# A Primer on Stochastic Partial Differential Equations

Davar Khoshnevisan*

Department of Mathematics, The University of Utah,
Salt Lake City, UT 84112–0090,
Email: davar@math.utah.edu
URL: http://www.math.utah.edu/~davar

**Summary.** These notes form a brief introductory tutorial to elements of Gaussian noise analysis and basic stochastic partial differential equations (SPDEs) in general, and the stochastic heat equation, in particular. The chief aim here is to get to the heart of the matter quickly. We achieve this by studying a few concrete equations only. This chapter provides sufficient preparation for learning more advanced theory from the remainder of this volume.

**Key words:** White noise, Gaussian processes, regularity of processes, martingale measures, stochastic partial differential equations

## 1 What is an SPDE?

Let us consider a perfectly even, infinitesimally-thin wire of length $L$. We lay it down flat, so that we can identify the wire with the interval $[0\,,L]$. Now we apply pressure to the wire 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 we are pressing down toward $y = -\infty$; and $F > 0$ means the opposite is true. Classical physics tells us that the position $u(t\,,x)$ of the wire solves the partial differential equation,

$$\frac{\partial^2 u(t\,,x)}{\partial t^2} = \kappa \frac{\partial^2 u(t\,,x)}{\partial x^2} + F(t\,,x) \qquad (t \geq 0,\ 0 \leq x \leq L), \tag{1}$$

where $\kappa$ is a physical constant that depends only on the linear mass density and the tension of the wire.

* Research supported in part by a grant from the National Science Foundation grant DMS-0404729.

Equation (1) is the so-called *one-dimensional wave equation*. Its solution—via separation of variables and superposition—is a central part of the classical theory of partial differential equations.

We are interested in addressing the question, "*What if F is random noise*"? There is an amusing interpretation, due to Walsh [30], of (1) for random noise $F$: If a guitar string is bombarded by particles of sand, then the induced vibrations of the string are determined by a suitable version of (1).

It turns out that in most cases of interest to us, when $F$ is random noise, Equation (1) does not have 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 mostly on the Saint-Flour lecture notes of Walsh from 1986 [30, Chapters 1–3]. Walsh's lecture notes remain as one of the exciting 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 $\boldsymbol{t}\cdot\boldsymbol{g} := \sum_{j=1}^n t_j g_j$ is a Gaussian random variable for all $\boldsymbol{t} := (t_1, \ldots, t_n) \in \mathbf{R}^n$. It turns out that $\boldsymbol{g}$ is Gaussian if and only if there exist $\boldsymbol{\mu} \in \mathbf{R}^n$ and an $n \times n$, symmetric nonnegative-definite matrix $\boldsymbol{C}$ such that

$$\mathrm{E}\left[\exp\left(i\boldsymbol{t}\cdot\boldsymbol{g}\right)\right] = \exp\left(i\boldsymbol{t}\cdot\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}\cdot\boldsymbol{C}\boldsymbol{t}\right). \tag{2}$$

**Exercise 2.1.** Prove this assertion. It might help to recall that $\boldsymbol{C}$ is *nonnegative definite* if and only if $\boldsymbol{t}\cdot\boldsymbol{C}\boldsymbol{t} \geq 0$ for all $\boldsymbol{t} \in \mathbf{R}^n$. That is, 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 might refer to $G$ as either a *random field*, or a [*stochastic*] *process indexed by* $T$.

We say that $G$ is a *Gaussian process*, or a *Gaussian random field*, if $(G(t_1), \ldots, G(t_k))$ is a $k$-dimensional Gaussian random vector for every $t_1, \ldots, t_k \in T$. The *finite-dimensional distributions* of the process $G$ are the collection of all probabilities obtained as follows:

$$\mu_{t_1,\ldots,t_k}(A_1, \ldots, A_k) := \mathrm{P}\left\{G(t_1) \in A_1, \ldots, G(t_k) \in A_k\right\}, \tag{3}$$

as $A_1, \ldots, A_k$ range over Borel subsets of $\mathbf{R}$ and $k$ ranges 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.

The consistency theorem of Kolmogorov [19] implies 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, $\mu$ is a real-valued function on $T$, whereas $C$ is a real-valued function on $T \times T$.

**Exercise 3.1.** Prove that if $G$ is a Gaussian process with mean function $\mu$ 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 3.2.** Prove that $C$ is *nonnegative definite*. That is, prove that for all $t_1, \ldots, t_k \in T$ and all $z_1, \ldots, z_k \in \mathbf{C}$,

$$\sum_{j=1}^{k} \sum_{l=1}^{k} C(t_j\,, t_l) z_j \overline{z_l} \geq 0. \tag{4}$$

**Exercise 3.3.** Prove that whenever $C : T \times T \to \mathbf{R}$ is nonnegative definite and symmetric,

$$|C(s\,,t)|^2 \leq C(s\,,s) \cdot C(t\,,t) \qquad \text{for all } s, t \in T. \tag{5}$$

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

**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 classical theorem—due in various degrees of generality to Herglotz, Bochner, Minlos, etc.—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 $\mu$ 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 $t_1, \ldots, t_k \geq 0$,

$$\begin{aligned} \sum_{j=1}^k \sum_{l=1}^k \min(t_j\,,t_l) z_j \overline{z_l} &= \sum_{j=1}^k \sum_{l=1}^k z_j \overline{z_l} \int_0^\infty \mathbf{1}_{[0,t_j]}(x) \mathbf{1}_{[0,t_l]}(x)\,dx \\ &= \int_0^\infty \left| \sum_{j=1}^k \mathbf{1}_{[0,t_j]}(x) z_j \right|^2 dx, \end{aligned} \tag{6}$$

which is 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 Bachelier [1].

Brownian motion has the following additional property. *Let $s > 0$ be fixed. Then the process $\{B(t+s) - B(s)\}_{t\geq 0}$ is independent of $\{B(u)\}_{0\leq u\leq s}$.* This is the socalled *Markov property* of Brownian motion, and is not hard to derive. Indeed, thanks to Exercise 3.4 it suffices to prove that for all $t \geq 0$ and $0 \leq u \leq s$,

$$\mathrm{E}[(B(t+s) - B(s))B(u)] = 0. \tag{7}$$

But this is easy to see because

$$\begin{aligned} \mathrm{E}[(B(t+s) - B(s))B(u)] &= \mathrm{Cov}(B(t+s)\,, B(u)) - \mathrm{Cov}(B(s)\,, B(u)) \\ &= \min(t+s\,,u) - \min(s\,,u) \\ &= u - u \\ &= 0. \end{aligned} \tag{8}$$

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 bridge).* The *Brownian bridge* is a mean-zero Gaussian process $\{b(x)\}_{0\leq x\leq 1}$ with covariance,

$$\mathrm{Cov}(b(x)\,, b(y)) := \min(x\,,y) - xy \qquad \text{for all } 0 \leq x, y \leq 1. \tag{9}$$

The next exercise shows that the process $b$ looks locally like a Brownian motion. Note also that $b(0) = b(1) = 0$; this follows because $\mathrm{Var}(b(0)) = \mathrm{Var}(b(1)) = 0$, and motivates the ascription "bridge." The next exercise explains why $b$ is "brownian."

**Exercise 3.8.** Prove that if $B$ is Brownian motion, then $b$ is Brownian bridge, where

$$b(x) := B(x) - xB(1) \qquad \text{for all } 0 \le x \le 1. \tag{10}$$

Also prove that the process $b$ is independent of $B(1)$.

*Example 3.9 (OU process).* Let $B := \{B(t)\}_{t\ge 0}$ denote a $d$-dimensional Brownian motion, and define

$$X(t) := \frac{B(e^t)}{e^{t/2}} \qquad \text{for all } t \ge 0. \tag{11}$$

The coordinate processes $X_1, \ldots, X_d$ are i.i.d. Gaussian processes with mean function $\mu(t) := 0$ and covariance function

$$\begin{aligned} C(s\,,t) &:= \mathrm{E}\left[\frac{B_1\left(e^s\right) B_1\left(e^t\right)}{e^{(s+t)/2}}\right] \\ &= \exp\left(-\tfrac{1}{2}|s-t|\right). \end{aligned} \tag{12}$$

Note that $C(s\,,t)$ depends on $s$ and $t$ only through $|s-t|$. Such processes are called *stationary Gaussian processes*. This particular stationary Gaussian process was predicted in the works of Dutch physicists Leonard S. Ornstein and George E. Uhlenbeck [29], and bears their name as a result. The existence of the Ornstein–Uhlenbeck process was proved rigorously in a landmark paper of Doob [10].

*Example 3.10 (Brownian sheet).* Let $T := \mathbf{R}_+^N := [0\,,\infty)^N$, $\mu(\boldsymbol{t}) := 0$ for all $\boldsymbol{t} \in \mathbf{R}_+^N$, and define

$$C(\boldsymbol{s}\,,\boldsymbol{t}) := \prod_{j=1}^N \min(s_j\,,t_j) \qquad \text{for all } \boldsymbol{s}, \boldsymbol{t} \in \mathbf{R}_+^N. \tag{13}$$

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*. This generalizes Brownian motion to an $N$-parameter random field. 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.

*Example 3.11 (OU 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:

$$X(\boldsymbol{t}) := \frac{B\left(e^{t_1}, \ldots, e^{t_N}\right)}{e^{(t_1+\cdots+t_N)/2}} \qquad \text{for all } \boldsymbol{t} := (t_1\,, \ldots, t_N) \in \mathbf{R}_+^N. \tag{14}$$

This is called the $N$-paramerter *Ornstein–Uhlenbeck sheet*, and generalizes the Ornstein–Uhlenbeck process of Example 3.9.

**Exercise 3.12.** Prove that the Ornstein–Uhlenbeck sheet is a mean-zero, $N$-parameter Gaussian process and its covariance function $C(\boldsymbol{s}\,,\boldsymbol{t})$ depends on $(\boldsymbol{s}\,,\boldsymbol{t})$ only through $|\boldsymbol{s}-\boldsymbol{t}| := \sum_{i=1}^N |s_i - t_i|$.

