Lecture notes for Introduction to SPDE part of 256B, Spring 2024

Lenya Ryzhik*

May 23, 2024

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.

Contents

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

1 Some preliminary basic questions

We would like to be able to make sense of the solutions to equations with rough forces, such as the additive 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 would be natural to think of $u(t,x)$ as a weak solution to the PDE as that would not require differentiation – but if the PDE (unlike the examples above) is nonlinear we would still typically need to know that $u(t,x)$ is a function to consider it as a weak solution, 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 but only a distribution and so is $F(t,x)$. Then how should we interpret a function $f(u(t,x))$? The same problem comes up even in linear equations when one deals with products of the form $u(t,x)F(t,x)$. If both u and F are distributions, then what kind of meaning can we assign to their product? This issue arises, of course, already in the theory of stochastic ordinary differential equations but is much more severe for stochastic partial differential equations.

Another obvious issue is to understand what we would mean by a "rough force" and what kind of rough forces we can allow. If, say, (1.1) were posed on the lattice $\mathbb{Z}^d$, and Δ were a discrete Laplacian, then (1.1) could be viewed as 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 such noises 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

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

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

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

This, however, would presumably require that $F(t,x)$ is a distribution bringing up all the issues we have discussed above.

1.1 Separation of scales between the noise and the solution

In order to explain how the correlation function (1.4) comes about naturally and why allowing such noises is helpful, let us consider the forced linear heat equation

$$u_t = \Delta u + Ng(t,x). \tag{1.5}$$

Here, we assume that $g(t,x)$ is a smooth space-time stationary mean-zero random field with a correlation function $R(t,x)$:

$$\mathbb{E}(g(t,x)) = 0, \quad \text{for all } t \in \mathbb{R},\ x \in \mathbb{R}^d, \tag{1.6}$$

and

$$\mathbb{E}(g(s,y)g(t,x)) = R(t-s,x-y), \quad \text{for all } t,s \in \mathbb{R} \text{ and } x,y \in \mathbb{R}^d. \tag{1.7}$$

The parameter N measures the strength of the noise.

The main assumption that quantifies the idea that $g(t,x)$ is a "small-scale noise" is that the initial condition for (1.5) varies on a scale much larger than that of $g(t,x)$. To formalize this, we introduce a small parameter $\varepsilon \ll 1$ that measures the ratio of these two scales and consider an initial condition for (1.5) of the form

$$u(0,x) = u_0(\varepsilon x). \tag{1.8}$$

Then, setting

$$u^{\varepsilon}(t,x) = u\Big(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\Big), \tag{1.9}$$

we obtain the following equation for the function $u^{\varepsilon}(t,x)$:

$$\begin{aligned} u_t &= \Delta u^{\varepsilon} + \frac{N}{\varepsilon^2} g\Big(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\Big), \\ u(0,x) &= u_0(x). \end{aligned} \tag{1.10}$$

Now, the noise

$$g_{\varepsilon}(t,x) = \frac{N}{\varepsilon^2} g\Big(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\Big) \tag{1.11}$$

varies on a small scale both in time and in space and is, therefore, some form of a rough forcing. Its correlation function is

$$R_{\varepsilon}(t,x) = \mathbb{E}(g_{\varepsilon}(s,y) g_{\varepsilon}(s+t, y+x)) = \frac{N^2}{\varepsilon^4} R\Big(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\Big). \tag{1.12}$$

Observe that if we take ε_N so that

$$\frac{N^2}{\varepsilon_N^4} = \frac{1}{\varepsilon_N^{d+2}}, \tag{1.13}$$

then

$$R_{\varepsilon}(t,x) \to \delta(t,x), \quad \text{as } \varepsilon \to 0, \tag{1.14}$$

recovering the δ-correlation in space and time in (1.4). In other words, if the length scale L_{ε} of the initial condition and the strength N of the noise are related by

$$L_{\varepsilon} = \varepsilon^{-1} = N^{2/(d-2)}, \tag{1.15}$$

then we would expect that $u^{\varepsilon}(t,x)$ converges, as $\varepsilon \to 0$, to a solution of the forced linear heat equation (1.1):

$$\frac{\partial u}{\partial t} = \Delta u + F(t,x), \tag{1.16}$$

with the δ-correlated noise $F(t,x)$. Moreover, such δ-correlated noise essentially has to appear in any reasonable situation where we expect to have a scale separation between the solution and the noise, as in the above example.

