_n^2(t) \le CK^2 \int_0^t \frac{H_{n-1}^2(s)}{\sqrt{t-s}}\,ds. \tag{6.25}$$

Choose and fix $p \in (0\,,2)$, and let q denote its conjugate so that $q^{-1}+p^{-1}=1$. Apply Hölder's inequality to find that there exists $A = A_T$ such that uniformly for all $t \in [0\,,T]$,

$$H_n^2(t) \le A\left(\int_0^t H_{n-1}^{2q}\,ds\right)^{1/q}. \tag{6.26}$$

Apply Gronwall's Lemma 6.5 with $\phi_n := H_n^{2q}$ to find that $\sum_{n=1}^{\infty} H_n(t) < \infty$. Therefore, $u_n(t\,,x)$ converges in $L^2(\mathrm{P})$ to some $u(t\,,x)$ for each t and x. This proves also that as n tends to infinity, $\int_0^t\int_0^L f(u_n(y\,,s))G_{t-s}(x\,,y)\,W(dy\,ds)$ converges in $L^2(\mathrm{P})$ to the same object but where u_n is replaced by u. This proves that u is a solution to (6.13). □

We are finally ready to complete the picture by proving that the solution to (6.1) is continuous [up to a modification, of course].

Theorem 6.7. *There exists a continuous modification $u(x\,,t)$ of (6.1).*

Remark 6.8. In Exercise 6.9 below you will be asked to improve this to the statement that there exists a Hölder-continuous modification.

Sketch of Proof. We need the following easy-to-check facts about the Green's function G: First of all,

$$G_t(x\,,y) = \Gamma(t\,;x-y) + H_t(x\,,y), \tag{6.27}$$

where $H_t(x\,,y)$ is smooth in $(t\,,x\,,y) \in \mathbf{R}_+ \times \mathbf{R} \times \mathbf{R}$, and Γ is the "heat kernel" defined in (6.8). Define

$$U(x\,,t) := \int_0^t\int_0^L f(u(y\,,s))\Gamma(t-s\,;x-y)\,W(dy\,ds). \tag{6.28}$$

The critical step is to prove that U has a continuous modification. Because u_0 is bounded it is then not too hard to complete the proof based on this, and the fact that the difference between Γ and G is smooth and bounded. From here on I prove things honestly.

Let $0 \le t \le t'$ and note that

$$\begin{aligned} U(x\,,t') - U(x\,,t) &= \int_0^t\int_0^L f(u(y\,,s))\left[\Gamma(t'-s\,;x-y) - \Gamma(t-s\,;x-y)\right]W(dy\,ds) \\ &\quad + \int_t^{t'}\int_0^L f(u(y\,,s))\Gamma(t'-s\,;x-y)\,W(dy\,ds). \end{aligned} \tag{6.29}$$

By Burkholder's inequality (Theorem 5.27) and the inequality $|a+b|^p \le 2^p|a|^p + 2^p|b|^p$,

$$\begin{aligned}
&\mathrm{E}\left[|U(x\,,t)-U(x\,,t')|^p\right] \\
&\le 2^p c_p \mathrm{E}\left[\left(\int_0^t\int_0^L f^2(u(y\,,s))\left[\Gamma(t'-s\,;x-y)-\Gamma(t-s\,;x-y)\right]^2\,dy\,ds\right)^{p/2}\right] \\
&\quad + 2^p c_p \mathrm{E}\left[\left(\int_t^{t'}\int_0^L f^2(u(y\,,s))\Gamma^2(t-s\,;x-y)\,dy\,ds\right)^{p/2}\right].
\end{aligned} \tag{6.30}$$