*Example 3.13 (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 $C(A\,,B) := \lambda^N(A\cap B)$, where $\lambda^N$ denotes the $N$-dimensional Lebesgue measure. Clearly, $C$ is symmetric. It turns out that $C$ is also nonnegative definite (Exercise 3.14 on page 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 3.14.** 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 3.15.** 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 nonrandom, then with probability one,

$$\dot W(A\cup B) = \dot W(A) + \dot W(B) - \dot W(A\cap B). \tag{15}$$

**Exercise 3.16.** 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,

$$\lim_{n\to\infty}\sum_{j=0}^{2^n-1}\left|\dot W\left(\left[\frac{j-1}{2^n}\,,\,\frac{j}{2^n}\right]\right)\right|^2 = 1. \tag{16}$$

Use this to prove that with probability one,

$$\lim_{n\to\infty}\sum_{j=0}^{2^n-1}\left|\dot W\left(\left[\frac{j-1}{2^n}\,,\,\frac{j}{2^n}\right]\right)\right| = \infty. \tag{17}$$

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 3.17 (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 $c_1,\ldots,c_k\in\mathbf{R}$,

$$\dot W\left(\sum_{j=1}^k c_j\mathbf{1}_{A_j}\right) := \sum_{j=1}^k c_j\dot W(A_j). \tag{18}$$

The random variables $\dot{W}(A_1), \ldots, \dot{W}(A_k)$ are independent, thanks to Exercise 3.15. Therefore,

$$\begin{aligned}\left\| \dot{W}\left(\sum_{j=1}^k c_j \mathbf{1}_{A_j}\right)\right\|_{L^2(\mathrm{P})}^2 &= \sum_{j=1}^k c_j^2 |A_j| \\ &= \left\| \sum_{j=1}^k c_j \mathbf{1}_{A_j}\right\|_{L^2(\mathbf{R}^N)}^2 .\end{aligned} \tag{19}$$

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 (19) 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

$$\left\| \dot{W}(h) \right\|_{L^2(\mathrm{P})} = \|h\|_{L^2(\mathbf{R}^N)}. \tag{20}$$

That is, $\dot{W} : L^2(\mathbf{R}^N) \to L^2(\mathrm{P})$ is an isometry; (20) is called *Wiener's isometry* [32]. [Note that we now know how to construct the stochastic integral $\int h\, dW$ only if $h \in L^2(\mathbf{R}^N)$ is *nonrandom.*] The process $\{\dot{W}(h)\}_{h \in L^2(\mathbf{R}^N)}$ is called the *isonormal process* [11]. 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 3.18.** Prove that for all [nonrandom] $h, g \in L^2(\mathbf{R}^N)$ and $a, b \in \mathbf{R}$,

$$\int (ah + bg)\, dW = a \int h\, dW + b \int h\, dW, \tag{21}$$

almost surely.

**Exercise 3.19.** 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 complete orthonormal system in $L^2(\mathrm{P})$. In particular, for all Gaussian random variables $Z \in L^2(\mathrm{P})$ that are measurable with respect to the white noise,

$$Z = \sum_{j=1}^\infty a_j \dot{W}(h_j) \quad \text{almost surely,} \quad \text{with} \quad a_j := \mathrm{Cov}\left(Z\,, \dot{W}(h_j)\right), \tag{22}$$

and the infinite sum converges in $L^2(\mathrm{P})$. This permits one possible entry into the "Malliavin calculus." For this, and much more, see the course by D. Nualart in this volume.

**Exercise 3.20.** Verify that (18) is legitimate. That is, prove that if $B_1, \ldots, B_\ell \in \mathscr{B}(\mathbf{R}^N)$ are disjoint, then

$$\dot{W}\left(\sum_{j=1}^{k} c_j \mathbf{1}_{A_j}\right) = \dot{W}\left(\sum_{l=1}^{\ell} d_l \mathbf{1}_{B_l}\right) \quad \text{almost surely,} \tag{23}$$

provided that $d_1, \ldots, d_\ell \in \mathbf{R}$ satisfy $\sum_{j=1}^{k} c_j \mathbf{1}_{A_j} = \sum_{l=1}^{\ell} d_l \mathbf{1}_{B_l}$.

## 4 Regularity of random processes

Our construction of Gaussian processes is very general. This generality makes our construction both useful and useless. It is useful because we can make sense of fundamental mathematical 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 problem has to do with the structure of Kolmogorov's existence theorem. But instead of discussing this technical subject directly, let us consider a simple example first.

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

$$B'(t) := \begin{cases} B(t) & \text{if } t \neq U, \\ 5000 & \text{if } t = U. \end{cases} \tag{24}$$

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.

**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*

$$\mathrm{P}\left\{X'(t) = X(t)\right\} = 1 \quad \textit{for all } t \in T. \tag{25}$$

**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 $\mathrm{P}\{X' = X\} = 0$.

A remarkable theorem of Wiener [31] states that we can always find a continuous modification of a Brownian motion. According to the previous exercise, this modification is itself a Brownian motion. Thus, a *Wiener process* is a Brownian motion $B$ such that the random function $t \mapsto B(t)$ is continuous; it is also some times known as *standard Brownian motion.*

### 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 $\mathscr{F}f$. We normalize the Fourier transform as follows:

$$(\mathscr{F}f)(z) := \int_{-\infty}^{\infty} e^{izx} f(x)\, dx \quad \text{for all } z \in \mathbf{R}. \tag{26}$$

Let $\mathscr{W}(\mathbf{R})$ denote the collection of all $f \in L^1(\mathbf{R})$ such that $\hat{f} \in L^1(\mathbf{R})$ as well. The space $\mathscr{W}$ is the so-called Wiener algebra on $\mathbf{R}$. If $f \in \mathscr{W}(\mathbf{R})$, then we can proceed, intentionally carelessly, and use the inversion formula to arrive at the following:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-izx} (\mathscr{F}f)(z)\, dz. \tag{27}$$

It follows readily from this and the dominated convergence theorem that $f$ is uniformly continuous. But this cannot be so! In order to see why, let us consider the function

$$g(x) = \begin{cases} f(x) & \text{if } x \neq 0, \\ f(0) + 1 & \text{if } x = 0. \end{cases} \tag{28}$$

If $f$ were a continuous function, 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 *a priori* (27) holds only for almost all $x \in \mathbf{R}$. Therefore,

$$x \mapsto \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-izx} (\mathscr{F}f)(z)\, dz \tag{29}$$

defines a "modification" of $f$ which happens to be uniformly continuous. That is, we have proven that every $f \in \mathscr{W}(\mathbf{R})$ has a uniformly-continuous modification.

### 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 [11; 12; 26; 27; 28]. 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.** *Suppose $\{X(\boldsymbol{t})\}_{\boldsymbol{t}\in T}$ is a stochastic process indexed by a compact cube $T := [a_1\,,b_1]\times\cdots\times[a_N\,,b_N]\subset\mathbf{R}^N$. Suppose also that there exist constants $C>0$, $p>0$, and $\gamma>N$ such that uniformly for all $\boldsymbol{s},\boldsymbol{t}\in T$,*

$$\mathrm{E}\left(|X(\boldsymbol{t})-X(\boldsymbol{s})|^p\right)\le C|\boldsymbol{t}-\boldsymbol{s}|^\gamma. \tag{30}$$

*Then $X$ has a continuous modification $\bar{X}$. Moreover, if $0\le\theta<(\gamma-N)/p$ then*

$$\left\|\sup_{\boldsymbol{s}\neq\boldsymbol{t}}\frac{|\bar{X}(\boldsymbol{s})-\bar{X}(\boldsymbol{t})|}{|\boldsymbol{s}-\boldsymbol{t}|^\theta}\right\|_{L^p(\mathrm{P})}<\infty. \tag{31}$$

*Remark 4.4.* Here, $|\boldsymbol{x}|$ could be any of the usual Euclidean $\ell^p$ norms for $x\in\mathbf{R}^k$. That is,

$$\begin{aligned}
|\boldsymbol{x}| &:= \max\left(|x_1|\,,\ldots,|x_k|\right);\\
|\boldsymbol{x}| &:= \left(|x_1|^p+\cdots+|x_k|^p\right)^{1/p} \qquad \text{for } p\ge 1;\\
|\boldsymbol{x}| &:= |x_1|^p+\cdots+|x_k|^p \qquad \text{for } 0<p<1.
\end{aligned} \tag{32}$$

*Proof.* We prove Theorem 4.3 in the case that $N=1$ and $T:=[0\,,1]$. The general case is not much more difficult to prove, but requires introducing further notation. Also, we extend the domain of the process by setting

$$X(t) := \begin{cases} X(0) & \text{if } t<0,\\ \mathrm{X}(1) & \text{if } t>1.\end{cases} \tag{33}$$

First we introduce some notation: For every integer $n\ge 0$ we define $\mathscr{D}_n := \{j2^{-n}\,:\,0\le j<2^n\}$ to be the collection of all dyadic points in $[0\,,1)$. The totality of all dyadic points is denoted by $\mathscr{D}_\infty := \cup_{n=0}^\infty\mathscr{D}_n$.

Suppose $n>k\ge 1$, and consider $u,v\in\mathscr{D}_n$ that are within $2^{-k}$ of one another. We can find two sequences of points $u_k,\ldots,u_n$ and $v_k,\ldots,v_n$ with the following properties:

1. $u_j,v_j\in\mathscr{D}_j$ for all $j=k,\ldots,n$;
2. $|u_{j+1}-u_j|\le 2^{-j-1}$ for all $j=k,\ldots,n$;
3. $|v_{j+1}-v_j|\le 2^{-j-1}$ for all $j=k,\ldots,n$;
4. $u_n=u$, $v_n=v$, and $u_k=v_k$.