Let us make a comment on the dependence of the above simple analysis on the spatial dimension. We write (1.15) as an expression of the "critical" noise strength in terms of the scale separation parameter $\varepsilon \ll 1$:

$$N_d = \varepsilon^{-(d-2)/2}, \tag{1.17}$$

and note that

$$\begin{aligned} &N_d \ll 1, \quad d = 1, \\ &N_d = O(1), \quad d = 2, \\ &N_d \gg 1, \quad d = 3. \end{aligned} \tag{1.18}$$

In other words, in one dimension, even a weak microscopic noise leads to a non-trivial effect, in two dimensions the microscopic noise has to have strength of order one lead to the delta-correlated noise, and in dimensions three and higher the microscopic noise needs to be strong to get the delta-correlated noise in the limit. This distinction between different dimensions is very typical in stochastic partial differential equations.

1.2 Is the solution a function?

Disregarding the question of a careful definition of such δ-correlated noise, 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 to (1.1) in $\mathbb{R}^d$ is

$$u(t,x)=\int_0^t\int_{\mathbb{R}^d}G(t-s,x-y)F(s,y)dsdy. \tag{1.19}$$

Here,

$$G(t,x)=\frac{1}{(4\pi t)^{d/2}}e^{-|x|^2/(4t)} \tag{1.20}$$

is the standard heat kernel. Once again, we do not for the moment define what exactly we mean by the integral in (1.20) but rather suppose that $F(t,x)$ is a a random field with the delta-correlation function given by (1.4).

The "function" $u(t,x)$ given by (1.19) has to be a stationary field in x. This can be seen immediately simply from the translation invariance of the heat equation and the space-time stationarity of the forcing $F(t,x)$. Alternatively, one may deduce that directly from (1.19) and the space-time stationarity of $F(t,x)$ if we knew exactly what the integral in (1.19) means. Hence, we can not possibly expect any decay of $u(t,x)$ as $|x|\to+\infty$ but we may still ask how large $u(t,x)$ should be. Let us formally compute its point-wise second moment, using the assumption (1.4) about the two-point correlation function of the field $F(t,x)$:

$$\begin{aligned}\mathbb{E}\Big[|u(t,x)|^2\Big]&=\int_0^t\int_0^t\int_{\mathbb{R}^{2d}}G(t-s,x-y)G(t-s',x-y')\mathbb{E}[F(s,y)F(s',y')]dsds'dydy'\\&=\int_0^t\int_{\mathbb{R}^d}|G(t-s,x-y)|^2dsdy=\int_0^t\int_{\mathbb{R}^d}|G(s,y)|^2dsdy\\&=\frac{1}{(4\pi)^d}\int_0^t\int_{\mathbb{R}^d}e^{-|y|^2/(2s)}\frac{dyds}{s^d}=C_d\int_0^t\frac{ds}{s^{d/2}}.\end{aligned} \tag{1.21}$$

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. This is related to the comments made at the end of last section: spatial dimension plays a crucial role in all such considerations.

