Lecture notes for Introduction to SPDE, Spring 2016

Lenya Ryzhik*

May 6, 2016

Nothing found here is original except for a few mistakes and misprints here and there. These notes are simply a record of what I cover in class, to spare the students the necessity of taking the lecture notes. The readers should consult the original books for a better presentation and context. We plan to follow the lecture notes by Davar Khoshnevisan and the book by John Walsh.

1 The white noise

We would like to be able to make sense of the solutions of equations with rough forces, such as the heat equation

$$\frac{\partial u}{\partial t} = \Delta u + F(t,x), \tag{1.1}$$

or the wave equation

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \Delta u + F(t,x), \tag{1.2}$$

with a highly irregular function $F(t,x)$, as well as nonlinear versions of these equations. However, if F is "very rough" then, presumably, the solution $u(t,x)$ will not be very smooth either, hence one would not be able to differentiate it in time or space, and the sense in which $u(t,x)$ solves the corresponding equation is not quite clear. It is natural to think of $u(t,x)$ as a weak solution of the PDE as that would not require differentiation – but if the PDE (unlike the examples above) is nonlinear we would still need to know that $u(t,x)$ is a function, which is not a priori obvious if the force $F(t,x)$ is too irregular. As we will see, it is often the case that $u(t,x)$ is actually not a function.

Another obvious issue is to understand what we would mean by a "rough force". If, say, (1.1) were posed on the lattice $\mathbb{Z}^d$, and Δ were a discrete Laplacian, then (1.1) would be an infinite system of SDE's at each lattice site, with $F(t,x)$ independent for each site $x \in \mathbb{Z}^d$. This makes clear sense, as long as x is discrete. In order to define this in $\mathbb{R}^d$, we need to develop some basics. A natural way to generalize independence at each site is to require that $F(t,x)$ is a stationary in time and space mean-zero process such that the two-point correlation function is

$$\mathbb{E}[F(t,x)F(t',x')] = \begin{cases} 1 & \text{if } t=t' \text{ and } x=x', \\ 0 & \text{otherwise.} \end{cases} \tag{1.3}$$

*Department of Mathematics, Stanford University, Stanford, CA 94305, USA; ryzhik@math.stanford.edu

Such random field, however, does not exist, and the next best choice would be a random field with a correlation function

$$\mathbb{E}[F(t,x)F(t',x')] = \delta(t-t')\delta(x-x'). \tag{1.4}$$

The first order of business would be to make this rigorous.

When is the solution a function?

Disregarding the question of a careful definition of such random process, let us see what we can expect about the solutions of, say, the heat equation (1.1) with such force. The Duhamel formula says that if $u(0,x)=0$, then the solution of (1.1) in $\mathbb{R}^d$ is

$$u(t,x) = \int_0^t \int_{\mathbb{R}^d} G(t-s,x-y)F(s,y)dsdy. \tag{1.5}$$

Here,

$$G(t,x) = \frac{1}{(4\pi t)^{d/2}} e^{-|x|^2/(4t)} \tag{1.6}$$

is the standard heat kernel. The function $u(t,x)$ is a stationary field in x, hence we can not possibly expect any decay in x but we may still ask how large it should be. Let us compute its point-wise second moment:

$$\mathbb{E}\Big[|u(t,x)|^2\Big] = \int_0^t \int_0^t \int_{\mathbb{R}^{2d}} G(t-s,x-y)G(t-s',x-y')\mathbb{E}[F(s,y)F(s',y')]dsds'dydy'$$
$$= \int_0^t \int_{\mathbb{R}^d} |G(t-s,x-y)|^2 dsdy = \int_0^t \int_{\mathbb{R}^d} |G(s,y)|^2 dsdy = \frac{1}{(4\pi)^d}\int_0^t \int_{\mathbb{R}^d} e^{-|y|^2/(2s)}\frac{dyds}{s^d}$$
$$= C_d \int_0^t \frac{ds}{s^{d/2}}. \tag{1.7}$$

We see that $u(t,x)$ has a point-wise second moment if and only if $d=1$ – therefore, it is only in one dimension that we may expect the solution of a typical SPDE be a function.

Is the solution a distribution?

Since $u(t,x)$ does not seem to be a function in $d>1$, let us see if the field given by (1.5) at least makes sense as a distribution in x, pointwise in time. We multiply (1.5) by a test function $\phi \in C_c^\infty(\mathbb{R}^d)$:

$$\Phi(t) = \langle u(t),\phi\rangle = \int_0^t \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} G(t-s,x-y)F(s,y)\phi(x)dsdydx,$$

and compute the second moment of $\Phi(t)$:

$$\mathbb{E}\Big[|\Phi(t)|^2\Big] = \int_0^t \int_0^t \int_{\mathbb{R}^{4d}} G(t-s,x-y)G(t-s',x'-y')\mathbb{E}[F(s,y)F(s',y')] \tag{1.8}$$
$$\times \phi(x)\phi(x')dsds'dydy'dxdx' = \int_0^t \int_{\mathbb{R}^{3d}} G(t-s,x-y)G(t-s,x'-y)\phi(x)\phi(x')dsdydxdx'$$

$$
\begin{aligned}
&= \int_0^t \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} G(t-s, y-x)\phi(x)dx \Big) \Big(\int_{\mathbb{R}^d} G(t-s, y-x')\phi(x')dx' \Big) ds \\
&= \int_0^t \int_{\mathbb{R}^d} v^2(t-s, y)dyds = \int_0^t \int_{\mathbb{R}^d} v^2(s, y)dyds.
\end{aligned}
$$

Here, the function $v(s, y)$ is the solution of the heat equation

$$
\frac{\partial v}{\partial t} = \Delta v, \tag{1.9}
$$

with the initial condition $v(0, y) = \phi(y)$. As $\phi \in C_c^\infty(\mathbb{R}^d)$, the function $v(t, x)$ is most beautifully smooth and rapidly decaying. It follows that $\Phi(t)$ has a finite second moment, meaning that it is likely one can make sense of (1.5) to make sense as a distribution in any dimension. Thus, in dimensions $d \geq 2$ one, generally, would expect solutions of SPDEs to be distributions and not functions.

This is a problem as we will often be interested in solutions of nonlinear SPDEs, and we do not know how to take nonlinear functions of distributions. In particular, SPDEs often arise as limit descriptions of the densities of systems of N particles. It is typical that in such models the particle density converges to the solution of a deterministic PDE as $N \to +\infty$, such as, in the simplest case, the heat equation:

$$
\frac{\partial u}{\partial t} = \Delta u. \tag{1.10}
$$

Accounting for a large but finite number of particles often leads to an SPDE that is a perturbation of the limiting deterministic problem, such as

$$
\frac{\partial u}{\partial t} = \Delta u + \text{Noise}. \tag{1.11}
$$

In that setting, the noise term typically has variance proportional, locally, to the total number of particles, that is, to $u(t, x)$ – this is a version of the central limit theorem Thus, the equation would have the form

$$
\frac{\partial u}{\partial t} = \Delta u + \sqrt{u} \cdot \text{Noise}. \tag{1.12}
$$

Hence, we would need to deal not only with nonlinearities but also with non-Lipschitz nonlinearities.

The randomly forced wave equation

Let us, as a next example, informally, consider the wave equation

$$
\frac{\partial^2 v}{\partial t^2} - \frac{\partial^2 v}{\partial x^2} = F(t, x), \tag{1.13}
$$

with the initial condition

$$
v(0, x) = \frac{\partial v(0, x)}{\partial t} = 0.
$$

Its solution is

$$v(t,x) = \frac{1}{2}\int_0^t \int_{x-(t-s)}^{x+(t-s)} F(s,y)dyds. \tag{1.14}$$

Let us now assume that $F(s,y)$ is a white noise, as i the heat equation example. Then, the solution $v(t,x)$ is the value of the white noise as a distribution on the characteristic function of the triangle, which is the domain of integration in (1.14) – therefore, we arrive at the same need to understand the white noise $F(t,x)$ as a distribution. In higher dimensions, similar expressions for the solution of the forced wave equation can be obtained via spherical means.

Gaussian processes

We now start being more careful than in the above informal discussion. A stochastic process $G(t)$, $t \in T$, indexed by a set T is a Gaussian random field if for every finite collection $t_1,\dots,t_k \in T$, the vector $(G(t_1),\dots,G(t_k))$ is a Gaussian random vector. The Kolmogorov consistency theorem implies that the finite-dimensional distributions of G are uniquely determined by the mean:

$$\mu(t) = \mathbb{E}(G(t)),$$

and the covariance

$$C(s,t) = \mathbb{E}[(G(s)-\mu(s))(G(t)-\mu(t))].$$

The covariance function is non-negative definite in the following sense: for any $t_1,\dots,t_k$ and $z_1,\dots,z_k \in \mathbb{C}$, we have

$$\sum_{j,m=1}^k C(t_j,t_m)z_j\bar{z}_m \geq 0. \tag{1.15}$$

This is because

$$0 \leq \mathbb{E}\Big|\sum_{j=1}^k (G(t_j)-\mu(t_j))z_j\Big|^2 = \sum_{j,m=1}^k C(t_j,t_m)z_j\bar{z}_m.$$

A classical result of Kolmogorov is that given any function $\mu(t)$ and a nonnegative-definite function $C(s,t)$ one can construct a Gaussian random field $G(t)$ with mean $\mu(t)$ and covariance $C(s,t)$.

Example 1: the Brownian motion. One of the most basic examples of a Gaussian process is the Brownian motion. In that case, $T = [0,+\infty)$, $\mu(t) \equiv 0$, and $C(s,t) = \min(s,t)$. Let us check that this covariance is nonnegative-definite:

$$\sum_{j,m=1}^k z_j\bar{z}_m \min(t_j,t_m) = \sum_{j,m=1}^k z_j\bar{z}_m \int_0^{+\infty} \chi_{[0,t_j]}(s)\chi_{[0,t_m]}(s)ds = \int_0^{+\infty} \Big|\sum_{j=1}^k z_j\chi_{[0,t_j]}(t)\Big|^2 dt \geq 0.$$

Of course, the Brownian motion can be constructed in many other ways, without invoking general abstract theorems.

Example 2: the Brownian bridge. The Brownian bridge $b(t)$ is a process on the interval $T = [0,1]$, with mean-zero: $\mu(t) = 0$, and the covariance

$$C(s,t) = \mathbb{E}(b(s)b(t)) = s \wedge t - st. \tag{1.16}$$

A remarkable property of the Brownian bridge is that

$$\mathbb{E}(b(0)^2) = \mathbb{E}(b(1)^2) = 0. \tag{1.17}$$

That is, the process $b(s)$ starts at $b(0) = 0$ and ends at $b(1) = 0$, almost surely. Instead of verifying directly that the function $C(s,t)$ in (1.16) is nonnegative-definite, we observe that if $B(t)$ is a Brownian motion, then

$$b(t) = B(t) - tB(1) \tag{1.18}$$

is a Brownian bridge. Indeed, it clearly has mean-zero, and for any $0 \le s.t \le 1$, we have

$$\mathbb{E}(b(s)b(t)) = \mathbb{E}[(B(s) - sB(1))(B(t) - tB(1))] = s \wedge t - st - ts + st = s \wedge t - st.$$

Example 3: the Ornstein-Uhlenbeck process. The Brownian motion does not have a finite invariant measure – it typically runs away to infinity. In order to confine it, let us define the process

$$X(t) = e^{-t/2}B(e^t), \tag{1.19}$$

for $t \ge 0$. The process $X(t)$ is mean-zero, and has the covariance for $0 \le t \le s$:

$$C(s,t) = e^{-(s+t)/2} \min\left(e^s, e^t\right) = e^{(t-s)/2}. \tag{1.20}$$

In other words, we have, for all $t \ge 0$ and $s \ge 0$:

$$C(s,t) = e^{-|t-s|/2}. \tag{1.21}$$

In particular, $C(s,t)$ depends only on $|t-s|$ – such processes are called stationary Gaussian processes. We also have

$$\mathbb{E}(X^2(t)) = 1, \tag{1.22}$$

for all $t \ge 0$, which indicates that $X(t)$ is, indeed, confined in some sense.

Example 4: a general white noise. In general, we define a white noise as follows. Let E be a set endowed with a measure ν and a collection $\mathcal{M}$ of measurable sets. Then a white noise is a random function on the sets $A \in \mathcal{M}$ of a finite ν-measure such that $\dot{W}(A)$ is a mean-zero Gaussian random variable with

$$\mathbb{E}(\dot{W}(A)^2) = \nu(A),$$

and if $A \cap B = \emptyset$, then $\dot{W}(A)$ and $\dot{W}(B)$ are independent, with

$$\dot{W}(A \cup B) = \dot{W}(A) + \dot{W}(B).$$

Then the covariance of $\dot{W}$ is

$$\mathbb{E}(\dot{W}(A)\dot{W}(B)) = \mathbb{E}[(\dot{W}(A\cap B) + \dot{W}(A\backslash B))(\dot{W}(A\cap B) + \dot{W}(B\backslash A))] = \mathbb{E}[\dot{W}(A\cap B)^2] = \nu(A\cap B).$$

This function is nonnegative-definite because

$$\sum_{i,j=1}^{k} z_i\bar{z}_j C(A_i, A_j) = \sum_{i,j=1}^{k} z_i\bar{z}_j \nu(A_i \cap A_j) = \sum_{i,j=1}^{k} \int_E z_i\bar{z}_j \chi_{A_i}(x)\chi_{A_j}(x) d\nu(x)$$
$$= \int_E \Big|\sum_{j=1}^{k} z_j \chi_{A_j}(x)\Big|^2 d\nu(x) \geq 0.$$

Example 5: a Brownian sheet. The Brownian motion $W(t)$ can now be alternatively defined as follows. Take $E = \mathbb{R}_+$, and ν the Lebesgue measure on $\mathbb{R}$, and set $W(t) = \dot{W}([0,t])$. This gives $\mathbb{E}(W(t)) = 0$, and

$$\mathbb{E}(W(t)^2) = |[0,t] \cap [0,s]| = s \wedge t.$$

A Brownian sheet $W(t)$, with $t \in \mathbb{R}^n_+$, can be defined as above, taking $E = \mathbb{R}^n_+$, and ν the Lebesgue measure on $\mathbb{R}^n$. For $t = (t_1, \ldots, t_n) \in \mathbb{R}^n$, we set

$$[0,t] = [0,t_1] \times \ldots \times [0,t_n],$$

and define the Brownian sheet as

$$W(t) = \dot{W}([0,t]).$$

Exercise. Consider $n = 2$ and denote $t = (t_1, t_2)$.
(i) Show that if t_1 is fixed, then $W_{t_1,t}$ is a Brownian motion.
(ii) Show that on the hyperbole $t_1 t_2 = 1$ we have that

$$X_t = W_{e^t, e^{-t}}$$

is an Ornstein-Uhlenbeck process.
(iii) Show that on the diagonal the process $W_{t,t}$ is a martingale, has independent increments but is not a Brownian motion.