(Draw a picture.) Because $|X(u)-X(u_k)|\le\sum_{j=k}^{n-1}|X(u_{j+1})-X(u_j)|$, this yields

$$|X(u)-X(u_k)|\le\sum_{j=k}^\infty\max_{s\in\mathscr{D}_{j+1}}\max_{t\in B(s,2^{-j-1})\cap\mathscr{D}_j}|X(s)-X(t)|, \tag{34}$$

where $B(x\,,r):=[x-r\,,x+r]$. The right-most term does not depend on $u$, nor on the sequences $\{u_j\}_{j=k}^n$ and $\{v_j\}_{j=k}^n$. Moreover, $|X(v)-X(v_k)|=$

$|X(v)-X(u_k)|$ is bounded above by the same quantity. Hence, by the triangle inequality,

$$|X(u)-X(v)| \leq 2\sum_{j=k}^{\infty} \max_{s\in\mathscr{D}_{j+1}} \max_{t\in B(s,2^{-j-1})\cap\mathscr{D}_j} |X(s)-X(t)|, \tag{35}$$

uniformly for all $u,v\in\mathscr{D}_n$ that are within $2^{-k}$ of one another. Because its right-hand side is independent of $n$, the preceding holds uniformly for all $u,v\in\mathscr{D}_\infty$ that are within distance $2^{-k}$ of one another. This and the Minkowski inequality together imply that

$$\begin{aligned}\Bigg\| \sup_{\substack{u,v\in\mathscr{D}_\infty\\ |u-v|\leq 2^{-k}}} &|X(u)-X(v)| \Bigg\|_{L^p(\mathrm{P})} \\ &\leq 2\sum_{j=k}^{\infty} \left\| \max_{s\in\mathscr{D}_{j+1}} \max_{t\in B(s,2^{-j-1})\cap\mathscr{D}_j} |X(s)-X(t)| \right\|_{L^p(\mathrm{P})}.\end{aligned} \tag{36}$$

A crude bound yields

$$\begin{aligned}\mathrm{E}\Bigg( \max_{s\in\mathscr{D}_{j+1}} &\max_{t\in B(s,2^{-j-1})\cap\mathscr{D}_j} |X(s)-X(t)|^p \Bigg) \\ &\leq \sum_{s\in\mathscr{D}_{j+1}} \sum_{t\in B(s,2^{-j-1})\cap\mathscr{D}_j} \mathrm{E}\left(|X(s)-X(t)|^p\right) \\ &\leq C \sum_{s\in\mathscr{D}_{j+1}} \sum_{t\in B(s,2^{-j-1})\cap\mathscr{D}_j} |s-t|^\gamma,\end{aligned} \tag{37}$$

thanks to Condition (30) of the theorem. For the range in question: $|s-t|^\gamma \leq 2^{-(j+1)\gamma}$; the sum over $t$ then contributes a factor of 2; and the sum over $s$ yields a factor of $2^{j+1}$. Therefore,

$$\mathrm{E}\left( \max_{s\in\mathscr{D}_{j+1}} \max_{t\in B(s,2^{-j-1})\cap\mathscr{D}_j} |X(s)-X(t)|^p \right) \leq \frac{2^{2-\gamma}C}{2^{j(\gamma-1)}}. \tag{38}$$

We can plug this into (36) to deduce that

$$\left\| \sup_{\substack{u,v\in\mathscr{D}_\infty\\ |u-v|\leq 2^{-k}}} |X(u)-X(v)| \right\|_{L^p(\mathrm{P})} \leq \frac{\tilde{C}}{2^{k\gamma/p}}, \tag{39}$$

where

$$\tilde{C} := \frac{2^{(2-\gamma+p)/p}C^{1/p}}{1-2^{-(\gamma-1)/p}}. \tag{40}$$

Now let us define

$$\bar{X}(s) := \limsup X(t), \tag{41}$$

where the lim sup is taken over all $t \in \mathscr{D}_\infty$ such that $t \to s$. Because $\bar{X}(s) = X(s)$ for all $s \in \mathscr{D}_\infty$, Equation (39) continues to hold, even if we replace $X$ by $\bar{X}$. In that case, we can also replace the condition "$s, t \in \mathscr{D}_\infty$" with "$s, t \in [0\,,1]$" at no extra cost. This proves, among other things, that $\bar{X}$ is a.s. continuous [Borel–Cantelli lemma].

It is not hard to check that $\bar{X}$ is a modification of $X$ because: (i) $X$ and $\bar{X}$ agree on $\mathscr{D}_\infty$; (ii) $X$ is continuous in probability[2] by (30); and (iii) $\bar{X}$ is continuous a.s., as we just proved.

It remains to verify (31). For $\theta$ as given, (39) implies that for all integers $k \geq 1$,

$$\left\| \sup_{\substack{0 \leq s \neq t \leq 1: \\ 2^{-k} < k \leq 2^{-k+1}}} \frac{|\bar{X}(s) - \bar{X}(t)|}{|s-t|^\theta} \right\|_{L^p(\mathrm{P})} \leq \frac{\tilde{C}}{2^{k(\gamma-\theta)/p}}. \tag{42}$$

Sum both sides of this inequality from $k = 1$ to infinity to deduce (31), and hence the theorem. □

**Exercise 4.5.** Suppose the conditions of Theorem 4.3 are met, but we have the following in place of (30):

$$\mathrm{E}\left(|X(\boldsymbol{t}) - X(\boldsymbol{s})|^p\right) \leq h\left(|\boldsymbol{t} - \boldsymbol{s}|\right), \tag{43}$$

where $h : [0\,,\infty) \to \mathbf{R}_+$ is continuous and increasing, and $h(0) = 0$. Prove that $X$ has a continuous modification provided that

$$\int_0^\eta \frac{h(r)}{r^{1+N}}\,dr < \infty \qquad \text{for some } \eta > 0. \tag{44}$$

**Definition 4.6 (Hölder continuity).** *A function $f : \mathbf{R}^N \to \mathbf{R}$ is said to be* globally Hölder continuous *with index $\alpha$ if there exists a constant $A$ such that for all $x, y \in \mathbf{R}^N$,*

$$|f(x) - f(y)| \leq A|x-y|^\alpha. \tag{45}$$

*It is said to be [locally]* Hölder continuous *with index $\alpha$ if for all compact sets $K \subset \mathbf{R}^N$ there exists a constant $A_K$ such that*

$$|f(x) - f(y)| \leq A_K|x-y|^\alpha \qquad \textit{for all } x, y \in K. \tag{46}$$

**Exercise 4.7.** Suppose $\{X(\boldsymbol{t})\}_{\boldsymbol{t}\in T}$ is a process indexed by a compact set $T \subset \mathbf{R}^N$ that satisfies (30) 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 $\alpha$.

[2] This means that $X(s)$ converges to $X(t)$ in probability as $s \to t$.

**Exercise 4.8.** Suppose $\{X(\boldsymbol{t})\}_{\boldsymbol{t}\in\mathbf{R}^N}$ 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

$$\mathrm{E}\left(|X(\boldsymbol{s}) - X(\boldsymbol{t})|^{p_T}\right) \le C_T |\boldsymbol{s} - \boldsymbol{t}|^{\gamma}, \qquad \text{for all } \boldsymbol{s}, \boldsymbol{t} \in T. \tag{47}$$

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

**Exercise 4.9 (Regularity of Gaussian processes).** Suppose $\{X(\boldsymbol{t})\}_{\boldsymbol{t}\in T}$ is a Gaussian random field, and $T \subseteq \mathbf{R}^N$ for some $N \ge 1$. Then, check that for all $p > 0$,

$$\mathrm{E}\left(|X(\boldsymbol{t}) - X(\boldsymbol{s})|^p\right) = c_p \left[\mathrm{E}\left(|X(\boldsymbol{t}) - X(\boldsymbol{s})|^2\right)\right]^{p/2}, \tag{48}$$

where

$$c_p := \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} |x|^p e^{-x^2/2}\, dx = \frac{2^{p/2}}{\pi^{1/2}} \Gamma\left(\frac{p+1}{2}\right). \tag{49}$$

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

$$\mathrm{E}\left(|X(\boldsymbol{t}) - X(\boldsymbol{s})|^2\right) \le A(K) |\boldsymbol{t} - \boldsymbol{s}|^{\varepsilon} \qquad \text{for all } \boldsymbol{t}, \boldsymbol{s} \in K. \tag{50}$$

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

*Example 4.10 (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 $\alpha < \frac{1}{2}$. This is due to Wiener [31].

**Warning:** This is not true for $\alpha = \frac{1}{2}$. Let $B$ denote the modification as well. [This should not be confusing.] Then, "the law of the iterated logarithm" of [18] asserts that

$$\mathrm{P}\left\{\limsup_{t \downarrow s} \frac{|B(t) - B(s)|}{\left(2(t-s) \ln|\ln(t-s)|\right)^{1/2}} = 1\right\} = 1 \qquad \text{for all } s > 0. \tag{51}$$

In particular, for all $s > 0$,

$$\mathrm{P}\left\{\limsup_{t \downarrow s} \frac{|B(t) - B(s)|}{|t - s|^{1/2}} = \infty\right\} = 1. \tag{52}$$

Thus, $B$ is not Hölder continuous of order $\frac{1}{2}$ at $s = 0$, for instance.

**Exercise 4.11.** Let $B$ denote $N$-parameter Brownian sheet. Prove that $B$ has a modification which is [locally] Hölder continuous with any nonrandom index $\alpha \in (0\,,\frac{1}{2})$. This generalized Wiener's theorem on Brownian motion.

**Exercise 4.12.** Let $B$ be a continuous Brownian motion, and define $\mathscr{H}_t$ to be the smallest sigma algebra that makes the random variables $\{B(s)\}_{s\in[0,t]}$ measurable. Then prove that the event in (51), whose probability is one, is measurable with respect to $\vee_{t\geq 0}\mathscr{H}_t$. Do the same for the event in (52).

Theorem 4.3 and the subsequent exercises all deal with distances on $\mathbf{R}^N$ that are based on norms. We will need a version based on another distance as well. This we state—without proof—in the case that $N = 2$.

Choose and fix some $p \in (0\,,1]$ and an integer $1 \leq k \leq 1/p$, and define for all $u, v, s, t \in [0\,,1]$,

$$|(s\,,t) - (u\,,v)| := |s-u|^p + |t-v|^{kp}. \tag{53}$$

This defines a distance on $[0\,,1]^2$, but it is inhomogeneous, when $k > 1$, in the sense that it scales differently in different directions. The following is essentially 1.4.1 of Kunita [23, p. 31]; see also Corollary A.3 of [6]. I omit the proof.

**Theorem 4.13.** *Let $\{Y(s\,,t)\}_{s,t\in[0,1)^2}$ be a 2-parameter stochastic process taking value in $\mathbf{R}$. Suppose that there exist $C, p > 1$ and $\gamma > (k+1)/k$ such that for all $s, t, u, v \in [0\,,1)$,*

$$\|Y(s\,,t) - Y(u\,,v)\|_{L^p(\mathrm{P})} \leq C\,|(s\,,t) - (u\,,v)|^{\gamma}\,. \tag{54}$$

*Then, $Y$ has a Hölder-continuous modification $\bar{Y}$ that satisfies the following for every $\theta \geq 0$ which satisfies $k\gamma - (k+1) - k\theta > 0$:*

$$\left\| \sup_{(s,t)\neq(u,v)} \frac{|Y(s\,,t) - Y(u\,,v)|}{|(s\,,t) - (u\,,v)|^{\theta}} \right\|_{L^p(\mathrm{P})} < \infty. \tag{55}$$

## 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 [nonrandom] sets $A_1, A_2, \ldots \in \mathscr{B}(\mathbf{R}^N)$,

$$\mathrm{P}\left\{\dot{W}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \dot{W}(A_i)\right\} = 1, \tag{56}$$

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

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

**Lemma 5.3.** $\{W_t(A)\}_{t\geq 0, A\in\mathscr{B}(\mathbf{R}^{N-1})}$ *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 $A \in \mathscr{B}(\mathbf{R}^{N-1})$, $\{W_t(A)\}_{t\geq 0}$ 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).