1.3 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 (1.19) at least could potentially make sense as a distribution in x, pointwise in time. We multiply (1.19) by a test function $\phi\in C_c^\infty(\mathbb{R}^d)$ and integrate:

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

Next, we compute the second moment of $\Phi(t)$, once again using the assumption (1.4) about the two-point correlation function of the field $F(t,x)$:

$$
\begin{aligned}
\mathbb{E}\Big[|\Phi(t)|^2\Big] &= \int_0^t \int_0^t \int_{\mathbb{R}^{4d}} G(t-s,x-y)G(t-s',x'-y')\mathbb{E}[F(s,y)F(s',y')] \\
&\times \phi(x)\phi(x')dsds'dydy'dxdx' \\
&= \int_0^t \int_{\mathbb{R}^{3d}} G(t-s,x-y)G(t-s,x'-y)\phi(x)\phi(x')dsdydxdx' \\
&= \int_0^t \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} G(t-s,y-x)\phi(x)dx\Big)\Big(\int_{\mathbb{R}^d} G(t-s,y-x')\phi(x')dx'\Big)dyds \\
&= \int_0^t \int_{\mathbb{R}^d} v^2(t-s,y)dyds = \int_0^t \int_{\mathbb{R}^d} v^2(s,y)dyds.
\end{aligned}
\tag{1.22}
$$

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

$$
\frac{\partial v}{\partial t} = \Delta v, \tag{1.23}
$$

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 that one can make sense of $u(t,x)$ given by (1.19) as a distribution in $x \in \mathbb{R}^d$, for a fixed time $t \geq 0$, in any spatial dimension. Thus, in dimensions $d \geq 2$ one, generally, would expect solutions of SPDEs to be distributions and not functions.

This phenomenon is a serious obstacle, as one is often interested in solutions to 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 a solution to a deterministic PDE as $N \to +\infty$, such as, in the simplest case, the heat equation:

$$
\frac{\partial u}{\partial t} = \Delta u. \tag{1.24}
$$

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

$$
\frac{\partial u}{\partial t} = \Delta u + \text{Noise}. \tag{1.25}
$$

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

$$
\frac{\partial u}{\partial t} = \Delta u + \sqrt{u} \cdot \text{Noise}. \tag{1.26}
$$

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

1.4 The randomly forced wave equation

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

$$\frac{\partial^2 v}{\partial t^2} - \frac{\partial^2 v}{\partial x^2} = F(t,x), \tag{1.27}$$

in one dimension, and with the initial condition

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

Its solution is given by the d'Alembert formula

$$v(t,x) = \frac{1}{2}\int_0^t \int_{x-(t-s)}^{x+(t-s)} F(s,y)dyds. \tag{1.28}$$

Let us now assume that $F(s,y)$ is a white noise with the two-point correlation function (1.4), as in 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 (1.28). 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 to the forced wave equation can be obtained via spherical means, leading to similar issues.

2 The white noise and the Wiener integral

2.1 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. If G is a Gaussian random field, 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 of any such process is non-negative definite in the following sense: for any $t_1,\dots,t_k$ and $z_1,\dots,z_k \in \mathbb{R}$, we have

$$\sum_{j,m=1}^{k} C(t_j,t_m)z_j z_m \geq 0. \tag{2.1}$$

This is because

$$\begin{aligned} 0 \leq \mathbb{E}\Big|\sum_{j=1}^{k}(G(t_j)-\mu(t_j))z_j\Big|^2 &= \sum_{m,j=1}^{k}\mathbb{E}\big[(G(t_j)-\mu(t_j))(G(t_m)-\mu(t_m))\big]z_j z_m \\ &= \sum_{j,m=1}^{k} C(t_j,t_m)z_j z_m. \end{aligned} \tag{2.2}$$

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 classical examples of a Gaussian process is the Brownian motion. In that case, the index set is $T = [0,+\infty)$, the mean is $\mu(t) \equiv 0$, and two-point correlation function is $C(s,t) = \min(s,t)$. Let us check that this covariance is nonnegative-definite: given any $t_1,\dots,t_k \in [0,+\infty)$ and $z_1,\dots,z_k \in \mathbb{R}$, we compute

$$\begin{aligned}\sum_{j,m=1}^{k} C(t_j,t_m)z_jz_m &= \sum_{j,m=1}^{k} z_jz_m \min(t_j,t_m) = \sum_{j,m=1}^{k} z_jz_m \int_0^{+\infty} \mathbb{1}_{[0,t_j]}(s)\mathbb{1}_{[0,t_m]}(s)ds \\ &= \int_0^{+\infty} \Big|\sum_{j=1}^{k} z_j \mathbb{1}_{[0,t_j]}(t)\Big|^2 dt \geq 0.\end{aligned} \tag{2.3}$$