Because of (6.2), $\sup|f| \le K$; see (6.2). Therefore,

$$\begin{aligned}
&\mathrm{E}\left[|U(x\,,t)-U(x\,,t')|^p\right] \\
&\le (2K)^p c_p \left(\int_0^t\int_{-\infty}^\infty \left[\Gamma(t'-s\,;x-y)-\Gamma(t-s\,;x-y)\right]^2\,dy\,ds\right)^{p/2} \\
&\quad + (2K)^p c_p \left(\int_t^{t'}\int_{-\infty}^\infty \Gamma^2(t-s\,;x-y)\,dy\,ds\right)^{p/2}.
\end{aligned} \tag{6.31}$$

[Notice the change from $\int_0^L$ to $\int_{-\infty}^\infty$.] Because $\int_{-\infty}^\infty \Gamma^2(t-s\,;a)\,da = C/\sqrt{|t-s|}$,

$$\left(\int_t^{t'}\int_{-\infty}^\infty \Gamma^2(t-s\,;x-y)\,dy\,ds\right)^{p/2} = C_p|t'-t|^{p/4}. \tag{6.32}$$

For the other integral we use a method that is motivated by the ideas in Dalang (1999): Recall Plancherel's theorem: For all $g \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$,

$$\|g\|_{L^2(\mathbf{R})}^2 = \frac{1}{2\pi}\|\mathscr{F}g\|_{L^2(\mathbf{R})}^2, \tag{6.33}$$

where $(\mathscr{F}g)(z) := \int_{-\infty}^\infty g(x)e^{ixz}\,dx$ denotes the Fourier transform [in the space variable]. Because $(\mathscr{F}\Gamma)(t\,;\xi) = \exp(-t\xi^2)$,

$$\begin{aligned}
\int_{-\infty}^\infty \left[\Gamma(t'-s\,;x-y)-\Gamma(t-s\,;x-y)\right]^2\,dy &= \frac{1}{2\pi}\int_{-\infty}^\infty \left[e^{-(t'-s)\xi^2}-e^{-(t-s)\xi^2}\right]^2\,d\xi \\
&= \frac{1}{2\pi}\int_{-\infty}^\infty e^{-2(t-s)\xi^2}\left[1-e^{-(t'-t)\xi^2}\right]^2\,d\xi.
\end{aligned} \tag{6.34}$$

Therefore,

$$\begin{aligned}
&\int_0^t\int_{-\infty}^\infty \left[\Gamma(t'-s\,;x-y)-\Gamma(t-s\,;x-y)\right]^2\,dy\,ds \\
&\quad = \frac{1}{2\pi}\int_{-\infty}^\infty \left(\int_0^t e^{-2(t-s)\xi^2}\,ds\right)\left[1-e^{-(t'-t)\xi^2}\right]^2\,d\xi \\
&\quad = \frac{1}{4\pi}\int_{-\infty}^\infty \frac{1-e^{-2t\xi^2}}{\xi^2}\left[1-e^{-(t'-t)\xi^2}\right]^2\,d\xi.
\end{aligned} \tag{6.35}$$

A little thought shows that $(1-e^{-2t\xi^2})/\xi^2 \le C_T/(1+\xi^2)$, uniformly for all $0 \le t \le T$. Also, $[1-e^{-(t'-t)\xi^2}]^2 \le 2\min[(t'-t)\xi^2\,,1]$. Therefore,

$$
\begin{aligned}
&\int_0^t \int_{-\infty}^\infty \left[\Gamma(t'-s\,;x-y)-\Gamma(t-s\,;x-y)\right]^2\,dy\,ds \\
&\qquad \le \frac{C_T}{\pi}\int_0^\infty \frac{\min[(t'-t)\xi^2\,,1]}{1+\xi^2}\,d\xi \\
&\qquad \le \frac{C_T}{\pi}\left\{\int_{1/\sqrt{t'-t}}^\infty \frac{d\xi}{\xi^2} + \int_0^{1/\sqrt{t'-t}} \frac{(t'-t)\xi^2}{1+\xi^2}\,d\xi\right\}.
\end{aligned}
\tag{6.36}
$$