Equation (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 $t \geq s \geq u \geq 0$,

$$\begin{aligned}
&\mathrm{E}\big[\,(W_t(A) - W_s(A))\, W_u(A)\big] \\
&\quad= \mathrm{E}\left[\Big(\dot{W}\left([0\,,t] \times A\right) - \dot{W}\left([0\,,s] \times A\right)\Big)\, \dot{W}\left([0\,,u] \times A\right)\right] \\
&\quad= \min(t\,,u)|A| - \min(s\,,u)|A| = 0.
\end{aligned} \tag{57}$$

Therefore, $W_t(A) - W_s(A)$ is independent of $\mathscr{F}_s$ (Exercise 3.4, page 7). As a result, with probability one,

$$\begin{aligned}\mathrm{E}\left[W_t(A) \mid \mathscr{F}_s\right] &= \mathrm{E}\left[W_t(A) - W_s(A) \mid \mathscr{F}_s\right] + W_s(A) \\ &= \mathrm{E}\left[W_t(A) - W_s(A)\right] + W_s(A) \\ &= W_s(A).\end{aligned} \tag{58}$$

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 $\{cW_t(A)\}_{t\geq 0}$ 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 $\{W_t(h)\}_{t\geq 0}$ is a continuous martingale with quadratic variation

$$\langle W_\bullet(h)\, ,\, W_\bullet(h)\rangle_t = t\int_{\mathbf{R}^{N-1}} h^2(x)\, dx. \tag{59}$$

It might help to recall that if $\{Z_t\}_{t\geq 0}$ is a continuous $L^2(\mathrm{P})$-martingale, then its quadratic variation is uniquely defined as the continuous increasing process $\{\langle Z\, , Z\rangle_t\}_{t\geq 0}$ 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 $t \geq 0$ 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 $\mathscr{F} := \{\mathscr{F}_t\}_{t\geq 0}$ be a filtration of sigma-algebras. We assume that $\mathscr{F}$ is *right-continuous*; i.e.,

$$\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s \qquad \text{for all } t \geq 0. \tag{60}$$

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

**Definition 5.6 (Martingale measures).** *A process $\{M_t(A)\}_{t\geq 0, A\in\mathscr{B}(\mathbf{R}^n)}$ 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 $A \in \mathscr{B}(\mathbf{R}^n)$, $\{M_t(A)\}_{t\geq 0}$ 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 $\mu$ be a sigma-finite $L^2(\mathrm{P})$-valued signed measure on $\mathscr{B}(\mathbf{R}^n)$, and $\mathscr{F} := \{\mathscr{F}_t\}_{t\geq 0}$ a right-continuous filtration. Define $\mu_t(A) := \mathrm{E}[\mu(A) \,|\, \mathscr{F}_t]$ for all $t \geq 0$ and $A \in \mathscr{B}(\mathbf{R}^n)$. Then prove that $\{\mu_t(A)\}_{t\geq 0, A\in\mathscr{B}(\mathbf{R}^n)}$ is a martingale measure.

**Exercise 5.9.** Let $\{M_t(A)\}$ be a martingale measure. Prove that for all $T \geq t \geq 0$, $M_t(A) = \mathrm{E}[M_T(A) \,|\, \mathscr{F}_t]$ 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 [30, 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*

$$f(x\,,t\,,\omega) = X(\omega)\mathbf{1}_{(a,b]}(t)\mathbf{1}_A(x), \tag{61}$$

*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 (61), then we define the stochastic-integral process of $f$ as

$$(f\cdot M)_t(B)(\omega) := X(\omega)\left[M_{t\wedge b}(A\cap B) - M_{t\wedge a}(A\cap B)\right](\omega). \tag{62}$$

**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 $f = c_1f_1 + \cdots + c_kf_k$ where $c_1, \ldots, c_k \in \mathbf{R}$ and $f_1, \ldots, f_k$ are elementary. We can then define

$$(f\cdot M)_t(B) := \sum_{j=1}^k c_j(f_j\cdot M)_t(B). \tag{63}$$

**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*

$$\overline{Q}_t(A\,,B) := \langle M_\bullet(A)\,, M_\bullet(B)\rangle_t, \qquad \textit{for all } t \geq 0,\ A, B \in \mathscr{B}(\mathbf{R}^n). \tag{64}$$

**Exercise 5.16.** Prove that:

1. $\overline{Q}_t(A\,,B) = \overline{Q}_t(B\,,A)$ almost surely;
2. If $B \cap C = \varnothing$ then $\overline{Q}_t(A\,,B\cup C) = \overline{Q}_t(A\,,B) + \overline{Q}_t(A\,,C)$ almost surely;
3. $|\overline{Q}_t(A\,,B)|^2 \leq \overline{Q}_t(A\,,A)\overline{Q}_t(B\,,B)$ almost surely; and
4. $t \mapsto \overline{Q}_t(A\,,A)$ 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 $t \geq 0$ and $A \in \mathscr{B}(\mathbf{R}^{N-1})$. Verify that the quadratic functional of this martingale measure is described by $\overline{Q}_t(A\,,B) := t\lambda^{N-1}(A\cap B)$, where $\lambda^k$ denotes the Lebesgue measure on $\mathbf{R}^k$.

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

$$Q\left(A\,,B\,;(s\,,t]\right) := \overline{Q}_t(A\,,B) - \overline{Q}_s(A\,,B). \tag{65}$$

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

$$Q\left(\bigcup_{i=1}^n \left(A_i \times B_i \times (s_i\,,t_i]\right)\right) := \sum_{i=1}^n Q\left(A_i\,,B_i\,;(s_i\,,t_i]\right). \tag{66}$$

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" [30]. 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,*

$$\mathrm{E}\left[\left((f\cdot M)_t(B)\right)^2\right] = \mathrm{E}\left[\iiint_{B\times B\times(0,t]} f(x\,,t)f(y\,,t)\,Q(dx\,dy\,dt)\right]. \tag{67}$$

*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 (61). Then,

$$\begin{aligned}
&\mathrm{E}\left[(f\cdot M)_t^2(B)\right]\\
&=\mathrm{E}\left[X^2\left(M_{t\wedge b}(A\cap B)-M_{t\wedge a}(A\cap B)\right)^2\right]\\
&=\mathrm{E}\left[X^2M_{t\wedge b}^2(A\cap B)\right]-2\mathrm{E}\left[X^2M_{t\wedge b}(A\cap B)M_{t\wedge a}(A\cap B)\right]\\
&\qquad\qquad+\mathrm{E}\left[X^2M_{t\wedge a}^2(A\cap B)\right].
\end{aligned}\tag{68}$$

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

$$\begin{aligned}
&\mathrm{E}\left[X^2\left(M_{t\wedge b}^2(A\cap B)-\langle M(A\cap B)\,,M(A\cap B)\rangle_{t\wedge b}\right)\right]\\
&\qquad=\mathrm{E}\left[X^2\left(M_{t\wedge a}^2(A\cap B)-\langle M(A\cap B)\,,M(A\cap B)\rangle_{t\wedge a}\right)\right].
\end{aligned}\tag{69}$$

Similarly,

$$\begin{aligned}
&\mathrm{E}\left[X^2\left(M_{t\wedge b}(A\cap B)M_{t\wedge a}(A\cap B)-\langle M(A\cap B)\,,M(A\cap B)\rangle_{t\wedge a}\right)\right]\\
&\qquad=\mathrm{E}\left[X^2\left(M_{t\wedge a}^2(A\cap B)-\langle M(A\cap B)\,,M(A\cap B)\rangle_{t\wedge a}\right)\right].
\end{aligned}\tag{70}$$

Combine to deduce the result in the case that $f$ has form (61).

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 $t>0$,*
$$|Q(A\times B\times(0\,,t])|\leq K(A\times B\times(0\,,t])\qquad a.s.$$

*[As usual, we drop the dependence on $\omega$.] 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 $K(A \times B \times C) := \lambda^{N-1}(A \cap B)\lambda^1(C)$, where $\lambda^k$ 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*

$$\begin{aligned} Q_{f\cdot M}(dx\,dy\,dt) &= f(x\,,t)f(y\,,t)\,Q_M(dx\,dy\,dt), \\ K_{f\cdot M}(dx\,dy\,dt) &= |f(x\,,t)f(y\,,t)|\,K_M(dx\,dy\,dt). \end{aligned} \tag{71}$$

*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 (61), and note that for all $t \geq 0$ and $B, C \in \mathscr{B}(\mathbf{R}^n)$,

$$\begin{aligned} &(f \cdot M)_t(B)(f \cdot M)_t(C) \\ &\quad = X^2\left[M_{t\wedge b}(A \cap B) - M_{t\wedge a}(A \cap B)\right] \\ &\qquad\qquad \times \left[M_{t\wedge b}(A \cap C) - M_{t\wedge a}(A \cap C)\right] \\ &\quad = \text{martingale } + X^2\left\langle M(A \cap B)\,, M(A \cap C)\right\rangle_{t\wedge b} \\ &\qquad\qquad - X^2\left\langle M(A \cap B)\,, M(A \cap C)\right\rangle_{t\wedge a} \\ &\quad = \text{martingale } + X^2 Q_M\left((A \cap B) \times (A \cap B) \times (s\,,t]\right) \\ &\quad = \text{martingale } + \iiint_{B\times C\times(0\,,t]} f(x\,,s)f(y\,,s)\,Q_M(dx\,dy\,ds). \end{aligned} \tag{72}$$

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 $\|f\|_M$, for all predictable function $f$, via

$$\|f\|_M^2 := \mathrm{E}\left[\iiint_{\mathbf{R}^n\times\mathbf{R}^n\times(0,T]} |f(x\,,t)f(y\,,t)|\,K_M(dx\,dy\,dt)\right]. \tag{73}$$

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

**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 [30, p. 293, Proposition 2.3].

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

Note from Proposition 5.18 that

$$\mathrm{E}\left[(f\cdot M)_t^2(B)\right] \leq \|f\|_M^2 \quad \text{for all } t\in(0\,,T],\ f\in\mathscr{S},\ B\in\mathscr{B}(\mathbf{R}^n). \tag{74}$$