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

$$C(s,t) = \mathbb{E}(b(s)b(t)) = s \wedge t - st. \tag{2.4}$$

To check that this function is positive-definite, given any $t_1,\dots,t_k \in [0,1]$ and $z_1,\dots,z_k \in \mathbb{R}$, we set

$$\varphi(t) = \sum_{j=1}^{k} z_k \mathbb{1}_{[0,t_j]}(t), \tag{2.5}$$

and observe that

$$\begin{aligned}\sum_{j,m=1}^{k} C(t_j,t_m)z_jz_m &= \sum_{j,m=1}^{k} z_jz_m\big(\min(t_j,t_m) - t_jt_m\big) \\ &= \sum_{j,m=1}^{k} z_jz_m\Big(\int_0^1 \mathbb{1}_{[0,t_j]}(s)\mathbb{1}_{[0,t_m]}(s)ds - \int_0^1\int_0^1 \mathbb{1}_{[0,t_j]}(s)\mathbb{1}_{[0,t_m]}(s')dsds'\Big) \\ &= \int_0^1 \varphi^2(s)ds - \Big(\int_0^1 \varphi(s)ds\Big)^2 \geq 0,\end{aligned} \tag{2.6}$$

by the Cauchy-Schwartz inequality.

A remarkable property of the Brownian bridge is that

$$\mathbb{E}(b(0)^2) = \mathbb{E}(b(1)^2) = 0, \tag{2.7}$$

as follows immediately from (2.4). That is, the process $b(s)$ starts at $b(0) = 0$ and ends at $b(1) = 0$, almost surely. Let us also observe that if $B(t)$ is a Brownian motion, then

$$b(t) = B(t) - tB(1) \tag{2.8}$$

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

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

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 consider the process

$$X(t) = e^{-t/2} B(e^t), \tag{2.9}$$

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

$$C(s,t) = e^{-(s+t)/2} \min\Big(e^s, e^t\Big) = e^{(t-s)/2}. \tag{2.10}$$

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

$$C(s,t) = e^{-|t-s|/2}. \tag{2.11}$$

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

$$\mathbb{E}(X^2(t)) = 1, \tag{2.12}$$

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

2.2 The white noise

2.2.1 Definition of the white noise

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 with the following two properties:
(1) For any collection of measurable sets $A_1, \dots, A_k$ the random vector $(\dot{W}(A_1), \dots, \dot{W}(A_k))$ is mean-zero Gaussian with

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

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

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

Under these two assumptions, we can compute the covariance of $\dot{W}$: for any two measurable sets A and B we have

$$\begin{aligned}\mathbb{E}(\dot{W}(A)\dot{W}(B)) &= \mathbb{E}\big[(\dot{W}(A \cap B) + \dot{W}(A \setminus B))(\dot{W}(A \cap B) + \dot{W}(B \setminus A))\big] \\ &= \mathbb{E}[\dot{W}(A \cap B)^2] = \nu(A \cap B).\end{aligned} \tag{2.13}$$