The white noise does not have a bounded total variation

Let us note that $\dot{W}$ does not have a bounded total variation almost surely. The proof is very much as what we do in the proof of the Ito formula. To see this, we first show that

$$\lim_{n\to\infty} \sum_{j=0}^{2^n-1} \Big|\dot{W}\Big(\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big)\Big|^2 = 1, \tag{1.23}$$

almost surely. Indeed, let us set

$$S_n = \sum_{j=0}^{2^n-1} \Big|\dot{W}\Big(\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big)\Big|^2,$$

then

$$\mathbb{E}S_n = \sum_{j=0}^{2^n-1} \Big|\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big| = 1,$$

and

$$\mathbb{E}(S_n-1)^2 = \mathbb{E}\Big(\sum_{j=0}^{2^n-1}\Big[\Big|\dot W\Big(\Big[\frac{j}{2^n},\frac{j+1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[\frac{j}{2^n},\frac{j+1}{2^n}\Big]\Big|\Big]\Big)^2$$

$$= \sum_{j,m=0}^{2^n-1}\mathbb{E}\Big\{\Big[\Big|\dot W\Big(\Big[\frac{j}{2^n},\frac{j+1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[\frac{j}{2^n},\frac{j+1}{2^n}\Big]\Big|\Big]\Big[\Big|\dot W\Big(\Big[\frac{m}{2^n},\frac{m+1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[\frac{m}{2^n},\frac{m+1}{2^n}\Big]\Big|\Big]\Big\}$$

$$= \sum_{j=0}^{2^n-1}\mathbb{E}\Big\{\Big[\Big|\dot W\Big(\Big[\frac{j}{2^n},\frac{j+1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[\frac{j}{2^n},\frac{j+1}{2^n}\Big]\Big|\Big]^2\Big\} = 2^n\mathbb{E}\Big\{\Big[\Big|\dot W\Big(\Big[0,\frac{1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[0,\frac{1}{2^n}\Big]\Big|\Big]^2\Big\}$$

$$= 2^n\Big(3\cdot\frac{1}{2^{2n}} - \frac{2}{2^{2n}} + \frac{1}{2^{2n}}\Big) = \frac{1}{2^{n-1}}.$$

We used the stationarity of the white noise in the next to last step above, and the fact that for a Gaussian random variable X we have

$$\mathbb{E}(X^4) = 3(\mathbb{E}(X^2))^2$$

in the last step. It follows that

$$\mathbb{P}(|S_n-1|>\varepsilon) \le \frac{1}{2^{n-1}\varepsilon^2}.$$

The Borel-Cantelli lemma implies that for every $\varepsilon>0$ the event $\{|S_n-1|>\varepsilon\}$ occurs only for finitely many n, almost surely. Thus, $S_n\to 1$ almost surely, or (1.23) holds.
Exercise. Use (1.23) to obtain

$$\lim_{n\to\infty}\sum_{j=0}^{2^n-1}\Big|\dot W\Big(\Big[\frac{j}{2^n},\frac{j+1}{2^n}\Big]\Big)\Big| = +\infty, \tag{1.24}$$

also almost surely. It follows that $\dot W$ does not have a bounded total variation almost surely.

2 Regularity of random processes

We will now prove the Kolmogorov theorem that reduces the question of continuity of a stochastic process to a computation of some moments. Recall that a process $X'(t)$, $t\in\mathbf{T}$, is a modification of a process $X(t)$, $t\in\mathbf{T}$ if

$$P[X'(t)=X(t)] = 1 \text{ for all } t\in\mathbf{T}.$$

Exercise. Construct an example where X' is a modification of X but

$$P[X'(t)=X(t) \text{ for all } t\in\mathbf{T}] = 0.$$

Modulus of continuity from an integral inequality

We first prove a real-analytic result that allows to translate an integral bound on a function into a point-wise modulus of continuity. We need some notation. In the theorem below, the functions $\psi(x)$ and $p(x)$, $x \in \mathbb{R}$, are even, $p(x)$ is increasing for $x > 0$, with $p(0) = 0$, and $\psi(x)$ is convex. We denote by R_1 the unit cube in $\mathbb{R}^d$.

Theorem 2.1 *Let f be a measurable function on $R_1 \subset \mathbb{R}^d$ such that*

$$B := \int_{R_1}\int_{R_1} \psi\Big(\frac{f(y)-f(x)}{p(|y-x|/\sqrt{d})}\Big)dxdy < +\infty, \tag{2.1}$$

then there is a set K of measure zero such that if $x, y \in R_1 \setminus K$, then

$$|f(y)-f(x)| \le 8\int_0^{|y-x|} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big)dp(u). \tag{2.2}$$

If f is continuous, then (2.2) holds for all x and y.

Proof. We denote the side of a cube Q in R_1 by $e(Q)$. Note that

$$\text{if } x, y \in Q, \text{ then } |y-x| \le \sqrt{d}e(Q). \tag{2.3}$$

The functions p and ψ are increasing for positive arguments, hence (2.1) and (2.3) imply

$$\int_Q\int_Q \psi\Big(\frac{f(y)-f(x)}{p(e(Q))}\Big)dxdy \le B, \tag{2.4}$$

for any cube Q in R_1. Next, take a nested sequence of cubes $Q_0 \supseteq Q_1 \supseteq Q_2 \supseteq \dots$ such that

$$p(e(Q_j)) = \frac{1}{2}p(e(Q_{j-1})), \tag{2.5}$$

and denote

$$f_j = \frac{1}{|Q_j|}\int_{Q_j} fdx, \quad r_j = e(Q_j).$$

It follows from (2.5) that $r_j \to 0$, and the cubes converge down to a point. As the function ψ is convex, we have

$$\begin{aligned}\psi\Big(\frac{f_j-f_{j-1}}{p(r_{j-1})}\Big) &\le \frac{1}{|Q_{j-1}|}\int_{Q_{j-1}} \psi\Big(\frac{f_j-f(x)}{p(r_{j-1})}\Big)dx \\ &\le \frac{1}{|Q_{j-1}||Q_j|}\int_{Q_{j-1}}\int_{Q_j} \psi\Big(\frac{f(y)-f(x)}{p(r_{j-1})}\Big)dxdy \\ &\le \frac{1}{|Q_{j-1}||Q_j|}\int_{Q_{j-1}}\int_{Q_{j-1}} \psi\Big(\frac{f(y)-f(x)}{p(r_{j-1})}\Big)dxdy \le \frac{B}{|Q_{j-1}||Q_j|},\end{aligned}$$

by (2.4). We conclude that

$$|f_j - f_{j-1}| \le p(r_{j-1})\psi^{-1}\Big(\frac{B}{|Q_{j-1}||Q_j|}\Big). \tag{2.6}$$

This starts to look like a modulus of continuity estimate but the two terms in the right side still compete – one is small, the other is large. We now re-write it to make it look like the right side of (2.2). The definition (2.5) of Q_j means that

$$p(r_{j-1}) = 4|p(r_{j+1}) - p(r_j)|.$$

hence we may write

$$|f_j - f_{j-1}| \le 4\psi^{-1}\Big(\frac{B}{|Q_{j-1}||Q_j|}\Big)|p(r_{j+1}) - p(r_j)|. \tag{2.7}$$

Next, note that for $r_{j+1} \le u \le r_j$, we have $|Q_{j-1}||Q_j| \ge u^{2d}$, hence

$$|f_j - f_{j-1}| \le 4\psi^{-1}\Big(\frac{B}{u^{2d}}\Big)|p(r_{j+1}) - p(r_j)|, \quad \text{for all } r_{j+1} \le u \le r_j. \tag{2.8}$$

We deduce that

$$|f_j - f_{j-1}| \le 4\int_{r_{j+1}}^{r_j} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big)dp(u). \tag{2.9}$$

Summing over j gives

$$\limsup_{j\to+\infty} |f_j - f_0| \le 4\int_0^{r_0} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big)dp(u). \tag{2.10}$$

Now, by the Lebesgue theorem, except for x in an exceptional set K of measure zero, the sequence f_j converges to $f(x)$ for any sequence of cubes Q_j decreasing to the point x. If x and y are not in K, and Q_0 is the smallest cube containing both x and y, then, as $r_0 \le |x-y|$, we have both

$$|f(x) - f_0| \le 4\int_0^{|x-y|} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big)dp(u), \tag{2.11}$$

and

$$|f(y) - f_0| \le 4\int_0^{|x-y|} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big)dp(u), \tag{2.12}$$

proving (2.2). □

The Kolmogorov theorem

We may now apply Theorem 2.1 to various stochastic processes. We begin with the Kolmogorov theorem.

Theorem 2.2 *Let X_t, $t \in \mathbf{T} = [a_1, b_1] \times \ldots, [a_d, b_d] \subset \mathbb{R}^d$ be a real-valued stochastic process. Suppose there are constants $k > 1$, $C > 0$ and $\varepsilon > 0$ so that for all $s, t \in \mathbf{T}$, we have*

$$\mathbb{E}(|X(t) - X(s)|^k) \le C|t-s|^{d+\varepsilon}. \tag{2.13}$$

Then $X(t)$ has a continuous modification $\bar{X}(t)$. Moreover, $X(t)$ has the following modulus of continuity:

$$|X(t) - X(s)| \le Y|t-s|^{\varepsilon/k}\Big(\log\frac{c_1}{|t-s|}\Big)^{2/k}, \tag{2.14}$$

with a deterministic constant $c_1 > 0$*, and a random variable* Y *such that* $\mathbb{E}(Y^k) \le C'$*. Finally, if* $\mathbb{E}(|X_t|^k) < +\infty$ *for some* t*, then*

$$\mathbb{E}(\sup_{t\in\mathbf{T}} |X_t|^k) < +\infty.$$

Proof. Without loss of generality we will assume that $\mathbf{T}$ is the unit cube Q_1. We will use Theorem 2.1 with $\psi(x) = |x|^k$, and

$$p(x) = |x|^{(2d+\varepsilon)/k}\Big(\log\frac{c_1}{|x|}\Big)^{2/k}.$$

This function is increasing on $[0, \sqrt{d}]$ with an appropriately large choice of c_1. The function f in Theorem 2.1 is taken to be $X(t;\omega)$, for a fixed realization ω. This gives

$$B(\omega) = \int_{Q_1}\int_{Q_1} \frac{|X(t;\omega) - X(s;\omega)|^k}{[p(|t-s|/\sqrt{d})]^k} dtds = C\int_{Q_1}\int_{Q_1} \frac{|X(t;\omega) - X(s;\omega)|^k}{|t-s|^{2d+\varepsilon}\log^2(c_1/|t-s|)} dtds.$$

Taking the expectation, and using (2.13), we obtain

$$\mathbb{E}(B) \le C\int_{Q_1}\int_{Q_1} \frac{\mathbb{E}|X(t;\omega) - X(s;\omega)|^k}{|t-s|^{2d+\varepsilon}\log^2(c_1/|t-s||)} dtds \le C\int_{Q_1}\int_{Q_1} \frac{1}{|t-s|^{d}\log^2(c_1/|t-s|)} dtds. \tag{2.15}$$

The integral in the right, for a fixed t. and when c_1 is sufficiently large, behaves as

$$\int_0^{1/2} \frac{r^{n-1}}{r^n \log^2 r} dr = \int_{\log 2}^{\infty} \frac{dr}{r^2} < +\infty.$$

Therefore, $\mathbb{E}(B) < \infty$, and B is finite almost surely. Going back to (2.2) we get

$$|X(t) - X(s)| \le 8B^{1/k}\int_0^{|t-s|} \frac{1}{u^{2d/k}} p'(u)du. \tag{2.16}$$

It is now a calculus an exercise to check that (2.14) holds with $Y = CB^{1/k}$. The last statement in the theorem is an immediate consequence of the modulus of continuity estimate (2.14) and the moment bound on Y. □

Regularity of a Gaussian process

Theorem 2.3 *Let* X_t*,* $t \in Q_1 \subset \mathbb{R}^d$ *be a mean zero Gaussian process, and set*

$$p(u) = \max_{|s-t|\le|u|/\sqrt{d}} \mathbb{E}\{|X_t - X_s|^2\}^{1/2}. \tag{2.17}$$

If

$$\int_0^1 \Big(\log\frac{1}{u}\Big)^{1/2} dp(u) < +\infty,$$

then X *has a continuous version with the modulus of continuity*

$$m(\delta) \le C\int_0^{\delta} \Big(\log\frac{1}{u}\Big)^{1/2} dp(u) + Yp(\delta), \tag{2.18}$$

with a universal constant $C > 0$ *and a random variable* Y *such that* $\mathbb{E}(\exp(c_1 Y^2)) < +\infty$ *for a sufficiently small universal constant* $c_1 > 0$*.*

Proof. We will use Theorem 2.1 with $\psi(x) = e^{|x|^2/4}$, and $p(u)$ as in (2.17). The function $p(u)$ has to satisfy the assumption of the aforementioned theorem – otherwise, the modulus of continuity (2.18) becomes vacuous. Note that (2.17) implies that

$$D_{st} = \frac{X_s - X_t}{p(|t-s|/\sqrt{d})}$$

is a mean zero Gaussian variable with variance $\sigma_{st} \le 1$. It follows that

$$\mathbb{E}(B) = \int_{Q_1}\int_{Q_1} \mathbb{E}(\exp(D_{st}^2/4))dsdt \le C < +\infty.$$

Therefore, $B < +\infty$ a.s., and the rest is a direct application of Theorem 2.1 and integration by parts, as in the proof of Theorem 2.2. Note that in our case, for $u \in (0,1)$ we have

$$\psi^{-1}(u) = 2\sqrt{\log(1/u)},$$

which explains the modulus of continuity in (2.18). The same computation shows that the random variable Y can be taken as $Y = C\sqrt{|\log B|}$. □

A useful example is when $p(u) = |u|^\varepsilon$, that is, if we assume that a Gaussian process satisfies

$$\mathbb{E}\{|X_t - X_s|^2\} \le C|t-s|^{2\varepsilon}.$$

Then (2.18) says that $X(t)$ is Hölder continuous for any exponent smaller than ε.

3 Stochastic Integrals

In order to talk about the weak solutions of the stochastic partial differential equations, we need to define carefully stochastic integration with respect to white noises.