Consequently, if $\{f_m\}_{m=1}^\infty$ is a Cauchy sequence in $(\mathscr{S}\,,\|\cdot\|_M)$ then the sequence $\{(f_m\cdot M)_t(B)\}_{m=1}^\infty$ 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 (71). Moreover, for all $t\in(0\,,T]$ and $A,B\in\mathscr{B}(\mathbf{R}^n)$,*

$$\begin{aligned}\big\langle (f\cdot M)(A)\,,(f\cdot M)(B)\big\rangle_t &= \iiint_{A\times B\times(0,t]} f(x\,,s)f(y\,,s)\,Q_M(dx\,dy\,ds),\\ \mathrm{E}\left[(f\cdot M)_t^2(B)\right] &\leq \|f\|_M^2.\end{aligned} \tag{75}$$

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

**Theorem 5.27 (Burkholder's inequality [3]).** *For all $p\geq 2$ there exists $c_p\in(0\,,\infty)$ such that for all predictable $f$ and all $t>0$,*

$$\begin{aligned}&\mathrm{E}\left[\,|(f\cdot M)_t(B)|^p\,\right]\\ &\quad\leq c_p\mathrm{E}\left[\left(\iiint_{\mathbf{R}^n\times\mathbf{R}^n\times(0,T]} |f(x\,,t)f(y\,,t)|\,K_M(dx\,dy\,dt)\right)^{p/2}\right].\end{aligned} \tag{76}$$

*Proof (Special Case).* It is enough to prove that if $\{N_t\}_{t\geq 0}$ is a martingale with $N_0:=0$ and quadratic variation $\langle N\,,N\rangle_t$ at time $t$, then

$$\|N_t\|_{L^p(\mathrm{P})}^p \leq c_p\|\langle N\,,N\rangle_t\|_{L^{p/2}(\mathrm{P})}^{p/2}, \tag{77}$$

but this is precisely the celebrated Burkholder inequality [3]. Here is why it is true in the case that $N$ is a *bounded* and continuous martingale. Recall Itô's formula [15; 16; 17]: For all $f$ that is $C^2$ a.e.,

$$f(N_t) = f(0) + \int_0^t f'(N_s)\, dN_s + \frac{1}{2}\int_0^t f''(N_s)\, d\langle N\,, N\rangle_s. \tag{78}$$

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

$$|N_t|^p = \frac{p(p-1)}{2}\int_0^t |N_s|^{p-2}\, d\langle N\,, N\rangle_s + \text{mean-zero martingale}. \tag{79}$$

Take expectations to find that

$$\mathrm{E}\left(|N_t|^p\right) \le \frac{p(p-1)}{2}\mathrm{E}\left(\sup_{0\le u\le t} |N_u|^{p-2}\langle N\,, N\rangle_t\right). \tag{80}$$

Because $|N_t|^p$ is a submartingale, Doob's maximal inequality asserts that

$$\mathrm{E}\left(\sup_{0\le u\le t}|N_u|^p\right) \le \left(\frac{p}{p-1}\right)^p \mathrm{E}\left(|N_t|^p\right). \tag{81}$$

Therefore, $\phi_p(t) := \mathrm{E}(\sup_{0\le u\le t}|N_u|^p)$ satisfies

$$\begin{aligned}\phi_p(t) &\le \frac{p(p-1)}{2}\left(\frac{p}{p-1}\right)^p \mathrm{E}\left(\sup_{0\le u\le t}|N_u|^{p-2}\langle N\,, N\rangle_t\right)\\ &:= a_p\mathrm{E}\left(\sup_{0\le u\le t}|N_u|^{p-2}\langle N\,, N\rangle_t\right).\end{aligned} \tag{82}$$

Apply Hölder's inequality to find that

$$\phi_p(t) \le a_p\left(\phi_p(t)\right)^{(p-2)/p}\left(\mathrm{E}\left[\langle N\,, N\rangle_t^{p/2}\right]\right)^{2/p}. \tag{83}$$

We can solve this inequality for $\phi_p(t)$ to finish. □

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

$$\mathrm{E}\left(\sup_{0\le u\le T}|N_u|^p\right) \le c_p\mathrm{E}\left(\langle N\,, N\rangle_T^{p/2}\right). \tag{84}$$

In addition, prove that 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 $p \ge 2$ there exists $c_p' \in (0\,,\infty)$ such that for all bounded stopping times $T$,

$$\mathrm{E}\left(\langle N\,, N\rangle_T^{p/2}\right) \le c_p'\mathrm{E}\left(\sup_{0\le u\le T}|N_u|^p\right). \tag{85}$$

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

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

$$(f \cdot M)_t(A) := \iint_{A\times(0,t]} f\, dM := \iint_{A\times(0,t]} f(x\,,s)\, M(dx\, ds). \tag{86}$$

[N.B.: The last $f(x\,,s)$ is actually $f(x\,,s\,,\omega)$, but we have dropped the $\omega$ 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 the following expectation is finite:*

$$\int\cdots\int_{\Omega\times\mathbf{R}^n\times\mathbf{R}^n\times[0,T]\times A} |f(x\,,t\,,\omega\,,u) f(y\,,t\,,\omega\,,u)|\, K(dx\, dy\, dt)\, \mu(du)\, \mathrm{P}(d\omega). \tag{87}$$

*Then almost surely,*

$$\begin{aligned}\int_A \left( \iint_{\mathbf{R}^n\times[0,t]} f(x\,,s\,,\bullet\,,u)\, M(dx\, ds) \right) \mu(du) \\ = \iint_{\mathbf{R}^n\times[0,t]} \left( \int_A f(x\,,s\,,\bullet\,,u)\, \mu(du) \right) M(dx\, ds).\end{aligned} \tag{88}$$

It suffices to prove this for elementary functions of the form (61). You can do this yourself, or consult the lecture notes of Walsh [30, p. 297].

## 6 A nonlinear heat equation

We are ready to try and study a class of nonlinear elliptic SPDEs that is an example of the equations studied by Baklan [2], Daleckiĭ [7], Dawson [8; 9], Pardoux [24; 25], Krylov and Rozovski [20; 21; 22], and Funaki [13; 14]. It is possible to adapt the arguments to study hyperbolic SPDEs as well. For an introductory example see the paper by Cabaña [4]. The second chapter, by R. C. Dalang, of this volume contains more advanced recent results on hyperbolic SPDEs.

Let $L > 0$ be fixed, and consider

$$\left|\begin{aligned} &\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, && t > 0,\ x \in [0\,,L], \\ &\frac{\partial u}{\partial x}(0\,,t) = \frac{\partial u}{\partial x} u(L\,,t) = 0, && t > 0, \\ &u(x\,,0) = u_0(x), && x \in [0\,,L], \end{aligned}\right. \tag{89}$$

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 nonrandom, measurable, and bounded function. As regards the function $f : \mathbf{R} \to \mathbf{R}$, we assume that

$$K := \sup_{0\leq x\neq y\leq L} \frac{|f(x)-f(y)|}{|y-x|} + \sup_{0\leq x\leq L} |f(x)| < \infty. \tag{90}$$

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

$$\begin{aligned}
&\int_0^L u(x\,,t)\phi(x)\,dx - \int_0^L u_0(x)\phi(x)\,dx \\
&= \int_0^t \int_0^L \frac{\partial^2 u}{\partial x^2}(x\,,s)\phi(x)\,dx\,ds + \int_0^t \int_0^L f\left(u(x\,,s)\right)\phi(x) W(dx\,ds).
\end{aligned} \tag{91}$$

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

$$\int_0^t \int_0^L \frac{\partial^2 u}{\partial x^2} u(x\,,s)\phi(x)\,dx\,ds = \int_0^t \int_0^L u(x\,,s)\phi''(x)\,dx\,ds. \tag{92}$$

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

$$\begin{aligned}
&\int_0^L u(x\,,t)\phi(x)\,dx - \int_0^L u_0(x)\phi(x)\,dx \\
&= \int_0^t \int_0^L u(x\,,s)\phi''(x)\,dx\,ds + \int_0^t \int_0^L f\left(u(x\,,s)\right)\phi(x) W(dx\,ds).
\end{aligned} \tag{93}$$

**Exercise 6.2 (Important).** Argue that if $u$ solves (93), then for all $C^\infty$ functions $\psi(x\,,t)$ with $\partial_x\psi(0\,,t) = \partial_x\psi(L\,,t) = 0$,

$$\begin{aligned}
&\int_0^L u(x\,,t)\psi(x\,,t)\,dx - \int_0^L u_0(x)\psi(x\,,0)\,dx \\
&= \int_0^t \int_0^L u(x\,,s)\left[\frac{\partial^2 u}{\partial x^2}\psi(x\,,s) + \frac{\partial\psi}{\partial s}(x\,,s)\right]dx\,ds \\
&\quad + \int_0^t \int_0^L f\left(u(x\,,s)\right)\psi(x\,,s) W(dx\,ds).
\end{aligned} \tag{94}$$

This is formal, but important.

Let $G_t(x\,,y)$ denote the Green's function for the linear heat equation. [The subscript $t$ is *not* a derivative, but a variable.] Then it follows from the method of images that

$$G_t(x\,,y) = \sum_{n=-\infty}^{\infty} \big[\Gamma(t\,;x-y-2nL) + \Gamma(t\,;x+y-2nL)\big], \tag{95}$$

where $\Gamma$ is the fundamental solution to the linear heat equation (89); i.e.,

$$\Gamma(t\,;a) = \frac{1}{(4\pi t)^{1/2}} \exp\left(-\frac{a^2}{4t}\right). \tag{96}$$

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

$$G_t(\phi\,,y) := \int_0^L G_t(x\,,y)\phi(x)\,dx, \tag{97}$$

if $t > 0$, and $G_0(\phi\,,y) := \phi(y)$. We can integrate (89)—with $f(u\,,t) \equiv 0$—by parts for all $C^\infty$ functions $\phi : [0\,,L] \to \mathbf{R}$ such that $\phi'(0) = \phi'(L) = 0$, and obtain the following:

$$G_t(\phi\,,y) = \phi(y) + \int_0^t G_s\left(\phi'' - \phi\,,y\right)\,ds. \tag{98}$$

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

$$\frac{\partial^2\psi}{\partial x^2}(x\,,s) + \frac{\partial\psi}{\partial s}(x\,,s) = 0, \quad \psi(x\,,t) = \phi(x),\;\; \psi(x\,,0) = G_t(\phi\,,x). \tag{99}$$

Use this $\psi$ in Exercise 6.2 to find that any solution to (89) must satisfy

$$\begin{aligned}\int_0^L u(x\,,t)\phi(x)\,dx - \int_0^L u_0(y)G_t(\phi\,,y)\,dy \\ = \int_0^t\int_0^L f\left(u(y\,,s)\right)G_{t-s}(\phi\,,y)W(dy\,ds).\end{aligned} \tag{100}$$

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

$$\begin{aligned}u(x\,,t) - \int_0^L u_0(y)G_t(x\,,y)\,dy \\ = \int_0^t\int_0^L f\left(u(y\,,s)\right)G_{t-s}(x\,,y)W(dy\,ds).\end{aligned} \tag{101}$$

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* (89) *we mean a solution $u$ to* (101) *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* (93) *subject to* (90) *has an a.s.-unique solution $u$ that satisfies the following for all $T>0$:*

$$\sup_{0\le x\le L}\sup_{0\le t\le T}\mathrm{E}\left(|u(x\,,t)|^2\right)<\infty. \tag{102}$$

For its proof we will need the following well-known result.