Alternatively, we could have defined the white noise as a mean-zero Gaussian random field on $\mathcal{M}$ with the covariance

$$C(A,B) = \mathbb{E}(\dot{W}(A)\dot{W}(B)) = \nu(A \cap B). \tag{2.14}$$

Once again, we recall that according to the Kolmogorov extension theorem, the only requirement needed to make sure that the white noise is well-defined is that the covariance function given by (2.14) is nonnegative-definite. Indeed, for any finite collection $A_1, \dots, A_k \in \mathcal{M}$ of measurable sets and any $z_1, \dots, z_k \in \mathbb{R}$, we have

$$\begin{aligned}\sum_{i,j=1}^{k} z_i z_j C(A_i, A_j) = \sum_{i,j=1}^{k} z_i z_j \nu(A_i \cap A_j) = \sum_{i,j=1}^{k} \int_E z_i z_j \mathbb{1}_{A_i}(x) \mathbb{1}_{A_j}(x) d\nu(x) \\ = \int_E \Big| \sum_{j=1}^{k} z_j \mathbb{1}_{A_j}(x) \Big|^2 d\nu(x) \geq 0.\end{aligned} \tag{2.15}$$

The Brownian motion can be constructed from white noise as follows. Let $\dot{W}$ be a white noise on $\mathbb{R}$ and define $B_t = \dot{W}([0,t])$. We claim that B_t is a Brownian motion. Indeed, for any $t_1, \dots, t_k \geq 0$, the vector $(B_{t_1}, \dots, B_{t_k})$ is Gaussian by the definition of white-noise, and we have $\mathbb{E}(B_t) = 0$ for all $t \geq 0$. The corresponding covariance function is

$$\mathbb{E}(B_t B_s) = \mathbb{E}(\dot{W}([0,t])\dot{W}([0,s])) = |[0,t] \cap [0,s]| = t \wedge s. \tag{2.16}$$

Thus, B_t is a Brownian motion.

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 $t = (t_1, \dots, t_n) \in \mathbb{R}^n$, we set

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

and define the Brownian sheet as

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

This definition can be naturally extended to t having some negative components by define the rectangle R_t as the n-dimensional rectangle with the extreme vertices 0 and t and setting

$$W(t) = \dot{W}(R_t).$$

Exercise 2.1 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.

2.2.2 White-noise as an L^2-valued measure

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. In general, we may start with a subset E of $\mathbb{R}^d$ and a σ-finite measure ν defined on the algebra $\mathcal{B}$ of ν-measurable subsets of E. We will restrict ourselves to the situation when the set $E = \mathbb{R}^d$, the measure ν is the Lebesgue measure on $\mathbb{R}^d$, and $\mathcal{B}$ is the σ-algebra of the Borel sets on $\mathbb{R}^d$.

Let Φ be a real-valued random set function on $\mathcal{B}$ such that $\Phi(A)$ is an $L^2(\Omega)$ random variable for each $A \in \mathcal{B}$. Here and below we will denote by Ω the underlying probability space. We denote by

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

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

$$\mathbb{R}^d = \bigcup_n E_n,$$

and

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

Then we say that the function Φ is σ-finite. It is finitely additive if for any finite collection of pairwise disjoint measurable sets $A_1, \ldots, A_k$, we have

$$\Phi\Big(\bigcup_{j=1}^{k} A_k\Big) = \sum_{j=1}^{k} \Phi(A_j). \tag{2.17}$$

The function Φ is countably additive if, for each n, given that $A_j \subset E_n$ and A_j is a decreasing sequence of sets of a finite Lebesgue measure with an empty intersection, then

$$\lim_{j\to+\infty} \Phi(A_j) = 0. \tag{2.18}$$

Then we say that Φ is a σ-finite L^2-valued measure. The equalities in (2.17) and (2.18) should hold in $L^2(\Omega)$.

Lemma 2.2 *The white noise $\dot{W}$ is an $L^2(\Omega)$-valued countably additive measure on $\mathcal{B}$.*

Proof. Let A_1 and A_2 be a pair of disjoint measurable sets. Then, we have

$$\begin{aligned}
&\mathbb{E}\Big(\big|\dot{W}(A_1 \cup A_2) - \dot{W}(A_1) - \dot{W}(A_2)\big|^2\Big) \\
&= \mathbb{E}\Big(\big(\dot{W}(A_1 \cup A_2) - \dot{W}(A_1) - \dot{W}(A_2)\big)\big(\dot{W}(A_1 \cup A_2) - \dot{W}(A_1) - \dot{W}(A_2)\big)\Big) \\
&= |A_1 \cup A_2| - |A_1| - |A_2| - |A_1| + |A_1| - |A_2| + |A_2| = 0.
\end{aligned} \tag{2.19}$$