The Ito integral

Before talking about stochastic integrals with respect to space-time white noises $\dot{W}(t,x)$, let us very briefly recollect the steps done in the definition of the Ito integration, which is what we will generalize to higher dimensions. When we define the Ito integral with respect the Brownian motion

$$\int_0^t f(s,\omega)dB_s,$$

this is first done for elementary functions of the form

$$f(t,\omega) = X(\omega)\chi_{[a,b]}(t).$$

Here, $X(\omega)$ is an $\mathcal{F}_a$-measurable function – recall that this innocent sounding assumption is absolutely essential for the Ito integral (as opposed to the Stratonovich and other stochastic integrals) to be a martingale, with a finite second moment:

$$\mathbb{E}[X^2] < +\infty.$$

For such elementary functions we define the Ito integral as

$$\int_0^t f(s,\omega)dB_s = \begin{cases} 0, & \text{if } 0 < t < a, \\ X(\omega)(B_t - B_a), & \text{if } a < t < b, \\ X(\omega)(B_b - B_a), & \text{if } b < t. \end{cases} \tag{3.1}$$

This may be written more succinctly as

$$\int_0^t f(s,\omega)dB_s = X(\omega)[B_{t\wedge b} - B_{t\wedge a}]. \tag{3.2}$$

This expression can be immediately generalized to simple functions – these are linear combinations of elementary functions f_j

$$f(t,\omega) = \sum_{j=1}^{n} c_j f_j(t,\omega),$$

with deterministic constants c_j. The main observation that allows to go further and define the Ito integral for more general functions is the Ito isometry: it is easy to check from the above definition that for a simple function $f(t,\omega)$ we have

$$\mathbb{E}\Big(\int_0^t f(s,\omega)dB_s\Big)^2 = \mathbb{E}(X^2(\omega))(t\wedge b - t\wedge a) = \int_0^t \mathbb{E}(f^2(s,\omega))ds. \tag{3.3}$$

Then one can verify that simple functions are dense in $L^2(\Omega\times[0,T])$ and define the stochastic integral for all such functions as an object in $L^2(\Omega)$. This is also where we will construct the stochastic integral with respect to space-time white noises.

Another important aspect of the Ito integral is that, as it is defined for $\mathcal{F}_t$-adapted functions f, the integral

$$I_t = \int_0^t f(s,\omega)dB_s$$

is a martingale. Its quadratic variation is

$$\langle I, I\rangle_t = \int_0^t f^2(s,\omega)ds.$$

This follows from the Ito formula:

$$d(I_t^2) = f^2(t,\omega)dt + 2I_t f(t,\omega)dB_t.$$

Martingale measures

In order to construct the stochastic integral with respect to a white noise, we need the notion of a σ-finite L^2-valued measure. This is defined as follows. Let E be a subset of $\mathbb{R}^d$, and $\mathcal{B}$ be an algebra of measurable sets. Typically, unless specified otherwise, we will take E as $\mathbb{R}^d$ and $\mathcal{B}$ as the collection of the Borel sets. Let U be a real-valued random set function on $\mathcal{B}$. We denote by

$$\|U(A)\|_2 = (\mathbb{E}(U^2(A))^{1/2}.$$

Assume there exists an increasing sequence of measurable sets E_n such that

$$E = \bigcup_n E_n,$$

and

$$\sup\{\|U(A)\|_2 : \; A \subset E_n\} < +\infty, \text{ for all } n.$$

Then we say that the function U is σ-finite. It is countably additive if, for each n, given that $A_j \subset E_n$ and A_j is a decreasing sequence of sets with an empty intersection, then

$$\lim_{j \to +\infty} U(A_j) = 0.$$

Then we say that U is a σ-finite L^2-valued measure. It is easy to see that the white noise is an example of an L^2-valued σ-finite measure, as

$$\|\dot{W}(A)\|_2 = |A|^{1/2}.$$

We will now split one variable in the noise, and call it $t \geq 0$, while keeping all other variables as "spatial variables". A process $M_t(A)$, $A \in \mathcal{B}$ is a martingale measure if

(i) $M_0(A) = 0$ a.s., for all $A \in \mathcal{B}$.

(ii) For $t > 0$ fixed, $M_t(A)$ is a σ-finite L^2-valued measure.

(iii) For all $A \in \mathcal{B}$ fixed, $M_t(A)$ is a mean-zero martingale.

Let us check that the white noise process $W_t(A) = \dot{W}([0,t] \times A)$ is a martingale measure. First, we have $W_0(A) = 0$ a.s. because

$$\mathbb{E}[W_t(A)]^2 = t|A|.$$

This also implies that $W_t(A)$ is a σ-finite L^2-valued measure. Finally, to see that $W_t(A)$ is a martingale for each $A \in \mathcal{B}$ fixed, we observe that for all $t \geq s \geq u \geq 0$ we have

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

It follows that the increment $W_t(A) - W_s(A)$ is independent of $\mathcal{F}_s$, the σ-algebra generated by $W_r(A)$, with $0 \leq r \leq s$, $A \in \mathcal{B}$. Hence, we have

$$\mathbb{E}(W_t(A)|\mathcal{F}_s) = \mathbb{E}(W_t(A) - W_s(A)|\mathcal{F}_s) + W_s(A) = \mathbb{E}(W_t(A) - W_s(A)) + W_s(A) = W_s(A),$$

and $W_t(A)$ is a martingale.

The stochastic integral for simple functions

In order to define the stochastic integration we begin with the simple functions, as for the Ito integral. We say that a function $f(t,x,\omega)$ is elementary if it has the form

$$f(t,x,\omega) = X(\omega)\mathbb{1}_{(a,b]}(t)\mathbb{1}_A(x). \tag{3.4}$$

Here, A is a Borel set, and the random variable X is bounded and $\mathcal{F}_a$-measurable – the latter condition is very important, as it was for the Ito integral. A simple function is a linear combination of finitely many elementary functions (with deterministic coefficients). We will denote by $\mathcal{P}$ the σ-algebra generated by all simple functions. It is called the predictable σ-algebra.

Given an elementary function f, we define the stochastic-integral process of f as

$$(f\cdot M)_t(B)(\omega) = X(\omega)[M_{t\wedge b}(A\cap B) - M_{t\wedge a}(A\cap B)](\omega). \tag{3.5}$$

On the informal level, this agrees with

$$\int_0^t \int_B f dM(s,x) = X(\omega)\int_0^t\int_B \chi_{(a,b]}(s)\chi_A(x)dM(s,x)$$
$$= \begin{cases} 0, & \text{if } 0<t<a<b, \\ X(\omega)[M_t(A\cap B) - M_a(A\cap B)], & \text{if } 0<a<t<b, \\ X(\omega)[M_b(A\cap B) - M_a(A\cap B)], & \text{if } 0<a<b<t, \end{cases}$$

This is a direct generalization of (3.1)-(3.2) for the Ito integral. We can extend the definition (3.5) to simple functions in a straightforward way as linear combinations. Note that if f is a simple function and $M_t(A)$ is a martingale measure then $f\cdot M_t$ is a martingale measure as well.

Let us now compute the second moment of the stochastic integral process of an elementary function (3.4). We take a Borel set B and find

$$\mathbb{E}\Big[((f\cdot M_t)(B))^2\Big] = \mathbb{E}\Big[X^2[M_{t\wedge b}(A\cap B) - M_{t\wedge a}(A\cap B)]^2\Big] \tag{3.6}$$
$$= \mathbb{E}[X^2M^2_{t\wedge b}(A\cap B)] + \mathbb{E}[X^2M^2_{t\wedge a}(A\cap B)] - 2\mathbb{E}[X^2M_{t\wedge b}(A\cap B)M_{t\wedge a}(A\cap B)].$$

Let us recall that X is $\mathcal{F}_a$-measurable. Hence, by the definition of the quadratic variation we have

$$\mathbb{E}\Big[X^2(M^2_{t\wedge b}(A\cap B) - \langle M(A\cap B), M(A\cap B)\rangle_{t\wedge b})\Big] \tag{3.7}$$
$$= \mathbb{E}\Big[X^2(M^2_{t\wedge a}(A\cap B) - \langle M(A\cap B), M(A\cap B)\rangle_{t\wedge a})\Big],$$

and, since M_t is martingale measure, we also have

$$\mathbb{E}\Big[X^2(M_{t\wedge b}(A\cap B)M_{t\wedge a}(A\cap B) = \mathbb{E}\Big[X^2(M^2_{t\wedge a}(A\cap B))\Big]. \tag{3.8}$$

Using this in (3.6) gives

$$\mathbb{E}\Big[((f\cdot M_t)(B))^2\Big] = \mathbb{E}\Big[X^2(M^2_{t\wedge a}(A\cap B) - \langle M(A\cap B), M(A\cap B)\rangle_{t\wedge a})\Big] \tag{3.9}$$
$$+\mathbb{E}\Big[X^2\langle M(A\cap B), M(A\cap B)\rangle_{t\wedge b})\Big] + \mathbb{E}[X^2M^2_{t\wedge a}(A\cap B)] - 2\mathbb{E}\Big[X^2(M^2_{t\wedge a}(A\cap B))\Big]$$
$$= \mathbb{E}\Big[X^2(\langle M(A\cap B), M(A\cap B)\rangle_{t\wedge b} - \langle M(A\cap B), M(A\cap B)\rangle_{t\wedge a})\Big]. \tag{3.10}$$

Let us define the covariance functional of a martingale measure M_t as

$$\bar{Q}_t(A,B) = \langle M(A), M(B)\rangle_t. \tag{3.11}$$

Note that $\bar{Q}_t(A,B)$ is symmetric in A and B:

$$\bar{Q}_t(A,B) = \bar{Q}_t(B,A).$$

Moreover, if B and C are disjoint sets, then

$$\begin{aligned}\bar{Q}_t(A, B\cup C) &= \langle M(A), M(B\cup C)\rangle_t = \langle M(A), M(B)+M(C)\rangle_t \\ &= \langle M(A), M(B)\rangle_t + \langle M(A), M(C)\rangle_t = \bar{Q}_t(A,B)+\bar{Q}_t(A,C).\end{aligned}$$

Finally, it satisfies the Cauchy inequality:

$$|\bar{Q}_t(A,B)|^2 \le \bar{Q}_t(A,A)\bar{Q}_t(B,B).$$

With this notation, going back to (3.9), we can write

$$\mathbb{E}\Big[((f\cdot M_t)(B))^2\Big] = \mathbb{E}\Big(X^2[Q_{t\wedge b}(A\cap B, A\cap B) - Q_{t\wedge b}(A\cap B, A\cap B)]\Big). \tag{3.12}$$

In other words, for simple function we have

$$\mathbb{E}\Big[\int_0^t\int_B f dM(s,x)\Big]^2 = \mathbb{E}\Big[\int_0^t\int_{B\times B} f(s,x)f(s,y)Q(dxdyds)\Big]. \tag{3.13}$$

Here, we have defined

$$Q([s,t]\times A\times B) = \bar{Q}_t(A,B) - \bar{Q}_s(A,b). \tag{3.14}$$

Hence, the potential majorizer for the L^2-estimate on the stochastic integral is $Q(dxdyds)$. As the quadratic variation has bounded total variation, it is a good integrator. For the white noise we may compute the quadratic variation explicitly: take a set $A\in\mathcal{B}$, and set $W_t(A) = \dot{W}([0,t]\times A)$. We claim that in this case

$$\bar{Q}_t(A,B) = t|A\cap B|. \tag{3.15}$$

Indeed, let $X(\omega)$ be $\mathcal{F}_s$ measurable, and consider, for $t>s$:

$$\begin{aligned}&\mathbb{E}((W_t(A)^2 - t|A|)X) = \mathbb{E}[((W_s(A))^2 - s|A|)X] + \mathbb{E}[((W_t(A)-W_s(A))^2 - (t-s)|A|)X] \\ &+2\mathbb{E}[W_s(A)(W_t(A)-W_s(A))X)] = \mathbb{E}[((W_s(A))^2 - s|A|)X],\end{aligned}$$

hence

$$\bar{Q}_t(A,A) = t|A|. \tag{3.16}$$

On the other hand, for disjoint A and B, we know that

$$\bar{Q}_t(A,B) = 0, \tag{3.17}$$

since $M_t(A)$ and $M_t(B)$ are martingales that have increments independent of each other, so that $M_t(A)M_t(B)$ is also a martingale. It follows that for a general pair of sets A and B we may write

$$\begin{aligned}\bar{Q}_t(A,B) &= \langle M(A), M(B)\rangle_t = \langle (M(A\setminus B) + M(A\cap B)), (M(B\setminus A) + M(A\cap B)\rangle_t \\ &= \langle M(A\cap B), M(A\cap B)\rangle_t = t|A\cap B|,\end{aligned}$$

which is (3.15). The measure (3.14) that corresponds to the white noise is, therefore, simply

$$Q([s,t]\times A\times B) = |t-s||A\cap B|. \tag{3.18}$$

This may be written formally as

$$Q(dxdydt) = \delta(x-y)dxdydt, \tag{3.19}$$

that was our intent from the very beginning!

The general construction of a stochastic integral proceeds for the so-called worthy measures. These are the measures such that $Q(dxdydt)$ defined via (3.14) can be majorized by a positive-definite signed measure $K(dxdydt)$. We will for simplicity of notation restrict ourselves the stochastic integral with respect to the white noise.

The stochastic integral for predictable functions

Let us recall that we denote by $\mathcal{P}$ the σ-algebra generated by the simple functions. A function is predictable if it is $\mathcal{P}$-measurable. We can define the norm for predictable functions (keep in mind that we are only considering the white noise case here) as

$$\|f\|^2 = \mathbb{E}\Big(\int_0^T\int_{\mathbb{R}^d}|f(t,x,\omega)|^2dxdt\Big). \tag{3.20}$$

We will denote by P_2 the space of predictable function of a finite norm (3.20). It is an exercise to verify that P_2 is a Banach space. Another exercise shows that the simple functions are dense in P_2. Let us go back to (3.13), which for the white noise takes the form

$$\mathbb{E}\Big[\int_0^t\int_B f(s,x,\omega)W(ds,dx)\Big]^2 = \mathbb{E}\Big[\int_0^t\int_{B\times B}|f(s,x,\omega)|^2dxds\Big]. \tag{3.21}$$

Here, f is a simple function but this allows us to generalize the notion of the stochastic integral to functions in P_2. Indeed, if f_n is a Cauchy sequence of simple functions in P_2, then (3.21) shows that the sequence

$$\int_0^t\int_B f(s,x,\omega)W(ds,dx) \tag{3.22}$$

is Cauchy in $L^2(P)$. Hence, for any function $f\in P_2$ we may define the stochastic integral (3.22) as the limit in $L^2(P)$ of the

$$\int_0^t\int_B f_n(s,x,\omega)W(ds,dx), \tag{3.23}$$

where f_n is a sequence of simple functions in P_2 that converges to f in P_2. This is essentially the same procedure as in the definition of the usual Ito integral.

4 The stochastic heat equation with a Lipschitz nonlinearity: the basic theory

We now consider a very basic example of a parabolic SPDE

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{4.1}$$

posed on $\mathbb{R}$, with the initial condition $u(0,x) = u_0(x)$. The function $u_0(x)$ is deterministic and compactly supported. The nonlinearity $f(u)$ is globally Lipschitz:

$$|f(u_1) - f(u_2)| \leq K|u_1 - u_2|. \tag{4.2}$$

This assumption is extremely important both for u small and u large: the "interesting cases" are what happens when $f(u) \sim \sqrt{u}$ for small u – this will lead to compactly supported solutions, and when $f(u) \sim u^2$ for u large – this may lead to blow-up of solutions in a finite time. For now, we deliberately avoid both, and stay within the realm of Lipschitz nonlinearities for simplicity, but will come back to them later. It is sometimes helpful to assume that f is, in addition, bounded but we will avoid this for the moment, as we would like to include the standard stochastic heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u\dot{W}(t,x). \tag{4.3}$$

On an intuitive level, the noise acts as a huge and very irregular force in the heat equation, making even very familiar properties of the solutions of the heat equation somewhat non-obvious. For example, since $\dot{W}$ can be "huge and positive" – may that bring about growth at infinity that would knock the solution out of the $L^p(\mathbb{R})$ space? On the other hand, as the noise can be very negative, a priori it is by no means obvious that the strong maximum principle would hold: given that $u_0(x) \geq 0$, do we know that $u(t,x) > 0$? As the reader will see, getting the answers even to these questions will require some non-trivial arguments. We should also stress that the restriction to one spatial dimension is not technical or accidental, as the solutions have a very different nature in dimension $d > 1$.