**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 $t\in[0\,,T]$,*

$$\phi_{n+1}(t)\le A\int_0^t\phi_n(s)\,ds. \tag{103}$$

*Then,*

$$\phi_n(t)\le\phi_1(T)\frac{(At)^{n-1}}{(n-1)!}\quad\text{for all } n\ge 1 \text{ and } t\in[0\,,T]. \tag{104}$$

The preceding is proved by applying induction. I omit the details.

*Remark 6.6.* As a consequence of Gronwall's lemma, any positive power of $\phi_n(t)$ is summable in $n$. Also, if $\phi_n$ does not depend on $n$ then it follows that $\phi_n\equiv 0$.

*Proof (Theorem 6.4: Uniqueness).* Suppose $u$ and $v$ both solve (101), and both satisfy the integrability condition (102). We wish to prove that $u$ and $v$ are modifications of one another. Let $d(x\,,t):=u(x\,,t)-v(x\,,t)$. Then,

$$d(x\,,t)=\int_0^t\int_0^L\Big[f(u(y\,,s))-f(v(y\,,s))\Big]G_{t-s}(x\,,y)\,W(dy\,ds). \tag{105}$$

According to Theorem 5.26 (p. 25) and (90),

$$\mathrm{E}\left(|d(x\,,t)|^2\right)\le K^2\int_0^t\int_0^L\mathrm{E}\left(|d(y\,,s)|^2\right)G^2_{t-s}(x\,,s)\,dy\,ds. \tag{106}$$

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

$$H(t)\le K^2\int_0^t H(s)\left(\int_0^L G^2_{t-s}(x\,,y)\,dy\right)ds. \tag{107}$$

Now from (95) and the semigroup properties of $\Gamma$ it follows that

$$\int_0^L G_t(x\,,y)G_s(y\,,z)\,dy = G_{t+s}(x\,,z), \text{ and } G_t(x\,,y) = G_t(y\,,x). \tag{108}$$

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

$$H(t) \le CK^2 \int_0^t \frac{H(s)}{|t-s|^{1/2}}\,ds. \tag{109}$$

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 $t \in [0\,,T]$,

$$H(t) \le A\left(\int_0^t H^q(s)\,ds\right)^{1/q}. \tag{110}$$

We can apply Gronwall's Lemma 6.5 with $\phi_1 = \phi_2 = \phi_3 = \cdots = H^q$ to find that $H(t) \equiv 0$. □

*Proof (Theorem 6.4: Existence).* Note from (95) that $\int_0^L G_t(x\,,y)\,dy$ is a number in $[0\,,1]$. Because $u_0$ is assumed to be bounded $\int_0^L u_0(y)G_t(x\,,y)\,dy$ is bounded; this is the first term in (101). Now we proceed with a Picard-type iteration scheme. Let $u_0(x\,,t) := u_0(x)$, and then iteratively define

$$\begin{aligned} &u_{n+1}(x\,,t) \\ &\quad= \int_0^L u_0(y)G_t(x\,,y)\,dy + \int_0^t\int_0^L f\left(u_n(y\,,s)\right)G_{t-s}(x\,,y)W(dy\,ds). \end{aligned} \tag{111}$$

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

$$\begin{aligned} &d_n(x\,,t) \\ &\quad= \int_0^t\int_0^L \left[f\left(u_{n+1}(y\,,s)\right) - f\left(u_n(y\,,s)\right)\right]G_{t-s}(x\,,y)\,W(dy\,ds). \end{aligned} \tag{112}$$

Consequently, by (90),

$$\mathrm{E}\left(|d_n(x\,,t)|^2\right) \le K^2\int_0^t\int_0^L \mathrm{E}\left(|d_{n-1}(y\,,s)|^2\right)G_{t-s}^2(x\,,y)\,dy\,ds. \tag{113}$$

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

$$H_n^2(t) \le CK^2\int_0^t \frac{H_{n-1}^2(s)}{|t-s|^{1/2}}\,ds. \tag{114}$$

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 $t \in [0\,,T]$,

$$H_n^2(t) \le A\left(\int_0^t H_{n-1}^{2q}(s)\,ds\right)^{1/q}. \tag{115}$$

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

$$\begin{aligned}\lim_{n\to\infty} \int_0^t \int_0^L & f(u_n(y\,,s))G_{t-s}(x\,,y)\,W(dy\,ds) \\ &= \int_0^t \int_0^L f(u(y\,,s))G_{t-s}(x\,,y)\,W(dy\,ds),\end{aligned} \tag{116}$$

where the convergence holds in $L^2(\mathrm{P})$. This proves that $u$ is a solution to (101). □

We are finally ready to complete the picture by proving that the solution to (89) is continuous [up to a modification, of course].

**Theorem 6.7.** *There exists a continuous modification* $u(x\,,t)$ *of* (89).

*Remark 6.8.* In Exercise 6.9, on page 35 below, you will be asked to improve this to the statement that there exists a Hölder-continuous modification.

*Proof (Sketch).* We need the following easy-to-check fact about the Green's function $G$:

$$G_t(x\,,y) = \Gamma(t\,;x-y) + H_t(x\,,y), \tag{117}$$

where $H_t(x\,,y)$ is smooth in $(t\,,x\,,y) \in \mathbf{R}_+ \times \mathbf{R} \times \mathbf{R}$, and $\Gamma$ is the "heat kernel" defined in (96). Define

$$U(x\,,t) := \int_0^t \int_0^L f(u(y\,,s))\Gamma(t-s\,;x-y)\,W(dy\,ds). \tag{118}$$

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 $\Gamma$ and $G$ is smooth and bounded. From here on I prove things honestly.

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

$$\begin{aligned}&U(x\,,t') - U(x\,,t) \\ &= \int_0^t \int_0^L f(u(y\,,s))\left[\Gamma(t'-s\,;x-y) - \Gamma(t-s\,;x-y)\right]\,W(dy\,ds) \\ &\qquad + \int_t^{t'} \int_0^L f(u(y\,,s))\Gamma(t'-s\,;x-y)\,W(dy\,ds).\end{aligned} \tag{119}$$

By Burkholder's inequality (Theorem 5.27, page 25) and the elementary inequality $|a+b|^p \le 2^p|a|^p + 2^p|b|^p$,

$$\begin{aligned}
&\mathrm{E}\left(|U(x\,,t) - U(x\,,t')|^p\right)\\
&\le 2^p c_p \mathrm{E}\left[\left(\int_0^t \int_0^L f^2(u(y\,,s))\Lambda(s\,,t\,,t'\,;x\,,y)\;dy\,ds\right)^{p/2}\right]\\
&\qquad + 2^p c_p \mathrm{E}\left[\left(\int_t^{t'} \int_0^L f^2(u(y\,,s))\Gamma^2(t-s\,;x-y)\,dy\,ds\right)^{p/2}\right].
\end{aligned} \tag{120}$$

where

$$\Lambda(s\,,t\,,t'\,;x\,,y) := \left[\Gamma(t'-s\,;x-y) - \Gamma(t-s\,;x-y)\right]^2. \tag{121}$$

Because of (90), $\sup|f| \le K$. Therefore,

$$\begin{aligned}
&\mathrm{E}\left(|U(x\,,t) - U(x\,,t')|^p\right)\\
&\le (2K)^p c_p \left(\int_0^t \int_{-\infty}^{\infty} \Lambda(s\,,t\,,t'\,;x\,,y)\,dy\,ds\right)^{p/2}\\
&\qquad + (2K)^p c_p \left(\int_t^{t'} \int_{-\infty}^{\infty} \Gamma^2(t-s\,;x-y)\,dy\,ds\right)^{p/2}.
\end{aligned} \tag{122}$$

[Notice the change from $\int_0^L$ to $\int_{-\infty}^{\infty}$.] Because $\int_{-\infty}^{\infty} \Gamma^2(t-s\,;a)\,da$ is a constant multiple of $|t-s|^{-1/2}$,

$$\left(\int_t^{t'} \int_{-\infty}^{\infty} \Gamma^2(t-s\,;x-y)\,dy\,ds\right)^{p/2} = C_p|t'-t|^{p/4}. \tag{123}$$

For the other integral we use a method that is motivated by the ideas in [5]. Recall Plancherel's theorem: For all $g \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$,

$$\|g\|^2_{L^2(\mathbf{R})} = \frac{1}{2\pi}\|\mathscr{F}g\|^2_{L^2(\mathbf{R})}, \tag{124}$$

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

$$\begin{aligned}
&\int_{-\infty}^{\infty} \left[\Gamma(t'-s\,;x-y) - \Gamma(t-s\,;x-y)\right]^2\;dy\\
&\qquad = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[e^{-(t'-s)\xi^2} - e^{-(t-s)\xi^2}\right]^2\;d\xi\\
&\qquad = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-2(t-s)\xi^2}\left[1 - e^{-(t'-t)\xi^2}\right]^2\;d\xi.
\end{aligned} \tag{125}$$

Therefore,

$$\int_0^t \int_{-\infty}^{\infty} \left[\Gamma(t'-s\,;x-y) - \Gamma(t-s\,;x-y)\right]^2 dy\,ds$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \left(\int_0^t e^{-2(t-s)\xi^2}\,ds\right)\left[1-e^{-(t'-t)\xi^2}\right]^2 d\xi \tag{126}$$

$$= \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{1-e^{-2t\xi^2}}{\xi^2}\left[1-e^{-(t'-t)\xi^2}\right]^2 d\xi.$$