The first term is equal to $A\sqrt{t'-t}$, and the second term is also bounded above by $\sqrt{t'-t}$ because $\xi^2/(1+\xi^2)\le 1$. This, (6.31) and (6.32) together prove that

$$
\mathrm{E}\left[|U(x\,,t)-U(x\,,t')|^p\right] \le C_p|t'-t|^{p/4}. \tag{6.37}
$$

Similarly, we prove that for all $x,x' \in [0\,,L]$,

$$
\begin{aligned}
&\mathrm{E}\left[|U(x\,,t)-U(x'\,,t)|^p\right] \\
&\quad \le c_p K^p \left(\int_0^t\int_{-\infty}^\infty \left|\Gamma(t-s\,;y)-\Gamma(t-s\,;x'-x-y)\right|^2 dy\,ds\right)^{p/2}.
\end{aligned}
\tag{6.38}
$$

[Do it!] By Plancherel's theorem, and because the Fourier transform of $x \mapsto g(x+a)$ is $e^{-i\xi a}(\mathscr{F}g)(\xi)$,

$$
\int_{-\infty}^\infty \left|\Gamma(t-s\,;y)-\Gamma(t-s\,;x'-x-y)\right|^2 dy = \frac{1}{2\pi}\int_{-\infty}^\infty e^{-2(t-s)\xi^2}\left|1-e^{i\xi(x'-x)}\right|^2 d\xi. \tag{6.39}
$$

Use the elementary bound, $|1-e^{i\theta}|^2 \le |\theta|^{1-\varepsilon}$, valid for all $\theta\in\mathbf{R}$ and $\varepsilon\in(0\,,1)$, to deduce that

$$
\sup_{0\le t\le T}\mathrm{E}\left[|U(x\,,t)-U(x'\,,t)|^p\right] \le a_{p,T}|x'-x|^{p(1-\varepsilon)/2}. \tag{6.40}
$$

For all $(x\,,t)\in\mathbf{R}^2$ define $|(x\,,t)| := |x|^{(1-\varepsilon)/2}+|t|^{1/4}$. This defines a norm on $\mathbf{R}^2$, and is equivalent to the usual Euclidean norm $\sqrt{x^2+t^2}$ in the sense that both generate the same topology, etc. Moreover, we have by (6.37) and (6.40): For all $t,t'\in[0\,,T]$ and $x,x'\in[0\,,L]$,

$$
\mathrm{E}\left(|U(x\,,t)-U(x'\,,t')|^p\right) \le A\,|(x\,,t)-(x'\,,t')|^p\,. \tag{6.41}
$$

This and Kolmogorov's continuity theorem (Theorem 4.3) together prove that U has a modification which is continuous, in our inhomogeneous norm on $(x\,,t)$, of any order < 1. Because our norm is equivalent to the usual Euclidean norm, this proves continuity in the ordinary sense. □

Exercise 6.9. Complete the proof. Be certain that you understand why we have derived Hölder continuity. For example, prove that there is a modification of our solution which is Hölder continuous in x of any given order $< 1/2$; and it is Hölder continuous in t of any given order $< 1/4$.

Exercise 6.10. Improve (6.40) to the following: There exists C_p such that for all $x, x' \in [0\,, L]$ and $t \in [0\,, T]$, $\mathrm{E}[|U(x\,,t) - U(x'\,,t)|^p] \le C_p|x - x'|^{p/2}$.