It follows that

$$\dot{W}(A_1 \cup A_2) = \dot{W}(A_1) + \dot{W}(A_2), \quad \text{a.s.} \tag{2.20}$$

More generally, if $A_1, \ldots, A_k$ is a collection of pairwise disjoint measurable sets, an induction argument shows that

$$\dot{W}\Big(\bigcup_{j=1}^{k} A_j\Big) = \sum_{j=1}^{k} \dot{W}(A_j), \quad \text{a.s.} \tag{2.21}$$

To finish the proof, we need to show that if $B_1 \supseteq B_2 \supseteq \dots \supseteq B_k \supseteq \dots$ is a countable collection of nested measurable sets of a finite measure, such that

$$\bigcap_{j=1}^{\infty} B_j = \emptyset,$$

then

$$\lim_{n\to+\infty} \mathbb{E}(\dot W(B_n))^2 = 0. \tag{2.22}$$

To see that (2.22) holds, we simply observe that

$$\mathbb{E}(\dot W(B_n))^2 = |B_n| \to 0,$$

by the nested sets property of the Lebesgue measure. □

2.2.3 The white noise does not have a bounded total variation

Let us now show that $\dot W$ does not have a bounded total variation almost surely. The proof is very much as what one does in the standard proof of the Ito formula. Consider the square increments sum

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

Note first that

$$\mathbb{E}S_n = \sum_{j=0}^{2^n-1} \Big|\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big| = 1. \tag{2.23}$$

Next, we claim that

$$\lim_{n\to\infty} S_n = 1, \quad \text{a.s.} \tag{2.24}$$

Indeed, on one hand we have (2.23), and on the other, we may compute the variance, using the independence of $\dot W(A)$ and $\dot W(B)$ for disjoint sets A and B:

$$\begin{aligned}
\mathbb{E}(S_n-1)^2 &= \mathbb{E}\Big(\sum_{j=0}^{2^n-1}\Big[\Big|\dot W\Big(\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big|\Big]\Big)^2 \\
&= \sum_{j,m=0}^{2^n-1} \mathbb{E}\Big\{\Big[\Big|\dot W\Big(\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big|\Big]\Big[\Big|\dot W\Big(\Big[\frac{m}{2^n}, \frac{m+1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[\frac{m}{2^n}, \frac{m+1}{2^n}\Big]\Big|\Big]\Big\} \\
&= \sum_{j=0}^{2^n-1} \mathbb{E}\Big\{\Big[\Big|\dot W\Big(\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big|\Big]^2\Big\} = 2^n\mathbb{E}\Big\{\Big[\Big|\dot W\Big(\Big[0, \frac{1}{2^n}\Big]\Big)\Big|^2 - \Big|\Big[0, \frac{1}{2^n}\Big]\Big|\Big]^2\Big\} \\
&= 2^n\Big(3\cdot\frac{1}{2^{2n}} - \frac{2}{2^{2n}} + \frac{1}{2^{2n}}\Big) = \frac{1}{2^{n-1}}.
\end{aligned} \tag{2.25}$$

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

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

in the last step. It follows from (2.25) 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, so that (2.24) holds.

Exercise 2.3 Use (2.24) to obtain

$$\lim_{n\to\infty} \sum_{j=0}^{2^n-1} \Big|\dot{W}\Big(\Big[\frac{j}{2^n}, \frac{j+1}{2^n}\Big]\Big)\Big| = +\infty, \tag{2.26}$$

also almost surely.

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

2.3 The Wiener integral

We have seen that the white noise does not have a bounded variation a.s. On the other hand, we also know that it is an $L^2(\Omega)$-valued measure, so that one can still hope that one can define integration with respect to it in an $L^2(\Omega)$ sense. This is what we do next.