The mild solutions: existence and uniqueness

Let us first make precise which notion of a solution we will use. The Duhamel formula tells us that if the noise were smooth, the solution of (4.1) would have the form

$$u(t,x) = \int_{\mathbb{R}} G(t,x-y)u_0(y)dy + \int_0^t \int_{\mathbb{R}} G(t-s,x-y)f(u(s,y))dW(s,y), \tag{4.4}$$

and this will be our starting point. That is, a solution of (4.1) is a solution of (4.4) that is adapted to the σ-algebra $\mathcal{F}_t$ generated by the white noise $\dot{W}$. Here, $G(t,x)$ is the standard heat kernel:

$$G(t,x) = \frac{1}{(4\pi t)^{1/2}} e^{-|x|^2/(4t)}. \tag{4.5}$$

These are also known as mild solutions.

Theorem 4.1 *The stochastic heat equation (4.1) with a globally Lipschitz nonlinearity $f(u)$ and a compactly supported initial condition $u_0(x)$ has a unique solution $u(t,x)$ such that for all $T>0$ we have*

$$\sup_{x\in\mathbb{R}}\sup_{0\le t\le T}\mathbb{E}(|u(t,x)|^2)<+\infty. \tag{4.6}$$

In other words, solutions exist and are unique in the space $P_{2,\infty}[0,T]$ with the norm

$$\|u\|^2_{P_{2,\infty}}=\sup_{x\in\mathbb{R}}\sup_{0\le t\le T}\mathbb{E}(|u(t,x)|^2). \tag{4.7}$$

Note that the stochastic integral in the right side of (4.4) makes sense for all functions in $P_{2,\infty}[0,T]$.

The proof of uniqueness

We first prove uniqueness. Suppose that u and v are two mild solutions of (4.1) – or, equivalently, of (4.4) in $P_{2,\infty}[0,T]$. We will show that v is a modification of u. Set

$$z(t,x)=u(t,x)-v(t,x),$$

and write

$$z(t,x)=\int_0^t\int_{\mathbb{R}}G(t-s,x,y)[f(u(s,y))-f(v(s,y))]W(dyds). \tag{4.8}$$

The Ito isometry implies that

$$\begin{aligned}\mathbb{E}(|z(t,x)|^2)&=\int_0^t\int_{\mathbb{R}}G^2(t-s,x,y)\mathbb{E}\Big[|f(u(s,y))-f(v(s,y))|^2\Big]dxds \\ &\le K\int_0^t\int_{\mathbb{R}}G^2(t-s,x,y)\mathbb{E}\Big[|u(s,y)-v(s,y)|^2\Big]dxds \\ &= K\int_0^t\int_{\mathbb{R}}G^2(t-s,x,y)\mathbb{E}\Big[|z(s,y)|^2\Big]dxds.\end{aligned} \tag{4.9}$$

We set

$$H(t)=\sup_{0\le s\le t}\sup_{x\in\mathbb{R}}\mathbb{E}(|z(s,x)|^2),$$

and get

$$H(t)\le K\int_0^t\int_{\mathbb{R}}G^2(t-s,x,y)H(s)dxds. \tag{4.10}$$

Note that

$$\int_{\mathbb{R}}G^2(s,y)dy=\frac{C}{s}\int_{\mathbb{R}}e^{-|y|^2/(4s)}ds=\frac{C'}{\sqrt{s}}.$$

It follows from (4.10) that

$$H(t)\le K'\int_0^t\frac{H(s)}{|t-s|^{1/2}}ds. \tag{4.11}$$

Hölder's inequality with any $p \in (1,2)$ and

$$\frac{1}{p} + \frac{1}{q} = 1, \tag{4.12}$$

implies that

$$H(t)^q \le K'' \int_0^t H(s)^q ds. \tag{4.13}$$

Grownwall's inequality implies now that $H(t) = 0$ for almost all s, hence u and v are modifications of each other.

The existence proof

The proof is via the usual Picard iteration scheme. We let $u_0(t,x) = u_0(x)$ and define iteratively

$$u_{n+1}(t,x) = \int_{\mathbb{R}} G(t, x-y) u_0(y) dy + \int_0^t \int_{\mathbb{R}} G(t-s, x-y) f(u_n(s,y)) dW(s,y). \tag{4.14}$$

It is easy to verify that all u_n are in $P_{2,\infty}[0,T]$. The increment

$$q_n(t,x) = u_{n+1}(t,x) - u_n(t,x)$$

satisfies

$$q_n(t,x) = \int_0^t \int_{\mathbb{R}} G(t-s, x-y)(f(u_n(s,y)) - f(u_{n-1}(s,y)) dW(s,y). \tag{4.15}$$

As f is Lipschitz, it follows that

$$\begin{aligned} \mathbb{E}(|q_n(t,x)|^2) &= \int_0^t \int_{\mathbb{R}} G^2(t-s, x-y) \mathbb{E}(f(u_n(s,y)) - f(u_{n-1}(s,y))^2 dy ds \\ &\le K^2 \int_0^t \int_{\mathbb{R}} G^2(t-s, x-y) \mathbb{E}|q_{n-1}(s,y)|^2 dy ds. \end{aligned} \tag{4.16}$$

Hence, the function

$$Z_n(t) = \sup_{x \in \mathbb{R}} \sup_{0 \le s \le t} \mathbb{E}|q_n(s,x)|^2$$

satisfies

$$Z_n(t) \le K^2 \int_0^t \int_{\mathbb{R}} G^2(t-s, x-y) Z_{n-1}(s) dy ds, \tag{4.17}$$

and thus

$$Z_n(t) \le C \int_0^t \frac{Z_{n-1}(s)}{|t-s|^{1/2}} ds. \tag{4.18}$$

Once again, with $p \in (1,2)$ and q as in (4.12) we get

$$Z_n(t)^q \leq C \int_0^t Z_{n-1}(s)^q ds. \tag{4.19}$$

Hence, Gronwall's lemma implies that

$$Z_n(t)^q \leq C_1 \frac{(Ct)^{n-1}}{(n-1)!}.$$

As a consequence, we get

$$\sum_{n=0}^{\infty} Z_n^{1/2}(t) < +\infty.$$

It follows that the sequence $u_n(t,x)$ converges in $P_{2,\infty}[0,T]$ to a limit $u(t,x)$. The same argument based on the Ito isometry and the global Lipschitz bound on the function f implies that

$$\int_0^t \int_{\mathbb{R}} G(t-s,x-y) f(u_n(s,y)) dW(s,y) \to \int_0^t \int_{\mathbb{R}} G(t-s,x-y) f(u(s,y)) dW(s,y),$$

also in $P_{2,\infty}[0,T]$. We conclude that $u(t,x)$ is a solution to

$$u(t,x) = \int_{\mathbb{R}} G(t,x-y) u_0(y) dy + \int_0^t \int_{\mathbb{R}} G(t-s,x-y) f(u(s,y)) dW(s,y), \tag{4.20}$$

finishing the existence proof. □

Higher moments of the solutions

One may ask if the solutions of the stochastic heat equation (3.16) that we have constructed lie in better spaces, such as $P_{s,\infty}$ with the norm

$$\|u\|_{P_{s,\infty}}^s = \sup_{x \in \mathbb{R},\, 0 \leq t \leq T} \sup \mathbb{E}(|u(t,x)|^s), \tag{4.21}$$

and $s > 2$. We will not prove existence and uniqueness of the solution in $P_{s,\infty}$ with $s > 2$ but rather estimate its norm in this space. The solution can be constructed as a combination of the Picard iteration and very similar arguments. We will need Burkholder's inequality.

Theorem 4.2 *[Burkholder's inequality] Let N_t be a continuous martingale such that $N_0 = 0$, then for each $p \geq 2$ we have*

$$\mathbb{E}|N_t|^p \leq c_p \mathbb{E}(\langle N, N \rangle_t)^{p/2}, \tag{4.22}$$

with a constant $c_p > 0$ that depends only on p.

As a consequence of Burkholder's inequality, for any predictable function f we have

$$\mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} f(s,x) dW(s,x)\Big]^p \leq c_p \mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} |f(s,x)|^2 ds dx\Big]^{p/2}. \tag{4.23}$$

Thus, the moments of the solution of the stochastic heat equation

$$u(t,x)=\int_{\mathbb{R}}G(t,x-y)u_0(y)dy+\int_0^t\int_{\mathbb{R}}G(t-s,x-y)f(u(s,y))dW(s,y) \tag{4.24}$$

can be estimated as

$$\begin{aligned}M_s(t,x):=\mathbb{E}|u(t,x)|^s&\le C_0+C\mathbb{E}\Big[\int_0^t\int_{\mathbb{R}}G^2(t-s,x-y)f^2(u(s,y))dsdy\Big]^{s/2} \qquad (4.25)\\&\le C_0+C\mathbb{E}\Big[\int_0^t\int_{\mathbb{R}}G^2(t-s,x-y)u^2(s,y)dsdy\Big]^{s/2}.\end{aligned}$$

Here we have used the Lipschitz property of f. Let us assume for simplicity that $s=4$, then we can write

$$\begin{aligned}M_4(t,x)&\le C_0+C\mathbb{E}\Big[\int_0^t\int_{\mathbb{R}}G^2(t-s,x-y)u^2(s,y)dsdy\Big]^2 \qquad (4.26)\\&=C_0+C\int_0^t\int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}}G^2(t-s,x-y)G^2(t-s',x-y')\mathbb{E}[u^2(s,y)u^2(s',y')]dsdyds'dy'.\end{aligned}$$

Set

$$\bar{M}_4(t)=\sup_{x\in\mathbb{R}}M_4(t,x),$$

then we have, using the Cauchy inequality

$$\begin{aligned}\bar{M}_4(t)&\le C_0+C\int_0^t\int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}}G^2(t-s,y)G^2(t-s',y')\bar{M}_4^{1/2}(s)\bar{M}_4^{1/2}(s')dsdyds'dy'\\&\le C_0+C\Big(\int_0^t\frac{\bar{M}_4(s)^{1/2}ds}{|t-s|^{1/2}}\Big)^2\le C_0+C\Big(\int_0^t\bar{M}_4^{q/2}(s)ds\Big)^{2/q}, \qquad (4.27)\end{aligned}$$

with any $q>2$. Gronwall's lemma implies now that

$$\sup_{0\le t\le T}\bar{M}_4(t)\le C_T. \tag{4.28}$$

This argument can be generalized to all even integers s and from there to all $s<+\infty$. It is straightforward to adapt it to show the existence of the solutions in $P_{s,\infty}[0,T]$ via Picard's iteration. Uniqueness of the solutions in $P_{s,\infty}[0,T]$ follows immediately from the uniqueness result in $P_{2,\infty}[0,T]$ that we have already proved.

Exercise 4.3 *With a little more careful analysis one may show the following bound: there exists a constant $C>0$ that depends only the Lipschitz consant of f and $\|u_0\|_{L^\infty}$ so that*

$$\sup_{x\in\mathbb{R}}\mathbb{E}(|u(t,x)|^k)\le C^ke^{Ck^3t}. \tag{4.29}$$

The spatial L^2-bound

Let us now consider the unique $P_{2,\infty}[0,T]$-solution to

$$u(t,x) = \int_{\mathbb{R}} G(t,x-y)u_0(y)dy + \int_0^t \int_{\mathbb{R}} G(t-s,x-y)f(u(s,y))dW(s,y), \tag{4.30}$$

and ask if we may expect the L^2-norm in space of $u(t,x)$ to remain bounded – recall that we assumed that $u_0(x)$ is compactly supported (though this assumption can be easily weakened to $u_0 \in L^1(\mathbb{R})$ in the existence and uniqueness proofs). Clearly, this is not true just under the assumption that $f(u)$ is Lipschitz: if we take $f \equiv 1$ and consider the solutions of

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \dot{W}(t,x),$$

then there is no reason to expect that the solution has any spatial decay whatsoever. Let us, therefore, assume that, in addition to being globally Lipschitz, $f(u)$ satisfies $f(0) = 0$, so that

$$|f(u)| \leq K|u|. \tag{4.31}$$

The first integral in the right side of (4.30) is obviously in any $L^p(\mathbb{R})$, $1 \leq p \leq +\infty$, hence we only look at

$$U(t,x) = \int_0^t \int_{\mathbb{R}} G(t-s,x-y)f(u(s,y))dW(s,y),$$

and compute

$$\begin{aligned}
&\mathbb{E}\int_{\mathbb{R}} u^2(t,x)dx \leq 2\int |u_0(x)|^2dx + 2\int_{\mathbb{R}}\int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)\mathbb{E}[f^2(u(s,y))]dydsdx \\
&\leq 2\int |u_0(x)|^2dx + 2K\int_{\mathbb{R}}\int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)\mathbb{E}[|u(s,y)|^2]dydsdx \\
&\leq 2\|u_0\|_2^2 + C\int_0^t \int_{\mathbb{R}} \mathbb{E}[|u(s,y)|^2]dy\frac{ds}{|t-s|^{1/2}}.
\end{aligned} \tag{4.32}$$

Thus,

$$Z(t) = \mathbb{E}\int_{\mathbb{R}} |u(t,x)|^2dx$$

satisfies

$$Z(t) \leq 2Z(0) + C\int_0^t \frac{Z(s)ds}{|t-s|^{1/2}}.$$

Hence, for any $q \in (2,+\infty)$ and $0 \leq t \leq T$, we have

$$Z^q(t) \leq C_0 + C_T\int_0^t Z^q(s)ds.$$

Gronwall's lemma implies that

$$\sup_{0\leq t\leq T} Z(t) \leq C_T',$$

and we conclude that for every $T > 0$ we have

$$\sup_{0\le t\le T} \mathbb{E}\int_{\mathbb{R}} |u(t,x)|^2 dx < +\infty. \tag{4.33}$$

In the special case of the stochastic heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u\dot{W}, \tag{4.34}$$

with the initial condition $u(0,x) = u_0(x)$, we have

$$u(t,x) = v(t,x) + \int_0^t \int_{\mathbb{R}} G(t-s, x-y)u(s,y)W(dsdy). \tag{4.35}$$

It follows that

$$Z(t,x) = \mathbb{E}|u(t,x)|^2$$

satisfies a closed equation

$$Z(t,x) = v^2(t,x) + \int_0^t \int_{\mathbb{R}} G^2(t-s, x-y)Z(s,y)dy. \tag{4.36}$$

Here, the function $v(t,x)$ is the solution of the heat equation

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2}, \tag{4.37}$$

with the initial condition $v(0,x) = u_0(x)$. Hence, the L^2-norm of $Z(t,x)$,

$$\bar{Z}(t) = \mathbb{E}\int_{\mathbb{R}} |u(t,x)|^2 dx$$

satisfies

$$\bar{Z}(t) = \|v(t)\|_{L^2}^2 + b\int_0^t \frac{\bar{Z}(s)}{\sqrt{t-s}} ds, \tag{4.38}$$

with an explicit constant $b > 0$. Let us define the function

$$Z_\gamma(t) = e^{-\gamma t}\bar{Z}(t),$$

with the constant $\gamma > 0$ to be chosen. Then $Z_\gamma(t)$ satisfies

$$Z_\gamma(t) = a(t) + \int_0^t g(t-s)Z_\gamma(s)ds, \tag{4.39}$$

with

$$a(t) = \|v(t)\|_{L^2}^2 e^{-\gamma t}, \quad g(t) = \frac{be^{-\gamma t}}{\sqrt{t-s}}.$$

Let us choose

$$\gamma = \pi b^2, \tag{4.40}$$

so that

$$\int_0^\infty g(s)ds = 1. \tag{4.41}$$

This is, clearly, a necessary condition for $Z_\gamma(t)$ to have a limit as $t \to +\infty$, since $a(t) \to 0$ as $t \to +\infty$.