Exercise 6.11. Consider the constant-coefficient, free-space stochastic heat equation in two space variables. For instance, here is one formulation: Let $\dot{W}(x\,,t)$ denote white noise on $(x\,,t) \in \mathbf{R}^2 \times \mathbf{R}_+$, and consider

$$\begin{aligned} \partial_t u &= \Delta u + \dot{W}, \qquad t > 0,\ x \in \mathbf{R}^2, \\ u(x\,,0) &= 0, \qquad\qquad x \in \mathbf{R}^2. \end{aligned} \tag{6.42}$$

Here $\Delta := \partial_{x_1x_1} + \partial_{x_2x_2}$ denotes the Laplacian. Try to interpret this SPDE as the adapted solution to the following:

$$u(x\,,t) = \int_0^t \int_{\mathbf{R}^2} \Gamma(t-s\,;x-y)\, W(dy\, ds), \tag{6.43}$$

subject to $(t\,,x) \mapsto \mathrm{E}[u^2(t\,,x)]$ being continuous (say!). Here, Γ is the heat kernel on $\mathbf{R}^2$; i.e., $\Gamma(t\,,x) := (4\pi t)^{-1}\exp(-\|x\|^2/(2t))$. Prove that $\mathrm{E}[u^2(x\,,t)] = \infty$ for all $x \in \mathbf{R}^2$ and $t > 0$. Prove also that if $u(x\,,t)$ were a proper stochastic process then it would have to be a Gaussian process, but this cannot be because Gaussian processes have finite moments. Therefore, in general, one cannot hope to find function-valued solutions to the stochastic heat equation in spatial dimensions ≥ 2.

References

Bachelier, Louis. 1900. *Théorie de la spéculation*, Ann. Sci. École Norm. Sup. **17**, 21–86, [See also the 1995 reprint. Sceaux: Gauthier–Villars].

Burkholder, D. L. 1971. *Martingale inequalities*, Martingales, Springer, Berlin, pp. 1–8. Lecture Notes in Math., Vol. 190.

Dalang, Robert C. 1999. *Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s*, Electron. J. Probab. **4**, no. 6, 29 pp. (electronic).

Doob, J. L. 1942. *The Brownian movement and stochastic equations*, Ann. of Math. (2) **43**, 351–369.

Dudley, R. M. 1967. *The sizes of compact subsets of Hilbert space and continuity of Gaussian processes*, J. Functional Analysis **1**, 290–330.

Fernique, X. 1975. *Regularité des trajectoires des fonctions aléatoires gaussiennes*, In: École d'Été de Probabilités de Saint-Flour, IV-1974, Springer, Berlin, pp. 1–96. Lecture Notes in Math., Vol. 480. (French)

Itô, Kiyosi. 1944. *Stochastic integral*, Proc. Imp. Acad. Tokyo **20**, 519–524.

______. 1950. *Stochastic differential equations in a differentiable manifold*, Nagoya Math. J. **1**, 35–47.

______. 1951. *On a formula concerning stochastic differentials*, Nagoya Math. J. **3**, 55–65.

Khintchine, A. Ya. 1933. *Asymptotische Gesetz der Wahrscheinlichkeitsrechnung*, Springer.

Kolmogorov, A. N. 1933. *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin.

Preston, Christopher. 1972. *Continuity properties of some Gaussian processes*, Ann. Math. Statist. **43**, 285–292.

Talagrand, Michel. 1985. *Régularité des processus gaussiens*, C. R. Acad. Sci. Paris Sér. I Math. **301***(7)*, 379–381. (French, with English summary)

______. 1987. *Regularity of Gaussian processes*, Acta Math. **159***(1-2)*, 99–149.

Walsh, John B. 1986. *An Introduction to Stochastic Partial Differential Equations*, In: École d'été de probabilités de Saint-Flour, XIV—1984, Lecture Notes in Math., vol. 1180, Springer, Berlin, pp. 265–439.

Wiener, N. 1923a. *The homogeneous chaos*, Amer. J. Math. **60**, 879–036.

______. 1923b. *Differential space*, J. Math. Phys. **2**, 131–174.

Davar Khoshnevisan: Department of Mathematics, The University of Utah, 155 S. 1400 E. Salt Lake City, UT 84112–0090

E-mail address: davar@math.utah.edu

URL: http://www.math.utah.edu/~davar