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,

$$\int_0^t \int_{-\infty}^{\infty} \left[\Gamma(t'-s\,;x-y) - \Gamma(t-s\,;x-y)\right]^2 dy\,ds$$

$$\le \frac{C_T}{\pi}\int_0^{\infty} \frac{\min[(t'-t)\xi^2\,,1]}{1+\xi^2}\,d\xi \tag{127}$$

$$\le \frac{C_T}{\pi}\left(\int_{|t'-t|^{-1/2}}^{\infty} \frac{d\xi}{\xi^2} + \int_0^{|t'-t|^{-1/2}} \frac{(t'-t)\xi^2}{1+\xi^2}\,d\xi\right).$$

The first term is equal to $A|t'-t|^{1/2}$, and the second term is also bounded above by $|t'-t|^{1/2}$ because $\xi^2/(1+\xi^2) \le 1$. This, (122) and (123) together prove that

$$\mathrm{E}\left(|U(x\,,t) - U(x\,,t')|^p\right) \le C_p|t'-t|^{p/4}. \tag{128}$$

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

$$\mathrm{E}\left(|U(x\,,t) - U(x'\,,t)|^p\right)$$

$$\le c_p K^p \left(\int_0^t \int_{-\infty}^{\infty} \left|\Gamma(t-s\,;y) - \Gamma(t-s\,;x'-x-y)\right|^2 dy\,ds\right)^{p/2}. \tag{129}$$

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

$$\int_{-\infty}^{\infty} \left|\Gamma(t-s\,;y) - \Gamma(t-s\,;x'-x-y)\right|^2 dy$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-2(t-s)\xi^2}\left|1-e^{i\xi(x'-x)}\right|^2 d\xi. \tag{130}$$

Consequently, we can apply Tonelli's theorem to find that

$$\int_0^t \int_{-\infty}^{\infty} \left|\Gamma(t-s\,;y) - \Gamma(t-s\,;x'-x-y)\right|^2 dy\,ds$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1-e^{-2t\xi^2}}{2\xi^2}\left|1-e^{i\xi(x'-x)}\right|^2 d\xi \tag{131}$$

$$= \frac{1}{4\pi}\int_0^{\infty} \frac{1-e^{-2t\xi^2}}{\xi^2}\left|1-\cos(\xi(x'-x))\right| d\xi.$$

We use the elementary bounds $1-\exp(-|\theta|) \le 1$, and $1-\cos\theta \le \min(1\,,\theta^2)$—valid for all $\theta \in \mathbf{R}$—in order to bound the preceding, and obtain

$$\begin{aligned}\int_0^t \int_{-\infty}^{\infty} \left|\Gamma(t-s\,;y) - \Gamma(t-s\,;x'-x-y)\right|^2 dy\,ds \\ \le \frac{1}{4\pi}\int_0^\infty \frac{\xi^2(x'-x)^2 \wedge 1}{\xi^2}\,d\xi.\end{aligned} \tag{132}$$

We split the domain of integration into two domains: Where $\xi < |x'-x|^{-1}$; and where $\xi \ge |x'-x|^{-1}$. Each of the two resulting integrals is easy enough to compute explicitly, and we obtain

$$\int_0^t \int_{-\infty}^{\infty} \left|\Gamma(t-s\,;y) - \Gamma(t-s\,;x'-x-y)\right|^2 dy\,ds \le \frac{|x'-x|}{2\pi} \tag{133}$$

as a result. Hence, it follows that

$$\sup_{t\ge 0} \mathrm{E}\left(|U(x\,,t) - U(x',t)|^p\right) \le a_p|x'-x|^{p/2}. \tag{134}$$

For all $(x\,,t) \in \mathbf{R}^2$ define $|(x\,,t)| := |x|^{1/2} + |t|^{1/4}$. This defines a norm on $\mathbf{R}^2$, and is equivalent to the usual Euclidean norm $(x^2+t^2)^{1/2}$ in the sense that both generate the same topology. Moreover, we have by (128) and (134): For all $t, t' \in [0\,,T]$ and $x, x' \in [0\,,L]$,

$$\mathrm{E}\left(|U(x\,,t) - U(x',t')|^p\right) \le A\,|(x\,,t)-(x',t')|^p\,. \tag{135}$$

This and Kolmogorov's continuity theorem (Theorem 4.13, page 18) 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 $< \frac{1}{2}$; and it is Hölder continuous in $t$ of any given order $< \frac{1}{4}$.

**Exercise 6.10.** 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

$$\left|\begin{array}{ll}\dfrac{\partial u}{\partial t} = \left(\dfrac{\partial^2 u}{\partial x_1^2} + \dfrac{\partial^2 u}{\partial x_2^2}\right) + \dot W & t>0,\ x\in\mathbf{R}^2, \\ u(x\,,0) = 0 & x \in \mathbf{R}^2.\end{array}\right. \tag{136}$$

Interpret the adapted solution to the preceding as

$$u(x\,,t)=\int_0^t\int_{\mathbf{R}^2}\Gamma(t-s\,;x-y)\,W(dy\,ds),\tag{137}$$

subject to $(t\,,x)\mapsto \mathrm{E}[u^2(t\,,x)]$ being continuous (say!). Here, $\Gamma$ is the heat kernel on $\mathbf{R}^2$; that is, $\Gamma(t\,,x):=(4\pi t)^{-1}\exp(-\|x\|^2/(4t))$. 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 $\geq 2$.

## 7 From chaos to order

Finally, I mention an example of SPDEs that produce smooth solutions for all times $t>0$, and yet the solution is white noise at time $t=0$. In this way, one can think of the solution to the forthcoming SPDE as a smooth deformation of white noise, where the deformation is due to the action of the heat operator.

Now consider the heat equation on $[0\,,1]$, but with random initial data instead of random forcing terms. More specifically, we consider the stochastic process $\{u(x\,,t)\}_{0\leq x\leq 1, t\geq 0}$ that is formally defined by

$$\left|\begin{array}{ll}\dfrac{\partial u}{\partial t}(x\,,t)=\dfrac{\partial^2 u}{\partial x^2}(x\,,t) & 0<x<1,\;\; t\geq 0\\ u(0\,,t)=u(1\,,t)=0 & t>0\\ u(x\,,0)=\dot W(x) & 0<x<1,\end{array}\right.\tag{138}$$

where $\dot W$ denotes white noise.

A classical interpretation of (138) follows: Consider an infinitesimally-thin wire of length one that has even density and width. Interpret this wire as the interval $[0\,,1]$, and apply totally random heat to the wire, the heat amount at $x$ being $\dot W(x)$ units. The endpoints of the wire are perfectly cooled. If we watch the wire cool as time passes, then the amount of heat retained at position $x$ at time $t>0$ is $u(x\,,t)$.

If $\dot W$ were replaced by a square-integrable function then the solution is classical, and is given by

$$u(x\,,t)=\sqrt{2}\sum_{n=1}^\infty \xi_n\sin(n\pi x)\exp\left(-n^2\pi^2 t\right),\tag{139}$$

where

$$\xi_n:=\sqrt{2}\int_0^1\dot W(x)\sin(n\pi x)\,dx,\tag{140}$$

and the infinite sum in (139) converges in $L^2(dx)$ for each $t>0$, for example. Although $\dot W$ is not a square-integrable function, one can first consider "weak solutions," and then proceed to integrate by parts, and thus arrive at the *mild solution* to (138). That is described by (139), but with (140) replaced by the Wiener stochastic integrals

$$\xi_n := \sqrt{2}\int_0^1 \sin(n\pi x) W(dx), \qquad n = 1, 2, \ldots . \tag{141}$$

It follows from our construction of Wiener integrals that $\{\xi_n\}_{n=1}^\infty$ is a mean-zero Gaussian process. Thanks to the Wiener isometry (20), we also can compute its covariance structure to find that

$$\mathrm{Cov}(\xi_n\,, \xi_m) = 2\int_0^1 \sin(n\pi x)\sin(m\pi x)\,dx = \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases} \tag{142}$$

Consequently, $\{\xi_n\}_{n=1}^\infty$ is an i.i.d. sequence of standard-normal variates. The following lemma controls the rate of growth of the $\xi_n$'s.

**Lemma 7.1.** *With probability one,*

$$|\xi_n| = O\left(\sqrt{\ln n}\right) \qquad \text{as } n \to \infty. \tag{143}$$

*Proof.* We can apply Chebyshev's inequality to find that for all $a, \lambda > 0$,

$$\mathrm{P}\{\xi_n \geq a\} \leq e^{-\lambda a}\mathrm{E}\exp(\lambda \xi_1) = \exp\left(-\lambda a + \frac{\lambda^2}{2}\right). \tag{144}$$

The optimal choice of $\lambda$ is $a/2$, and this yields the following well-known bound: $\mathrm{P}\{\xi_n \geq a\} \leq \exp(-a^2/2)$, valid for all $a > 0$. By symmetry,

$$\mathrm{P}\left\{|\xi_n| \geq a\right\} \leq 2\exp(-a^2/2) \qquad \text{for all } a > 0. \tag{145}$$

We plug in $a := 2\sqrt{\ln n}$ and deduce (143) from

$$\sum_{n \geq 100} \mathrm{P}\left\{|\xi_n| \geq 2\sqrt{\ln n}\right\} \leq \sum_{n \geq 100} \frac{2}{n^2} < \infty \tag{146}$$

and the Borel–Cantelli lemma. □

**Exercise 7.2.** Improve Lemma 7.1 to the statement that

$$\limsup_{n\to\infty} \frac{\xi_n}{\sqrt{2\ln n}} = -\liminf_{n\to\infty} \frac{\xi_n}{\sqrt{2\ln n}} = 1 \qquad \text{a.s.} \tag{147}$$

An immediate consequence of Lemma 7.1 is that for all fixed $0 < r < R$, the infinite series in (139) converges a.s., uniformly for $(x\,, t) \in [0\,, 1] \times [r\,, R]$.

Among other things, this proves that $u$ is infinitely differentiable in both variables, away from time zero.

Thus, the random function $u$ is smooth except near time zero, where its behavior is chaotic. In words, the heat operator takes the pure-noise initial condition "$u(x\,,0) = \dot{W}(x)$" and immediately smooths it to generate nice random functions $u(x\,,t)$, one for every $t > 0$. Thus, it is interesting to investigate the transition from "chaos" $[t = 0]$ to "order" $[t > 0]$ in greater depth.