Exercise 4.4 Show that with this choice of γ and this $a(t)$, the limit

$$\bar{Z}_\gamma = \lim_{t\to+\infty} Z_\gamma(t) \tag{4.42}$$

exists.

In order to find the limit, let us write

$$Z_\gamma(t) = \bar{Z}_\gamma + \beta(t),$$

with $\beta(t) \to 0$ as $t \to +\infty$:

$$\bar{Z}_\gamma + \beta(t) = a(t) + \bar{Z}_\gamma \int_0^t g(t-s)ds + \int_0^t g(t-s)\beta(s)dy, \tag{4.43}$$

so that

$$\beta(t) = a(t) - \bar{Z}_\gamma \int_t^\infty g(s)ds + \int_0^t g(t-s)\beta(s)dy. \tag{4.44}$$

Integrating (4.44) gives

$$\int_0^t \beta(s)ds = \int_0^t a(s)ds - \bar{Z}_\gamma \int_0^t \int_s^\infty g(s')ds'ds + \int_0^t \int_0^s g(s-s')\beta(s')ds'ds. \tag{4.45}$$

The long time limit of the second integral in the right side can be computed as

$$\int_0^t \int_s^\infty g(s')ds'ds = \int_0^t s'g(s')ds' + t\int_t^\infty g(s')ds' \to \int_0^\infty sg(s)ds, \text{ as } t \to +\infty, \tag{4.46}$$

while for the last integral in the right side of (4.45) we have

$$\begin{aligned} \int_0^t \int_0^s g(s-s')\beta(s')ds'ds &\to \int_0^\infty \int_0^s g(s-s')\beta(s')ds'ds \\ &= \int_0^\infty \beta(s') \int_{s'}^\infty g(s-s')dsds' = \int_0^\infty \beta(s)ds, \text{ as } t \to +\infty. \end{aligned} \tag{4.47}$$

Going back to (4.45) we conclude that

$$\bar{Z}_\gamma = \Big(\int_0^\infty sg(s)ds\Big)^{-1} \int_0^\infty a(s)ds. \tag{4.48}$$

Therefore, the solution of the stochastic heat equation (4.34) satisfies

$$\mathbb{E}\|u(t)\|_{L^2}^2 \sim \bar{Z}_\gamma e^{\gamma t}, \text{ as } t \to +\infty, \tag{4.49}$$

with $\gamma > 0$ given by (4.40).

Exercise 4.5 Generalize this argument to the solutions of equations of the form

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{4.50}$$

with a nonlinearity $f(u)$ such that $c_1|u| \le f(u) \le c_2|u|$. Obtain a lower and upper bound for the L^2-norm of the solutions as in (4.49).

The exponential growth of the second moment in (4.49) should be contrasted with the simple bound on the integral:

$$\mathbb{E}\int_{\mathbb{R}} u(t,x)dx = \int_{\mathbb{R}} u_0(x)dx. \tag{4.51}$$

We will discuss this again when we talk about the intermittency of the solutions. Roughly, the disparity of the L^1 and L^2 norms of the solutions indicates that there are small islands where the solution is huge. We should also note that we will later show that $u(t,x) > 0$ if $u_0(x) \geq 0$ and u_0 does not vanish identically. Hence, the integral in the left side of (4.51) is the L^1-norm of u.

The Hölder regularity of the solutions

In order to study the Hölder regularity of the solutions, let us first make a slightly simplifying assumption that in addition to being Lipschitz, the function $f(u)$ is globally bounded:

$$\sup_{u\in\mathbb{R}} |f(u)| \leq K. \tag{4.52}$$

We will later explain how this assumption can be removed, using the $P_{s,\infty}[0,T]$ bounds on the solution with $s \in (2,+\infty)$, rather than just the bounds in $P_{2,\infty}[0,T]$ that we will use in the proof.

Theorem 4.6 *There exists a modification of the solution of (4.4) that is Hölder continuous in x of any order less than $1/2$ and in t of any order less than $1/4$.*

We will need in the proof a slight generalization of the Kolmogorov continuity criterion – compare this to Theorem 2.2.

Theorem 4.7 *Let X_t, $t \in \mathbf{T} = [a_1,b_1] \times \ldots, [a_d,b_d] \subset \mathbb{R}^d$ be a real-valued stochastic process. Suppose there are constants $k > 1$, $C > 0$ and $\alpha_i > 0$, $i = 1,\ldots,d$, so that*

$$q := \sum_{i=1}^{d} \frac{1}{\alpha_i} < 1,$$

and for all $s,t \in \mathbf{T}$, we have

$$\mathbb{E}(|X(t) - X(s)|^k) \leq C\sum_{i=1}^{d} |t_i - s_i|^{\alpha_i}. \tag{4.53}$$

Then $X(t)$ has a continuous modification $\bar{X}(t)$. Moreover, $\bar{X}(t)$ is Hölder continuous in each variable t_i with any exponent $\gamma \in (0, \alpha_i(1-q)/k)$.

Note that when all $\alpha_i = \alpha$, then the assumptions require $\alpha > d$, and the process has the Hölder exponent less than

$$\alpha(1-\frac{d}{\alpha})\frac{1}{k} = \frac{\alpha-d}{k},$$

which is exactly Theorem 2.2. We will leave the proof as an exercise.

The proof of Theorem 4.6

Let us consider

$$U(t,x)=\int_0^t\int_{\mathbb{R}} G(t-s,x-y)f(u(s,y))dW(s,y). \tag{4.54}$$

We need to show that $U(t,x)$ has the required Hölder continuous modification. For $0\le t\le t'$ we write

$$\begin{aligned}U(t',x)-U(t,x)&=\int_0^t\int_{\mathbb{R}}[G(t'-s,x-y)-G(t-s,x-y)]f(u(s,y))dW(s,y)\\&+\int_t^{t'}\int_{\mathbb{R}} G(t'-s,x-y)f(u(s,y))dW(s,y).\end{aligned} \tag{4.55}$$

Young's and Burkholder's inequalities imply that

$$\begin{aligned}\mathbb{E}|U(t',x)-U(t,x)|^p&\le C_p\mathbb{E}\Big[\int_0^t\int_{\mathbb{R}}|G(t'-s,x-y)-G(t-s,x-y)|^2f^2(u(s,y))dyds\Big]^{p/2}\\&+C_p\mathbb{E}\Big[\int_t^{t'}\int_{\mathbb{R}}G^2(t'-s,x-y)f^2(u(s,y))dyds\Big]^{p/2}=I+II.\end{aligned} \tag{4.56}$$

Using the assumption that $|f(u)|\le K$, the second term can be estimated as

$$II\le C\Big[\int_t^{t'}\int_{\mathbb{R}}G^2(t'-s,x-y)dyds\Big]^{p/2}\le C\Big[\int_t^{t'}\frac{ds}{|t'-s|^{1/2}}\Big]^{p/2}\le C|t'-t|^{p/4}. \tag{4.57}$$

For the first term in the right side of (4.56) we write, using the Plancherel identity

$$\begin{aligned}\int_{\mathbb{R}}|G(t'-s,x-y)-G(t-s,x-y)|^2dy&=\int_{\mathbb{R}}|G(t'-s,y)-G(t-s,y)|^2dy \quad (4.58)\\ =C\int_{\mathbb{R}}\left|e^{-(t'-s)|\xi|^2}-e^{-(t-s)|\xi|^2}\right|^2d\xi&=C\int_{\mathbb{R}}e^{-2(t-s)|\xi|^2}\Big[1-e^{-(t'-t)|\xi|^2}\Big]^2d\xi.\end{aligned}$$

It follows that the first integral in there right side of (4.56) can be bounded as

$$I^{2/p}\le C\int_0^t\int_{\mathbb{R}}e^{-2(t-s)|\xi|^2}\Big[1-e^{-(t'-t)|\xi|^2}\Big]^2d\xi ds=C\int_{\mathbb{R}}\frac{1}{|\xi|^2}\Big(1-e^{-2t|\xi|^2}\Big)\Big[1-e^{-(t'-t)|\xi|^2}\Big]^2d\xi. \tag{4.59}$$

Now, we use the following two elementary estimates: first, there exists $C_T>0$ so that for all $0\le t\le T$ and all $\xi\in\mathbb{R}$ we have

$$\frac{1}{|\xi|^2}\Big(1-e^{-2t|\xi|^2}\Big)\le\frac{C_T}{1+|\xi|^2},$$

and, second,

$$1-e^{-(t'-t)|\xi|^2}\le 2\min[(t'-t)|\xi|^2,1].$$

Using these estimates in (4.59) gives

$$\begin{aligned}I^{2/p}&\le C\int_{\mathbb{R}}\frac{1}{1+|\xi|^2}\min[(t'-t)|\xi|^2,1]d\xi\\&=C_T\int_0^{|t'-t|^{-1/2}}\frac{(t'-t)|\xi|^2}{1+|\xi|^2}d\xi+C_T\int_{|t'-t|^{-1/2}}\frac{d\xi}{1+|\xi|^2}\le C_T|t'-t|^{1/2}.\end{aligned}$$

We conclude that

$$\mathbb{E}|U(t',x)-U(t,x)|^p \le C_T|t'-t|^{p/4}. \tag{4.60}$$

Exercise 4.8 Show that

$$\mathbb{E}|U(t,x)-U(t,x')|^p \le C_p\Big(\int_0^t\int_{\mathbb{R}}|G(t-s,y)-G(t-s,x-x'-y)|^2dyds\Big)^{p/2}, \tag{4.61}$$

and then use a similar computation to what we have done to show that

$$\mathbb{E}|U(t,x)-U(t,x')|^p \le C_T|x-x'|^{p/2}. \tag{4.62}$$

Summarizing, we have

$$\mathbb{E}|U(t',x')-U(t,x)|^p \le C_T\Big(|t'-t|^{p/4}+|x-x'|^{p/2}\Big). \tag{4.63}$$

Now, we use Theorem 4.7 with $k=p$, $\alpha_x=p/2$ and $\alpha_t=p/4$, so that

$$q=\frac{1}{\alpha_x}+\frac{1}{\alpha_t}=\frac{2}{p}+\frac{4}{p}=\frac{6}{p},$$

so we get Hölder continuity in t with any exponent smaller than

$$\bar{\gamma}_t=\frac{\alpha_t(1-q)}{p}=\frac{1}{4}\Big(1-\frac{6}{p}\Big),$$

and in x with any exponent smaller than

$$\bar{\gamma}_t=\frac{\alpha_x(1-q)}{p}=\frac{1}{2}\Big(1-\frac{6}{p}\Big).$$

As $p>2$ is arbitrary, it follows that $u(t,x)$ is Hölder continuous in x with any exponent smaller than $1/2$ and in t with any exponent smaller than $1/4$.

Exercise 4.9 Use the bounds on the higher moments $\mathbb{E}|u(t,x)|^p$ to improve the argument above to show that the almost sure Hölder regularity of $u(t,x)$ with the same exponents holds under the (weaker) assumption that the function $f(u)$ is Lipschitz rather than bounded, removing assumption (4.52).

The comparison principle

It is well known if a function $f(u)$ is Lipschitz and $f(0)=0$, then the parabolic equations of the form

$$\frac{\partial u}{\partial t}=\Delta u+f(u), \tag{4.64}$$

satisfy the comparison principle. That is, if $u(t,x)$ and $v(t,x)$ are two solutions of (4.64) and $u(0,x)\le v(0,x)$ for all $x\in\mathbb{R}$ then $u(t,x)\le v(t,x)$ for all $t\ge 0$ and $x\in\mathbb{R}$. Moreover, the strong comparison principle says that actually $u(t,x)<v(t,x)$ for all $t>0$ and $x\in\mathbb{R}$ provided that $u(0,x)\not\equiv v(0,x)$. These results are easily generalized to equations of the form

$$\frac{\partial u}{\partial t}=\Delta u+g(t,x)f(u), \tag{4.65}$$

with a regular function $g(t,x)$.

Here, we prove the following comparison theorem for the solutions of the stochastic heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{4.66}$$

with a Lipschitz nonlinearity $f(u)$. The difference with the classical PDE results is that the noise $\dot{W}$ is highly irregular. The "canonical" PDE proof with a bounded function $g(t,x)$ relies on the fact that the Hessian of a function at a minimum is non-negative definite matrix. Here, we can not use this strategy since the solutions are merely Hölder with exponent less than $1/2$ in space.

Theorem 4.10 *Let $u(t,x)$ and $v(t,x)$ be two solutions of (4.66) such that $u(0,x) \geq v(0,x)$. Then, almost surely, we have $u(t,x) \geq v(t,x)$ for all $t \geq 0$ and $x \in \mathbb{R}$.*

Note that we do not yet claim the strong comparison principle, which says that $u(t,x) > v(t,x)$ for all $t > 0$ and $x \in \mathbb{R}$ unless $u_0(x) \equiv v_0(x)$. This will be done slightly later.

The idea of the proof of Theorem 4.10 is to use numerical analysis. We construct approximate solutions $u_n(t,x)$ and $v_n(t,x)$ such that almost surely we have $u_n(t,x) \geq v_n(t,x)$ for all $t \geq 0$ and $x \in \mathbb{R}$ and then pass to the limit $n \to +\infty$. The approximation is done by time-splitting in time and discretizing space. This is also an alternative way to construct the solutions of the original SPDE.

The time splitting schemes

Solution of a linear equation of the form

$$\frac{du}{dt} = (A+B)u, \tag{4.67}$$

is given by

$$u(t) = e^{(A+B)t}u_0. \tag{4.68}$$

If the linear operators A and B commute then we have

$$u(t) = e^{At}v(t), \quad v(t) = e^{Bt}u_0. \tag{4.69}$$

This means that we can solve first

$$\frac{dv}{dt} = Bv, \quad v(0) = u_0, \quad 0 \leq t \leq T,$$

followed by

$$\frac{du}{dt} = Au, \quad u(0) = v(T), \quad 0 \leq t \leq T,$$

and obtain the correct $u(T)$. When the operators A and B do not commute, one relies on the Trotter formula

$$e^{(A+B)t} = \lim_{n\to+\infty}(e^{A/n}e^{B/n})^n. \tag{4.70}$$

The corresponding time-splitting scheme proceeds as follows. We divide the time axis $t > 0$ into intervals of the form

$$T_{nj} = \Big\{ \frac{j}{n^2} \le t < \frac{j+1}{n^2} \Big\}. \tag{4.71}$$

On each time interval T_{nj} we first solve

$$\frac{\partial v}{\partial t} = Bv, \quad v(\frac{j}{n^2}) = u(\frac{j}{n^2}), \quad \frac{j}{n^2} \le t \le \frac{j+1}{n^2}, \tag{4.72}$$

followed by

$$\frac{\partial u}{\partial t} = Av, \quad u(\frac{j}{n^2}) = v(\frac{j+1}{n^2}), \quad \frac{j}{n^2} \le t \le \frac{j+1}{n^2}, \tag{4.73}$$

which gives us $u((j+1)/n^2)$, and we can solve (4.72) on the time interval $T_{n,j+1}$, and so on. Convergence of $u(t)$ to the solution of (4.67) is guaranteed by the Trotter formula under certain assumptions on A and B.

The spatial discretization and time splitting for the stochastic heat equation

For the stochastic heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{4.74}$$

we would like to consider the following time-splitting scheme: the first step is solving a point-wise SDE

$$\frac{\partial u_{n,j+1/2}}{\partial t} = f(u_{n,j+1/2})\dot{W}, \quad \frac{j}{n^2} \le t < \frac{j+1}{n^2}, \tag{4.75}$$

with the initial condition

$$u_{n,j+1/2}(\frac{j}{n^2}, x) = u_{n,j}(\frac{j}{n^2}, x),$$

followed by the heat equation

$$\frac{\partial u_{n,j+1}}{\partial t} = \frac{\partial^2 u_{n,j+1}}{\partial x^2}, \quad \frac{j}{n^2} \le t < \frac{j+1}{n^2}, \tag{4.76}$$

with the initial condition

$$u_{n,j+1}(\frac{j}{n^2}, x) = u_{n,j+1/2}(\frac{j+1}{n^2}, x).$$

This would give us the initial condition $u_{n,j+1}((j+1)/n^2, x)$ for the next SDE step (4.75) on the time interval $T_{n,j+1}$, and we would be able to re-start.

One difficulty is making sense of (4.75) as it is neither an SDE nor an SPDE. Hence, in addition to the time-splitting, we will discretize in space. We will consider functions $u_{nj}(t,x)$ that are piecewise constant on the spatial intervals

$$I_{nk} = \Big\{ \frac{k}{n} - \frac{1}{2n} \le x < \frac{k}{n} + \frac{1}{2n} \Big\}.$$

Each u_{nj} is defined on the time interval T_{nj}. Given $u_{nj}(j/n^2,x)$, in order to define $u_{n,j+1/2}$ and $u_{n,j+1}$, we first solve a family of SDEs

$$u_{n,j+1/2}(t,\frac{k}{n}) = u_{nj}(\frac{j}{n^2},\frac{k}{n}) + n\int_{j/n^2}^{t}\int_{I_{nk}} f(u_{n,j+1/2}(s,\frac{k}{n}))W(dsdy). \tag{4.77}$$

In other words, on the time interval T_{nj}, the piece-wise constant in space function $u_{n,j+1/2}(t,x)$ satisfies the SDE

$$du_n(t,\frac{k}{n}) = f(u_n(t,\frac{k}{n}))dB_k, \tag{4.78}$$

where

$$B_k(t) = n\int_0^t\int_{I_{nk}} W(dsdy)$$

is the standard Brownian motion.

In order to incorporate the heat equation step, we consider the discrete Laplacian

$$\Delta_n u(\frac{k}{n}) = n^2\Big[u(\frac{k+1}{n}) + u(\frac{k-1}{n}) - 2u(\frac{k}{n})\Big].$$

The function $u_{n,j+1}(t,x)$, also defined on the time interval T_{nj} is the solution of

$$\frac{\partial u_{n,j+1}}{\partial t} = \Delta_n u_{n,j+1}, \quad t\in T_{nj}, \tag{4.79}$$

with the initial condition $u_{n,j+1}(j/n^2,x) = u_{n,j+1/2}((j+1)/n^2,x)$.

It is convenient to re-write the above scheme in terms of the Green's function $G_n(t,x,y)$ of the discrete Laplacian. It is defined for the lattice points of the form $x=k/n$, $y=m/n$, and is the solution of

$$\frac{\partial G_n}{\partial t} = \Delta_n G_n, \tag{4.80}$$

with the initial condition

$$G_n(0,\frac{k}{n},\frac{m}{n}) = \begin{cases} n, & \text{if } k=m,\\ 0, & \text{otherwise.}\end{cases}$$

We extend $G_n(t,x,y)$ to $x,y\in\mathbb{R}$ as

$$G_n(t,x,y) = G_n(t,\frac{k}{n},\frac{m}{n}), \qquad \text{if } \frac{k}{n}-\frac{1}{2n}\le x<\frac{k}{n}+\frac{1}{2n} \text{ and } \frac{m}{n}-\frac{1}{2n}\le y<\frac{m}{n}+\frac{1}{2n},$$

Let us now verify that the approximation $u_{nj}(t,x)$ that we have defined above via the time-splitting scheme is the solution of (dropping sub-script j)

$$u_n(t,x) = \int_{\mathbb{R}}\bar{G}_n(t,0,x,y)u_n(0,y)dy + \int_0^t\int_{\mathbb{R}}\bar{G}_n(t,s,x,y)f(u_n(s,y))W(dsdy), \tag{4.81}$$

with the initial condition

$$u_n(0,x) = n\int_{k-1/(2n)}^{k+1/(2n)} u(0,y)dy, \quad \text{for } \frac{k}{n}-\frac{1}{2n}\le x<\frac{k}{n}+\frac{2}{2n},$$

and $\bar{G}(t,s,x,y)$ defined as

$$\bar{G}_n(t,s,x,y) = G_n\Big(\frac{[n^2t]-[n^2s]}{n^2}, x, y\Big),$$

for all $t \geq s$.

Indeed, given $t \in T_{nj}$ we have $[n^2t] = j$, thus $\bar{G}_n(t,0,x,y) = G_n(j/n^2, x, y)$. In addition, for $0 \leq s \leq j/n^2$ we have $\bar{G}(t,s,x,y) = \bar{G}(j/n^2, s, x, y)$. Hence we may re-write (4.81) as

$$u_n(t,x) = u_n(\frac{j}{n^2}, x) + \int_{j/n^2}^t \int_{\mathbb{R}} \bar{G}_n(t,s,x,y) f(u_n(s,y)) W(dsdy). \tag{4.82}$$

Next, for $x \in I_{nk}$, $t \in T_{nj}$, and $j/n^2 \leq s \leq t$, we have

$$\bar{G}_n(t,s,x,y) = G_n(0, \frac{k}{n}, y) = n\mathbb{1}_{I_{nk}}(y).$$

Hence, (4.82) says

$$u_n(t,x) = u_n(\frac{j}{n^2}, x) + n\int_{j/n^2}^t \int_{I_{nk}} f(u_n(s, \frac{k}{n})) W(dsdy). \tag{4.83}$$

Therefore, on the time interval T_{nj}, the piece-wise constant (in space) function $u_n(t,x)$ satisfies the SDE (4.78). At the time $t = (j+1)/n^2$ the function $u_n(t,x)$ experiences a jump. To describe it, we go back to (4.81): the solution after the jump is given by

$$\begin{aligned} u_n(\frac{j+1}{n^2}+, x) = &\int_{\mathbb{R}} \bar{G}_n(\frac{j+1}{n^2}, 0, x, y) u_n(0,y) dy \\ &+ \int_0^{(j+1)/n^2} \int_{\mathbb{R}} \bar{G}_n(\frac{j+1}{n^2}, s, x, y) f(u_n(s,y)) W(dsdy), \end{aligned} \tag{4.84}$$

while just before the jump we have

$$\begin{aligned} u_n(\frac{j+1}{n^2}-, x) = &\int_{\mathbb{R}} \bar{G}_n(\frac{j}{n^2}, 0, x, y) u_n(0,y) dy \\ &+ \int_0^{(j+1)/n^2} \int_{\mathbb{R}} \bar{G}_n(\frac{j}{n^2}, s, x, y) f(u_n(s,y)) W(dsdy). \end{aligned} \tag{4.85}$$

The semi-group property for (4.80) implies that

$$G_n(t, \frac{k}{n}, \frac{p}{n}) = \frac{1}{n}\sum_m G(t-s, \frac{k}{n}, \frac{m}{n}) G(s, \frac{m}{n}, \frac{p}{n}), \tag{4.86}$$

for all $0 \leq s \leq t$, k and p. The continuous version of (4.86) is

$$G_n(t,x,y) = \int_{\mathbb{R}} G_n(t-s, x, z) G_n(s, z, y) dz. \tag{4.87}$$

It follows from (4.87) that for $s < (j+1)/n^2$ we have

$$\bar{G}_n(\frac{j+1}{n^2}, s, x, y) = G_n(\frac{j-[n^2 s]}{n^2} + \frac{1}{n^2}, x, y) = \int_{\mathbb{R}} G_n(\frac{1}{n^2}, x, z)\bar{G}_n(\frac{j}{n^2}, s, z, y)dz.$$

Using this in (4.84), together with the semi-group property gives

$$\begin{aligned} u_n(\frac{j+1}{n^2}+, x) &= \int_{\mathbb{R}}\int_{\mathbb{R}} G_n(\frac{1}{n^2}, x, z)\bar{G}_n(\frac{j}{n^2}, z, y)u_n(0, y)dydz \qquad (4.88)\\ &+ \int_0^{(j+1)/n^2}\int_{\mathbb{R}}\int_{\mathbb{R}} G_n(\frac{1}{n^2}, x, z)\bar{G}_n(\frac{j}{n^2}, s, z, y)f(u_n(s, y))W(dsdy)dz \\ &= \int_{\mathbb{R}} G_n(\frac{1}{n^2}, x, z)u_n(\frac{j+1}{n^2}-, z)dz. \end{aligned}$$

We see that, indeed, the passage from u_n at the time $t = (j+1)/n^2-$ to u_n at $t = (j+1)/n^2+$ is exactly via solving the discrete heat equation (4.79).

Convergence of the approximation

We will need the result of the following exercise, verified by a lengthy computation found in [4].

Exercise 4.11 Show that the following two bounds hold: first,

$$\int_0^t \int_{\mathbb{R}} [\bar{G}_n(t, s, x, y) - G(t-s, x-y)]^2 dsdy \le \frac{c}{n}, \qquad (4.89)$$

and, second,

$$\sup_{0\le s\le t}\sup_{x\in\mathbb{R}} \Big[\int_{\mathbb{R}} [\bar{G}_n(t, s, x, y) - G(t-s, x-y)]u(0, y)dy\Big]^2 \to 0 \text{ as } n \to +\infty. \qquad (4.90)$$

We now show that

$$M(t) := \sup_{0\le s\le t}\sup_{x\in\mathbb{R}} \mathbb{E}|u_n(s, x) - u(s, x)|^2 \to 0 \text{ as } n \to +\infty. \qquad (4.91)$$

To see this, let us recall that

$$u(t, x) = \int_{\mathbb{R}} G(t, x-y)u(0, y)dy + \int_0^t \int_{\mathbb{R}} G(t-s, x-y)f(u(s, y))W(dsy),$$

and

$$u_n(t, x) = \int_{\mathbb{R}} \bar{G}_n(t, 0, x, y)u_n(0, y)dy + \int_0^t \int_{\mathbb{R}} \bar{G}_n(t, s, x, y)f(u_n(s, y))W(dsdy),$$

Subtracting, we get

$$\bar{M}(s, x) = \mathbb{E}|u_n(s, x) - u(s, x)|^2 \le C(I + II), \qquad (4.92)$$

with

$$I = \Big| \int_{\mathbb{R}} G(t, x-y)u(0,y)dy - \int_{\mathbb{R}} \bar{G}_n(t,0,x,y)u_n(0,y)dy \Big|^2 \to 0, \quad \text{as } n \to +\infty,$$

because of (4.90) and the fact that $u_n(0,y)$ converges to $u(0,y)$ in every L^p-norm. The other term is

$$II = \int_0^t \int_{\mathbb{R}} \mathbb{E}|G(t-s,x-y)f(u(s,y)) - \bar{G}_n(t,s,x,y)f(u_n(s,y))|^2 dsdy,$$

and can be bounded as

$$\begin{aligned} II \le C \int_0^t \int_{\mathbb{R}} \mathbb{E}|(G(t-s,x-y) - \bar{G}_n(t,s,x,y))f(u_n(s,y))\big|^2 dsdy \\ + C \int_0^t \int_{\mathbb{R}} \mathbb{E}|G(t-s,x-y)(f(u(s,y)) - f(u_n(s,y)))\big|^2 dsdy = II_1 + II_2. \end{aligned} \tag{4.93}$$

It is straightforward to verify that there exists C_T so that

$$\sup_{0\le s\le T} \sup_{x\in\mathbb{R}} \mathbb{E}|u_n(s,x)|^2 \le C_T. \tag{4.94}$$

This, together with (4.89) and the Lipschitz bond on f means that

$$II_1 \le \frac{C_T}{n}. \tag{4.95}$$

The last term in right side of (4.93) is bounded as

$$II_2 \le C \int_0^t \frac{M(s)ds}{\sqrt{t-s}}. \tag{4.96}$$

Therefore, we have an estimate for $M(s)$:

$$M(t) \le \alpha(n) + \int_0^t \frac{M(s)ds}{\sqrt{t-s}}, \tag{4.97}$$

with $\alpha(n) \to 0$ as $n \to +\infty$. We conclude that

$$M(t) \to 0 \text{ as } n \to +\infty. \tag{4.98}$$

Back to the comparison principle

We have shown that $u(t,x)$ and $v(t,x)$ can be obtained via the time-splitting approximation scheme. The approximations $u_n(t,x)$ and $v_n(t,x)$ satisfy $u_n(t,x) \ge v_n(t,x)$ if $u(0,x) \ge v(0,x)$ for all $x \in \mathbb{R}$. This is because each of the steps in the time splitting scheme preserves the order. Indeed, the heat equation has the comparison principle, while an SDE

$$du = f(u)dB_t$$

also preserves the order because if u and v are two solutions, the difference $z = u - v$ satisfies

$$dz = g(t)zdB_t,$$

with the function

$$g(t) = \frac{f(u(t)) - f(v(t))}{u(t) - v(t)}.$$

It is easy to verify that $z(t) \geq 0$ for all $t > 0$ if $z(0) \geq 0$. This completes the proof of Theorem 4.10.