Here we study the mentioned blowup problem for average $x$-values, and plan to prove that there is a sense in which the following holds for all "typical values of $x$":

$$u(x\,,t) \approx t^{-1/4} \qquad \text{when } t \approx 0. \tag{148}$$

Define

$$\mathscr{E}(t) := \left( \int_0^1 |u(x\,,t)|^2 \, dx \right)^{1/2}. \tag{149}$$

A classical interpretation of $\mathscr{E}(t)$ is the average heat—in the sense of $L^2(dx)$—in the wire at time $t$, where the wire at time 0 is subjected to heat amount $\dot{W}(x)$ at position $x \in [0\,,1]$. The following rigorous interpretation of (148) is a rather simple result that describes roughly the nature of the blowup of the solution near time zero.

**Theorem 7.3.** *With probability one,*

$$\lim_{t \searrow 0} t^{1/4} \mathscr{E}(t) = \frac{1}{(2\pi)^{3/4}}. \tag{150}$$

The proof of Theorem 7.3 relies on a lemma from calculus.

**Lemma 7.4.** *The following holds:*

$$\lim_{\lambda \searrow 0} \lambda^{1/2} \sum_{n=1}^{\infty} e^{-n^2 \lambda} = \frac{1}{2\sqrt{\pi}}. \tag{151}$$

*Proof.* Because $\int_0^\infty \exp(-x^2 \lambda)\, dx = 1/(2\sqrt{\pi\lambda})$,

$$\int_8^\infty e^{-x^2\lambda}\, dx = O(1) + \frac{1}{2\sqrt{\pi\lambda}} \qquad \text{as } \lambda \searrow 0. \tag{152}$$

Because $\sum_{k=1}^{8} \exp(-n^2\lambda) = O(1)$ as $\lambda \searrow 0$, it therefore suffices to prove that

$$T := \left| \sum_{n=9}^{\infty} e^{-n^2\lambda} - \int_8^\infty e^{-x^2\lambda}\, dx \right| = o\left( \frac{1}{\sqrt{\lambda}} \right) \qquad \text{as } \lambda \searrow 0. \tag{153}$$

To prove this we first write $T$ as

$$T = \sum_{n=9}^{\infty} \int_{n-1}^{n} e^{-x^2\lambda} \left(1 - e^{-(n^2-x^2)\lambda}\right) dx. \tag{154}$$

Because $1 - \exp(-\theta) \le 1 \wedge \theta$ for all $\theta \ge 0$, and since $n^2 - x^2 \le 4x$ for all $x \in [n-1\,, n]$ and $n \ge 1$,

$$\begin{aligned} T &\le 4 \int_8^{\infty} e^{-x^2\lambda} \left(1 \wedge x\lambda\right) dx \\ &\le \frac{4}{\sqrt{\lambda}} \int_0^{\infty} e^{-y^2} \left(1 \wedge y\sqrt{\lambda}\right) dy, \end{aligned} \tag{155}$$

and this is $o(1/\sqrt{\lambda})$ by the dominated convergence theorem. This proves (153), and hence the lemma. □

Next we prove Theorem 7.3.

*Proof (Theorem 7.3).* Equation (142) and the uniform convergence of the series in (139) together imply that for all $t > 0$,

$$\mathscr{E}^2(t) = \sum_{n=1}^{\infty} \xi_n^2 e^{-2n^2\pi^2 t} \qquad \text{a.s.} \tag{156}$$

Consequently, Lemma 7.4 implies that

$$\mathrm{E}\left(|\mathscr{E}(t)|^2\right) = \sum_{n=1}^{\infty} e^{-2n^2\pi^2 t} = \frac{1+o(1)}{(2\pi)^{3/2}\sqrt{t}} \qquad \text{as } t \searrow 0. \tag{157}$$

Because the $\xi_n$'s are independent, a second application of Lemma 7.4 yields

$$\begin{aligned} \mathrm{Var}\left(|\mathscr{E}(t)|^2\right) &= \mathrm{Var}(\xi_1^2) \sum_{n=1}^{\infty} e^{-4n^2\pi^2 t} \\ &= O\left(\mathrm{E}\left(|\mathscr{E}(t)|^2\right)\right) \qquad \text{as } t \searrow 0. \end{aligned} \tag{158}$$

These remarks, together with the Chebyshev inequality, yield two constants $C, \varepsilon > 0$ such that for all $t \in (0\,, \varepsilon)$ and $\delta > 0$,

$$\mathrm{P}\left\{\left|\frac{\mathscr{E}^2(t)}{\mathrm{E}\left(|\mathscr{E}(t)|^2\right)} - 1\right| > \delta\right\} \le C\sqrt{t}. \tag{159}$$

We can replace $t$ by $k^{-4}$, sum both sides from $k = 1$ to $k = \infty$, apply the Borel–Cantelli lemma, and then finally deduce that

$$\lim_{k\to\infty} \frac{\mathscr{E}^2\left(k^{-4}\right)}{\mathrm{E}\left(\left|\mathscr{E}\left(k^{-4}\right)\right|^2\right)} = 1 \qquad \text{a.s.} \tag{160}$$

Because $\mathscr{E}^2$ is non-increasing, (157) and a monotonicity argument together finish the proof. □

**Exercise 7.5 (Rapid cooling).** Prove that with probability one,

$$\lim_{t\nearrow\infty} \exp\left(\pi^2 t\right) \mathscr{E}(t) = 1. \tag{161}$$

That is, the wire cools rapidly as time goes by, as it does for classical initial heat profiles. Thus, the only new phenomenon occurs near time zero.

**Exercise 7.6.** Define the *average heat flux* in the wire as

$$\mathscr{F}(t) := \left( \int_0^1 \left| \frac{\partial u}{\partial x}(x\,,t) \right|^2 dx \right)^{1/2}. \tag{162}$$

Describe the blowup rate of $\mathscr{F}(t)$ as $t$ tends down to zero.

For a greater challenge try the following.

**Exercise 7.7.** Prove that as $t \searrow 0$, and after suitable centering and normalization, $\mathscr{E}(t)$ converges in distribution to a non-degenerate law. Describe that law.

**Exercise 7.8.** Prove that $\{b(x)\}_{0\le x\le 1}$ is a Brownian bridge, where

$$b(x) := \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{u(x\,,t)}{\sqrt{t}}\, dt \qquad \text{for all } x \in [0\,,1]. \tag{163}$$

## References

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

[2] V. V. Baklan (1965). The existence of solutions for a class of equations involving variational derivatives, *Dopovidi Akad. Nauk Ukraïn. RSR* **1965**, 554–556 (Ukranian. Russian, English summary)

[3] D. L. Burkholder (1971). Martingale inequalities, In: *Lecture Notes in Math.* **190**, 1–8 Springer-Verlag, Berlin

[4] E. M. Cabaña (1970). The vibrating string forced by white noise, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **15**, 111–130

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

[6] Robert C. Dalang, Davar Khoshnevisan, and Eulalia Nualart (2007). Hitting probabilities for systems of non-linear stochastic heat equations with additive noise, *Latin American J. Probab. and Math. Statist.* (or *Alea*; `http://alea.impa.br/english`), Vol. III, 231–371

[7] Ju. L. Daleckiĭ(1967). Infinite-dimensional elliptic operators and the corresponding parabolic equations, *Uspehi Mat. Nauk* **22** *(4) (136)*, 3–54 (In Russian) [English translation in: *Russian Math. Surveys* **22** *(4)*, 1–53, 1967]

[8] D. A. Dawson (1975). Stochastic evolution equations and related measure processes, *J. Multivariate Anal.* **5**, 1–52.

[9] D. A. Dawson (1972). Stochastic evolution equations, *Math. Biosci.* **15**, 287–316.

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

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

[12] X. Fernique (1975). Regularité des trajectoires des fonctions aléatoires gaussiennes, In: *Lecture Notes in Math.* **480**, 1–96 Springer-Verlag, Berlin (in French)

[13] Tadahisa Funaki (1984). Random motion of strings and stochastic differential equations on the space $C([0,1], \mathbf{R}^d)$, In: *Stochastic Analysis (Katata/Kyoto, 1982)*, North-Holland Math. Library, **32**, 121–133, North-Holland, Amsterdam

[14] Tadahisa Funaki (1983). Random motion of strings and related stochastic evolution equations, *Nagoya Math. J.* **89**, 129–193

[15] Kiyosi Itô (1944). Stochastic integral, *Proc. Imp. Acad. Tokyo* **20**, 519–524

[16] Kiyosi Itô (1950). Stochastic differential equations in a differentiable manifold, *Nagoya Math. J.* **1**, 35–47

[17] Kiyosi Itô (1951). On a formula concerning stochastic differentials, *Nagoya Math. J.* **3**, 55–65

[18] A. Ya. Khintchine (1933). *Asymptotische Gesetz der Wahrscheinlichkeitsrechnung*, Springer, Berlin

[19] A. N. Kolmogorov (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin

[20] N. V. Krylov and B. L. Rozovskiĭ (1979a). Itô equations in Banach spaces and strongly parabolic stochastic partial differential equations, *Dokl. Akad. Nauk SSSR* **249** *(2)*, 285–289 (in Russian)

[21] N. V. Krylov and B. L. Rozovskiĭ (1979b). Stochastic evolution equations, In: *Current Problems in Mathematics, Vol. 14* Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 71–147, 256 (in Russian)

[22] N. V. Krylov, N. V. and B. L. Rozovskiĭ (1977). The Cauchy problem for linear stochastic partial differential equations, *Izv. Akad. Nauk SSSR Ser. Mat.* **41** *(6)*, 1329–1347, 1448 (in Russian)

[23] H. Kunita (1991). *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, Cambridge

[24] Étienne Pardoux (1975). *Equations aux dérivées partielles stochastiques non linéaires monotones—Étude de solutions fortes de type Itô*, Thése d'État, Univ. Paris XI, Orsay

[25] Étienne Pardoux (1972). Sur des équations aux dérivées partielles stochastiques monotones, *C. R. Acad. Sci. Paris Sér. A–B* **275**, A101–A103

[26] Christopher Preston (1972). Continuity properties of some Gaussian processes, *Ann. Math. Statist.* **43**, 285–292

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

[28] Michel Talagrand (1987). Regularity of Gaussian processes, *Acta Math.* **159***(1-2)*, 99–149

[29] G. E. Uhlenbeck and L. S. Ornstein (1930). On the theory of Brownian Motion, *Phys. Rev.* **36** 823–841

[30] John B. Walsh (1986). *An Introduction to Stochastic Partial Differential Equations*, In: Lecture Notes in Math. **1180**, 265–439, Springer, Berlin

[31] N. Wiener (1923). Differential space, *J. Math. Phys.* **2**, 131–174

[32] Norbert Wiener (1938). The Homogeneous Chaos, *Amer. J. Math.* **60***(4)*, 897–936