The strong maximum principle

The heat equation

$$\frac{\partial u}{\partial t} = \Delta u, \tag{4.99}$$

in addition to the comparison principle, has the strong maximum principle: if $u_0(x) \geq 0$ and $u_0(x) \not\equiv 0$ everywhere, then $u(t,x) > 0$ for all $t > 0$ and all $x \in \mathbb{R}^d$. In other words, solutions with compactly supported nonnegative initial data become positive everywhere instantaneously. On the other hand, the heat equation with a non-Lipschitz nonlinearity

$$\frac{\partial u}{\partial t} = \Delta u - \sqrt{u}, \tag{4.100}$$

does not satisfy the strong maximum principle: solutions have compact support at $t > 0$ if $u_0(x)$ is compactly supported. One can think of (4.100) as

$$\frac{\partial u}{\partial t} = \Delta u - g(t,x)u, \tag{4.101}$$

with $g(t,x) = 1/\sqrt{u}$ that is large when u is small. Thus, a "large" $g(t,x)$ can prevent $u(t,x)$ from having non-compact support, and, of course, white noise is a pretty large force. Nevertheless, solutions of the stochastic heat equation have non-compact support. We formulate the result for the linear equation but it holds for any Lipschitz nonlinearity $f(u)$ such that $f(0) = 0$.

Theorem 4.12 *Let $u(t,x)$ be the solution of*

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u\dot{W}, \tag{4.102}$$

with the initial condition $u_0(x) \geq 0$ for all $x \in \mathbb{R}$. If $u_0(x) \not\equiv 0$, and $u_0(x)$ is continuous, then, almost surely, for each $t > 0$, we have $u(t,x) > 0$ for all $x \in \mathbb{R}$.

Taking $f(u) = u$ in (4.102) is not necessary, and is made purely to simplify some steps in the proof. On the other hand, the Lipschitz assumption on f is crucial: the conclusion is false if $f(u) = \sqrt{u}$.

Let us assume without loss of generality that $t = 1$, and take some $R > 2$. Because of an application of the large deviations principle in the proof, it will be convenient to restrict the problem to a finite interval. Consider the solution of the stochastic heat equation (4.101) with the Dirichlet boundary conditions at $x = \pm 2R$:

$$\begin{aligned} &\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + u\dot{W}, \quad t > 0, \quad |x| < 2R, \\ &v(t, -2R) = v(t, 2R) = 0, \end{aligned} \tag{4.103}$$

and the initial condition $v(0, x) = u_0(x)$. The solution is, once again, understood in the mild sense:

$$v(t, x) = \int_{|y| \le 2R} G_R(t, x, y) u_0(y) dy + \int_0^t \int_{|y| \le 2R} G_R(t - s, x, y) f(u(s, y)) W(dyds). \tag{4.104}$$

Here, $G_R(t, x, y)$ is Green's function for the Dirichlet problem:

$$\begin{aligned} &\frac{\partial G_R}{\partial t} = \frac{\partial^2 G_R(t, x, y)}{\partial x^2}, \quad t > 0, \quad |x| < 2R, \\ &G_R(t, -2R, y) = G_R(t, 2R, y) = 0, \\ &G_R(0, x, y) = \delta(x - y). \end{aligned} \tag{4.105}$$

Exercise 4.13 Use the time-splitting argument used in the proof of the comparison principle to show that $u(t, x) \ge v(t, x)$ for all $t \ge 0$ and $|x| \le 2R$.

As R is arbitrary, it suffices to show that with probability one

$$v(t = 1, x) > 0 \text{ for all } |x| \le R. \tag{4.106}$$

We may assume without loss of generality that

$$u_0(x) \ge \delta_0 \mathbb{1}_{[-1,1]}(x),$$

for some $\delta_0 > 0$. We will proceed "step-by-step". Fix $N > 0$ and set $t_k = k/N$. Let $\mathcal{A}_k$ be the event that there exists some $\delta_k > 0$ so that

$$v(t_k, x) \ge \delta_k I_k(x), \quad \text{for all } x \in \mathbb{R},$$

where

$$I_k(x) = \mathbb{1}\Big(-1 - \frac{Rk}{N} \le x \le 1 + \frac{Rk}{N} \Big).$$

As $v(t_k, x)$ is almost surely Hölder continuous in x, the event $\mathcal{A}_k$ is simply that $v(t_k, x)$ is strictly positive on I_k. The support of I_k grows with k, or "in time", and at the last moment we have

$$I_N(x) > \mathbb{1}_{[-R,R]}(x).$$

We will show that for all $\varepsilon > 0$ we may choose N_ε so large that for all $k = 1, \ldots, N_\varepsilon$ we have

$$\mathbb{P}\Big(\mathcal{A}_{k+1}^c \Big| \mathcal{A}_1 \cap \ldots \cap \mathcal{A}_k\Big) < \frac{\varepsilon}{N_\varepsilon}. \tag{4.107}$$

This estimate shows that support of v has to grow with a large probability – in the end we will show that the support of v is all of $[-2R, 2R]$ but we are not there yet. With (4.107) in hand, we would have

$$\mathbb{P}(\mathcal{A}^c_{N_\varepsilon}) \le \sum_{k=0}^{N_\varepsilon-1} \mathbb{P}\Big(\mathcal{A}^c_{k+1}\Big|\mathcal{A}_1\cap\ldots\cap\mathcal{A}_k\Big) < \varepsilon. \tag{4.108}$$

However, we have then, for any $\varepsilon > 0$

$$\mathbb{P}(v(t=1,x) > 0 \text{ for all } x\in[-R,R]) \ge \mathbb{P}(\mathcal{A}_{N_\varepsilon}) \ge 1-\varepsilon. \tag{4.109}$$

As $\varepsilon > 0$ is arbitrary, we would have

$$\mathbb{P}(v(t=1,x) > 0 \text{ for all } x\in[-R,R]) = 1, \tag{4.110}$$

finishing the proof. Thus, it suffices to verify (4.107) to finish the proof of Theorem 4.12.

To prove (4.107), let us assume that $\mathcal{A}_k$ occurs, so that $v(t_k,x) \ge \delta_k I_k(x)$. By the comparison principle, it is enough then to show that

$$\mathbb{P}\Big[\text{there exists } \delta_{k+1} > 0 \text{ so that } v(t_{k+1},x) \ge \delta_{k+1} I_{k+1}(x) \text{ for all } |x|\le 2R\Big] \ge 1-\frac{\varepsilon}{N}. \tag{4.111}$$

Here, $v(t,x)$ is the solution of

$$\begin{aligned}&\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + v\dot W, \quad t>t_k, \;\; |x|<2R,\\ &v(t,-2R) = v(t,2R) = 0,\\ &v(t_k,x) = I_k(x).\end{aligned} \tag{4.112}$$

Let us then write

$$\begin{aligned}v(t_{k+1},x) &= \int_{-2R}^{2R} G_R(t_{k+1}-t_k,x,y)I_k(y)dy\\ &+\int_{t_k}^{t_{k+1}}\int_{-2R}^{2R} G_R(t_{k+1}-s,x,y)v(s,y)W(dsdy) = v_1(t_{k+1},x)+v_2(t_{k+1},x).\end{aligned} \tag{4.113}$$

Exercise 4.14 Verify that if N is sufficiently large then $v_1(t_{k+1},x) > 1/10$ on the interval I_{k+1}. This is because the distance between the edges of I_k and I_{k+1} is R/N while the time increment is $t_{k+1}-t_k = 1/N$. Thus the solution would spread over the distance $N^{-1/2} \gg R/N$ during this time, and the region where $v_1(t_{k+1},x) > 1/10$ would cover I_{k+1}.

Thus, to prove (4.107) we need to show that

$$\mathbb{P}\Big[\sup_{|x|\le 2R}|v_2(t_{k+1},x)| \ge \frac{1}{20}\Big] < \frac{\varepsilon}{N}. \tag{4.114}$$

This is reasonable to expect – when N is large, the time interval $[t_k, t_{k+1}]$ is not long enough to let $v_2(t_{k+1},x)$ become large with an overwhelming probability. We will need an exponential moment estimate on w_2 for this. More precisely, we will show that a stochastic integral of the form

$$N(t,x)=\int_0^t\int_{-2R}^{2R}G(t-s,x,y)g(s,y)W(dsdy)$$

with $|g(s,y)|\le K$ almost surely, satisfies a large deviations estimate:

$$\mathbb{P}\Big[\sup_{0\le t\le T}\sup_{|x|\le 2R}|N(t,x)|>\lambda\Big]\le C_R\exp\Big(-\frac{\lambda^2}{C_RT^{1/2}K^2}\Big). \tag{4.115}$$

A minor difficulty is that a priori we do not know that $v(s,y)$ is bounded almost surely, which is what we need to apply (4.115) to v_2. To deal with this, we can consider instead the solution of a modified equation

$$\begin{aligned}&\frac{\partial\tilde v}{\partial t}=\frac{\partial^2\tilde v}{\partial x^2}+\chi(\tilde v)\tilde v\dot W,\quad t>t_k,\quad x\in[-2R,2R]\\&\tilde v(t,-2R)=\tilde v(t,2R)=0,\\&\tilde v(t_k,x)=I_k(x).\end{aligned} \tag{4.116}$$

The smooth cut-off function $\chi(v)$ is such that $\chi(v)=1$ for $0\le v\le 5$ and $\chi(v)=0$ for $v>10$. Note that $v(t,x)=\tilde v(t,x)$ until a stopping time τ:

$$\tau=\inf\{t>t_k:\ \sup_{|x|\le 2R}\tilde v(t,x)=5.\}$$

In addition, we know that $|\tilde v(t,x)|\le 10$ almost surely. We claim that $\tau>t_{k+1}=t_k+1/N$ with a very large probability. Indeed, setting

$$\tilde N(t,x)=\int_{t_k}^t\int_{-2R}^{2R}G_R(t_{k+1}-s,x,y)\tilde v(s,y)W(dsdy)$$

we can use use (4.115) for $\tilde v(t,x)$ to see that

$$\mathbb{P}(\tau<1/N))=\mathbb{P}\Big[\sup_{t_k\le t\le t_{k+1}}\sup_{|x|\le 2R}\tilde v(t,x)>5\Big]\le C\exp\Big(-\frac{C\cdot 25}{(1/N)^{1/2}10^2}\Big)\le C\exp(-CN^{1/2}). \tag{4.117}$$

It is in this estimate on the stopping time that it is helpful from the very beginning to restrict to the Dirichlet problem on a finite interval $[-2R,2R]$. We also have, from (4.115)

$$\mathbb{P}\Big[\sup_{t_k\le t\le t_{k+1}}\sup_{|x|\le 2R}|\tilde N(t,x)|>\frac{1}{20}\Big]\le C\exp\Big(-CN^{1/2}\Big). \tag{4.118}$$

Therefore, we can estimate

$$\begin{aligned}\mathbb{P}\Big[\sup_{|x|\le 2R}|v_2(t_{k+1},x)|\ge\frac{1}{20}\Big]&\le\mathbb{P}(\tau<1/N)+\mathbb{P}\Big[\sup_{0\le t\le T}\sup_{|x|\le 2R}|\tilde N(t,x)|>\frac{1}{20}\Big]\\&\le C\exp(-CN^{1/2}).\end{aligned} \tag{4.119}$$

This proves (4.114), hence (4.107), finishing the proof of Theorem 4.12 except for the large deviations estimate (4.115).

Exercise 4.15 Explain why this proof fails and can not be generalized to the equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \sqrt{u}\dot{W},$$

with a compactly supported initial condition $u_0(x) \geq 0$. Do not worry about the existence and uniqueness issues.

Large deviations for stochastic integrals

Let us now explain where (4.115) comes from. We will work on the whole line for simplicity, and consider

$$N(t,x) = \int_0^t \int_{\mathbb{R}} G(t-s, x-y)u(s,y)W(dsdy),$$

under the assumption $|u(s,y)| \leq K$ almost surely. We will show the analog of (4.115):

$$\mathbb{P}\Big[\sup_{0\leq t\leq T}\sup_{|x|\leq R}|N(t,x)| > \lambda\Big] \leq C_R \exp\Big(-\frac{\lambda^2}{C_R T^{1/2} K^2}\Big). \tag{4.120}$$

Exercise 4.16 Use the scaling of both $G(t,x)$ and of the white noise to verify that it suffices to prove (4.120) for $T = 1$.

Let us freeze the t variable inside the integral and set

$$\bar{N}_t(s,x) = \int_0^s \int_{\mathbb{R}} G(t-r, x-y)u(r,y)W(drdy),$$

so that $\bar{N}_t(t,x) = N(t,x)$. This makes $\bar{N}_t(s,x)$ a martingale in s (with t fixed), by virtue of being a stochastic integral, as the integrand does not depend on s. Consider

$$M_s = \bar{N}_t(s,x) - \bar{N}_t(s,y),$$

so that $M_t = N(t,x) - N(t,y)$. As M_s is a martingale, it is a random time change of a Brownian motion, that is,

$$M_s = B_{\langle M\rangle_s},$$

and, in particular, we have

$$M_t = N(t,x) - N(t,y) = B_{\langle M\rangle_t}.$$

We may estimate the quadratic variation:

$$\langle M\rangle_s = \int_0^s \int_{\mathbb{R}} [G(t-r, x-z) - G(t-r, y-z)]u^2(r,z)dzdr,$$

and

$$\langle M\rangle_t = \int_0^t \int_{\mathbb{R}} [G(t-r, x-z) - G(t-r, y-z)]^2 u^2(r,z)dzdr \leq CK^2|x-y|,$$

for all $0 \le t \le 1$. We deduce that

$$\mathbb{P}(N(t,x) - N(t,y) > \lambda) = \mathbb{P}(B_{\langle M\rangle_t} > \lambda) \le C\mathbb{P}(B_{CK^2|x-y|} > \lambda) \le Ce^{-c\lambda^2/K^2|x-y|}.$$

Switching x and y gives

$$\mathbb{P}(|N(t,x) - N(t,y)| > \lambda) = \mathbb{P}(B_{\langle M\rangle_t} > \lambda) \le C\mathbb{P}(B_{CK^2|x-y|} > \lambda) \le Ce^{-c\lambda^2/K^2|x-y|}. \tag{4.121}$$

In a similar vein, we can write, for $t > s$ and $x \in \mathbb{R}$ fixed:

$$\begin{aligned} N(t,x) - N(s,x) &= \int_0^t \int_{\mathbb{R}} G(t-r, x-z)u(r,z)W(drdy) \\ &\quad - \int_0^s \int_{\mathbb{R}} G(s-r, x-z)u(r,z)W(drdy) \\ &= \int_0^s \int_{\mathbb{R}} [G(t-r, x-z) - G(s-r, x-z)]u(r,z)W(drdy) \\ &\quad + \int_s^t \int_{\mathbb{R}} G(t-r, x-z)u(r,z)W(drdy). \end{aligned}$$

We set

$$A_\tau = \int_0^\tau \int_{\mathbb{R}} [G(t-r, x-z) - G(s-r, x-z)]u(r,z)W(drdy)$$

and

$$B_\tau = \int_s^{s+\tau} \int_{\mathbb{R}} G(t-r, x-z)u(r,z)W(drdy).$$

These are both martingales in τ and

$$N(t,x) - N(s,x) = A_s + B_{t-s}.$$

A simple computation shows that their quadratic variations are bounded by

$$\langle A\rangle_s \le K^2 \int_0^s \int_{\mathbb{R}} |G(t-r, x-z) - G(s-r, x-z)|^2 drdz \le CK^2|t-s|^{1/2},$$

and

$$\langle B\rangle_{t-s} \le K^2 \int_s^t \int_{\mathbb{R}} G^2(t-r, x-z)|^2 dzdr \le CK^2|t-s|^{1/2}.$$

Therefore, we have

$$\mathbb{P}(A_s + B_{t-s} > \lambda) \le \mathbb{P}(A_s > \lambda/2) + \mathbb{P}(B_{t-s} > \lambda/2) \le Ce^{-c\lambda^2/K^2|t-s|^{1/2}}.$$

We conclude that

$$\mathbb{P}(|N(t,x) - N(s,x)| > \lambda) \le Ce^{-c\lambda^2/K^2|t-s|^{1/2}}. \tag{4.122}$$

Now, the proof of (4.115):

$$\mathbb{P}\Big[\sup_{0\le t\le T}\sup_{|x|\le R} |N(t,x)| > \lambda\Big] \le C_R \exp\Big(-\frac{\lambda^2}{C_R T^{1/2} K^2}\Big) \tag{4.123}$$

becomes a real analysis exercise. One connects the point (t,x) to $(0,0)$ on a grid of dyadic points in the (t,x)-plane. Then one estimates the increments between the nearest neighbors using (4.121) and (4.122). Summing up all the differences leads to (4.123).

Exercise 4.17 Fill in the details in the last step in the proof.

5 Spreading in the stochastic heat equation

Spreading in the deterministic equation

Before discussing spreading for the solutions of the stochastic heat equation, let us recall some very basic facts about the solutions of the heat equation with a deterministic linear forcing

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u, \tag{5.1}$$

and an initial condition $u_0(x) \geq 0$ decaying at infinity. We say that $u(t,x)$ spreads with a speed c if for any $c' > c$ we have

$$\limsup_{t\to+\infty} \sup_{|x|\geq c't} u(t,x) = 0, \tag{5.2}$$

while for any $0 \leq c' < c$ we have

$$\liminf_{t\to+\infty} \inf_{|x|\leq c't} u(t,x) = 0, \tag{5.3}$$

Of course, this definition can be applied to other problems than (5.1).

Solutions with compactly supported initial conditions

Solutions of (5.1) with a compactly supported initial condition $u_0(x) \geq 0$ spread with the speed $c_* = 2$. In order to see this, let us assume that $u_0(x) = \mathbb{1}_{[-1,1]}(x)$ and write

$$u(t,x) = e^t \int_{-1}^{1} e^{-|x-y|^2/(4t)} dy. \tag{5.4}$$

Then, for $c > 2$ we have an upper bound

$$u(t,ct) \leq e^t \int_{-1}^{1} e^{-|ct-1|^2/(4t)} dy = 2\exp\Big\{(1-\frac{c^2}{4})t + \frac{c}{2} - \frac{1}{4t}\Big\} \to 0 \text{ as } t\to+\infty. \tag{5.5}$$

On the other hand, for $c \in (0,2)$ we have

$$u(t,x) \geq 2e^t \int_{-1}^{1} e^{-|ct+1|^2/(4t)} dy = 2\exp\Big\{(1-\frac{c^2}{4})t - \frac{c}{2} - \frac{1}{4t}\Big\}, \tag{5.6}$$

hence (5.3) holds. Thus, the front of the solution is located around $x = 2t$, in the sense that the solution is exponentially large at $x \gg 2t$ and it is exponentially small at $x \gg 2t$.

In order to understand what happens around $x = 2t$, let us just look at the heat kernel:

$$u(t,x) = \frac{1}{\sqrt{4\pi t}} e^{t-|x|^2/(4t)}.$$

Let us write $x = 2t + \xi$, then

$$u(t, 2t+\xi) = \frac{1}{\sqrt{4\pi t}} e^{t-|(2t+\xi)|^2/(4t)} = \frac{1}{\sqrt{4\pi t}} e^{-\xi-\xi^2/(4t)}.$$

A general solution with a compactly supported initial condition $u_0(x)$ has an asymptotics

$$u(t,x) \sim \frac{M_0}{\sqrt{4\pi t}} e^{t-|x|^2/(4t)}, \quad M_0 = \int_{\mathbb{R}} u_0(x)dx.$$

Hence, it can be written as

$$u(t,2t+\xi) \sim \frac{M_0}{\sqrt{4\pi t}} e^{-\xi-\xi^2/(4t)} \sim \exp\Big[-\xi - \frac{|\xi|^2}{4t} - \frac{1}{2}\log t + \log M_0 - \frac{1}{2}\log(4\pi)\Big].$$

Therefore, we have an approximation

$$u(t, 2t - \frac{1}{2}\log t + x_0 + \xi) \to \exp(-\xi), \tag{5.7}$$

with the shift x_0 that depends on the initial condition. In other words, the "front" of the solution (the location where $u(t,x) = 1$) is located at

$$X(t) = 2t - \frac{1}{2}\log t + x_0, \tag{5.8}$$

and the solution around this point converges to an exponential $\bar{u}(\xi) = e^{-\xi}$. The profile around the front is not Gaussian – it is an exponential function. Another remarkable point is that the function

$$\tilde{u}(t,x) = \bar{u}(x - X(t)) = e^{-(x-X(t))}$$

is not an exact solution of (5.1): it satisfies an approximate equation

$$\frac{\partial \tilde{u}}{\partial t} + \frac{1}{2t}\frac{\partial \tilde{u}}{\partial x} = \frac{\partial^2 \tilde{u}}{\partial x^2} + \tilde{u}. \tag{5.9}$$

This is quite typical – the limiting profiles need not be exact solutions of the original problem, the can solve an approximate problem instead.

The exponential solutions and pulled propagation

There is another simple way to guess the spreading speed $c_* = 2$ for the solutions of (5.1) with compactly supported initial conditions. Let us look for exponential solutions of this equation of the form

$$u(t,x) = \exp\{-\lambda(x - ct)\}.$$

Inserting this into (5.1) gives

$$c\lambda = \lambda^2 + 1. \tag{5.10}$$

This equation has a positive solution $\lambda > 0$ exists for all $c \geq c_* = 2$. This identifies the spreading speed correctly.

This very simple idea of using the exponential solutions is very useful in all sorts of "pulled front" deterministic reaction-diffusion problems. A simple evidence that the propagation is pulled is the sensitivity of the spreading rate to the precise rate of decay of the initial condition. Let us assume that

$$u_0(x) \sim Ce^{-\lambda x}, \text{ as } x \to +\infty, \tag{5.11}$$

with a decay rate $\lambda < 1$. In other words, there exists $x_0 > 0$ and two constants $C_{1,2} > 0$ such that

$$C_1 e^{-\lambda x} \le u_0(x) \le C_2 e^{-\lambda x}, \text{ for all } x > x_0. \tag{5.12}$$

On the other hand, we assume that $u_0(x)$ is compactly supported on the left: there exists x_1 such that $u_0(x) = 0$ for all $x < x_1$. Then we can find exponential sub- and super-solutions for $u(t,x)$ spreading with the speed c given by (5.10). For the super-solution, we find C so that

$$u_0(x) \le Ce^{-\lambda x},$$

for all $x \in \mathbb{R}$. Then we have, from the maximum principle

$$u(t,x) \le C_0 e^{-\lambda(x-ct)}, \tag{5.13}$$

hence $u(t,x)$ spreads at most with the speed c. On the other hand, given $\lambda < 1$ we can find c from (5.10) but also $\lambda' = 1/\lambda > 1$ that satisfies the same quadratic equation. Then we can find C and C' such that

$$u_0(x) \ge \tilde{u}_0(x) = Ce^{-\lambda x} - C'e^{-\lambda' x},$$

and there is some interval (a,b) such that $\tilde{u}_0(y) > 0$ for all $y \in (a,b)$. It follows that

$$u(t,x) \ge Ce^{-\lambda(x-ct)} - C'e^{-\lambda'(x-ct)}, \quad \text{for all } t > 0 \text{ and } x \in \mathbb{R}. \tag{5.14}$$

In particular, we have

$$u(t, ct+y) > \alpha_0 \text{ for all } t > 0 \text{ and } y \in (a,b). \tag{5.15}$$

Exercise 5.1 Use (5.15) to show that (5.3) holds for all $c' \in [0,c)$.

Hence, solutions with an initial condition that has an exponential decay as in (5.11) with $\lambda < 1$ propagate with the speed $c > 2$ given by (5.10).

Exercise 5.2 Show that if the initial condition has an exponential decay with a rate faster than $\lambda_* = 1$, that is, if (5.11) holds with $\lambda > 1$, then the solution spreads with the speed $c_* = 2$, "as if it were compactly supported".

Spreading in the nonlinear case

As the heat equation fronts are pulled, we have the following phenomenon. Consider the solutions of the nonlinear heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u), \tag{5.16}$$

and an initial condition $u_0(x) \ge 0$ decaying at infinity. We can interpret the function

$$r(u) = \frac{f(u)}{u}$$

as the rate of growth of u.

Exercise 5.3 Assume that the function $r(u)$ is decreasing for $u > 0$, and that $u_0(x)$ is either compactly supported or is exponentially decaying as in (5.12). Show that the spreading rate of the solutions of (5.1) is the same as for the solutions of

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + f'(0)v, \tag{5.17}$$

with $v(0,x) = u_0(x)$.

Spreading in the stochastic case

Let us now consider solutions of the stochastic heat equation

$$\frac{\partial u}{\partial t} = \nu \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{5.18}$$

with a continuous compactly supported initial condition $u_0(x) \geq 0$ such that $u_0(x) \not\equiv 0$. The nonlinearity $f(u)$ is Lipschitz:

$$|f(u) - f(v)| \leq \bar{L}|u - v|, \tag{5.19}$$

and $f(0) = 0$. In addition, we will assume that

$$f(u) \geq \beta u \text{ for all } u \geq 0. \tag{5.20}$$

As the forcing in the stochastic heat equation has mean zero, there is no a priori reason to expect that the solution will spread at a linear speed – one may also expect a diffusive behavior, as in the standard heat equation. And, indeed, since

$$\mathbb{E}\int_{\mathbb{R}} u(t,x)dx = \int_{\mathbb{R}} u_0(x)dx$$

is conserved in time, there can not be "growth everywhere" we have seen in the deterministic equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u.$$

Rather, we should be tracking "propagation of the non-trivial behavior". That is, there is a certain spatial scale $L(t)$ such that for $x \gg L(t)$ "nothing happens yet" – the solution is still very small, while for $|x| \ll L(t)$ we should observe a "non-trivial" behavior, whatever that means.

Spreading of the moments

In order to make this precise, we will judge the non-triviality of the behavior by the size of the second moment. By a vague analogy with the deterministic case and the exponential solutions, we may expect that

$$\mathbb{E}|u(t,x)|^2 \sim \exp(-\lambda_*(x - c_* t), \quad \text{for } x > 0. \tag{5.21}$$

Then we would call c_* the spreading speed of the solutions. According to these expectations, let us define

$$\bar{S}(c) = \limsup_{t\to+\infty} \frac{1}{t} \sup_{|x|>ct} \log \mathbb{E}(|u(t,x)|^2),$$

and

$$\underline{S}(c) = \liminf_{t\to+\infty} \frac{1}{t} \sup_{|x|<ct} \log \mathbb{E}(|u(t,x)|^2).$$

That is, if $\bar{S}(c) < 0$, there can be no peaks in $u(t,x)$ for $x > ct$ with a large probability. On the other hand, if $\underline{S}(c) > 0$, there must be peaks in $u(t,x)$ for $|x| < ct$ with a large probability. Hence, it makes sense to consider

$$\bar{c}_2 = \inf\{c > 0 : \ \bar{S}(c) < 0\}, \tag{5.22}$$

and

$$\underline{c}_2 = \inf\{c > 0 : \ \underline{S}(c) > 0\}. \tag{5.23}$$

If

$$\bar{c}_2 = \underline{c}_2, \tag{5.24}$$

it is natural to call $c_* = \bar{c}_2$ the speed of propagation – the solution is small for $x \gg c_*t$ and there are large peaks at positions $|x| \ll c_*t$. Note that these large peaks still occur with a very small probability – the first moment of the solution is not growing. Hence, the spreading of the second moment does not reflect a typical behavior at a given point. Nevertheless, these are interesting objects to study.

Recall that in the deterministic case (5.1) we know that $\bar{c}_2 = \underline{c}_2 = 2\sqrt{\nu}$. In the stochastic case, it is known that (5.24) holds in the special case $f(u) = \beta u$ – see the recent paper by Chen and Dalang [1]. However, to the best of my knowledge, this question is open for more general nonlinearities $f(u)$ even under our extra assumption (5.20).

The choice of the second moment as a measuring stick is subjective. One could define

$$\bar{S}_p(c) = \limsup_{t\to+\infty} \frac{1}{t} \sup_{|x|>ct} \log \mathbb{E}(|u(t,x)|^p),$$

and

$$\underline{S}_p(c) = \liminf_{t\to+\infty} \frac{1}{t} \sup_{|x|<ct} \log \mathbb{E}(|u(t,x)|^p),$$

and the corresponding speeds $\bar{c}_p$ and $\underline{c}_p$. It has been recently shown in a paper by Nualart [3] that $\bar{c}_p > \bar{c}_2$, and $\bar{c}_p = \underline{c}_p$ for $p > 2$ when $f(u) = \beta u$. Hence, the legitimacy of taking $\bar{c}_2$ even if $\bar{c}_2 = \underline{c}_2$ as the speed is not obvious.

We will prove the following result of Conus and Khoshnevisan [2].

Theorem 5.4 *There exists c_0 so that $\underline{S}(c) > 0$ for all $c \in (0, c_0)$. On the other hand, for any speed $c > c_* = \bar{L}^2/2$ we have $\bar{S}(c) < 0$.*

References

[1] Le Chen and R. Dalang, Moments and growth indices for the nonlinear stochastic heat equation with rough initial conditions, Ann. Prob., **43**, 2015, 3006-3051.

[2] D. Conus, and D. Khoshnevisan, On the existence and position of the farthest peaks of a family of stochastic heat and wave equations, Probab. Theory Related Fields, **152**, 2012, 681–701.

[3] J. Huang, K. Lê, D. Nualart, Large time asymptotic for the parabolic Anderson model driven by spatially correlated noise, Preprint, 2015.

[4] C. Mueller, On the support of solutions to the heat equation with noise, Stochastics and Stochastic Reports, **37**, 1991, 225–245.