2.3.1 Definition of the Wiener integral

We first define the Wiener integral for simple functions of the form

$$h(x) = \sum_{k=1}^{N} c_k \mathbb{1}(x \in A_k).$$

Here, we assume that $A_k \subseteq \mathbb{R}^d$ are Borel sets such that $|A_k| < +\infty$ for all $k = 1, \dots, N$, and $A_j \cap A_k = \emptyset$ for $j \ne k$. Then, the Wiener integral of $h(x)$ is defined as

$$\int_{\mathbb{R}^d} h(x)\dot{W}(dx) = \sum_{k=1}^{N} c_k \dot{W}(A_k). \tag{2.27}$$

This definition has the following properties: first, for any two simple functions $h_1(x)$ and $h_2(x)$ and any $a_1, a_2 \in \mathbb{R}$ we have

$$\int (a_1 h_1(x) + a_2 h_2(x))\dot{W}(dx) = a_1 \int h_1(x)\dot{W}(dx) + \int h_2(x)\dot{W}(dx). \tag{2.28}$$

Second, the random variables

$$\int h(x)\dot{W}(dx) \tag{2.29}$$

form a mean-zero Gaussian random field over the collection $\mathcal{H}(\mathbb{R}^d)$ of all simple Lebesgue measurable functions, with the covariance

$$\mathbb{E}\Big(\int h_1\dot{W}(dx) \int h_2(x)\dot{W}(dx)\Big) = \int h_1(x)h_2(x)dx. \tag{2.30}$$

To check the covariance formula (2.30), we take two simple functions

$$h_1(x) = \sum_{k=1}^{N} c_k \mathbb{1}(x \in A_k), \quad h_2(x) = \sum_{k=1}^{M} d_k \mathbb{1}(x \in B_k)$$

and compute

$$\begin{aligned}
&\mathbb{E}\Big(\int h_1 \dot{W}(dx) \int h_2(x)\dot{W}(dx)\Big) = \sum_{j=1}^{N}\sum_{k=1}^{M} c_j d_m \mathbb{E}\Big(\dot{W}(A_j)\dot{W}(B_k)\Big)\\
&= \sum_{j=1}^{N}\sum_{k=1}^{M} c_j d_m |A_j \cap B_k| = \sum_{j=1}^{N}\sum_{k=1}^{M} c_j d_m \int \mathbb{1}(x \in A_j)\mathbb{1}(x \in B_k)dx = \int h_1(x)h_2(x)dx,
\end{aligned} \tag{2.31}$$

which is (2.30).

It follows from (2.30) that given two simple functions $h, g \in \mathcal{H}(\mathbb{R}^d)$ and the corresponding integrals

$$I_h(\omega) = \int h(x)\dot{W}(dx), \quad I_g(\omega) = \int g(x)\dot{W}(dx), \quad \omega \in \Omega,$$

we have the isometry

$$\mathbb{E}|I_h - I_g|^2 = \int (h^2(x) + g^2(x) - 2h(x)g(x))dx = \|h - g\|_{L^2(\mathbb{R}^d)}. \tag{2.32}$$

This is the direct analog of the Ito isometry for the Ito integral in the special case of deterministic integrands.

Now, with (2.32) in hand, given any function $h \in L^2(\mathbb{R}^d)$, we can find a sequence of simple functions $h_n \in \mathcal{H}(\mathbb{R}^d)$ that converges to h in $L^2(\mathbb{R}^d)$. Then, the sequence

$$I_{h_n} = \int h_n(x)\dot{W}(dx) \tag{2.33}$$

is Cauchy in $L^2(\Omega)$, by the Ito isometry property (2.32). Thus, there exists the limit

$$I(h) := \lim_{n\to+\infty} I_{h_n}, \quad \text{in } L^2(\Omega), \tag{2.34}$$

that we define to be the Wiener integral

$$I(h) = \int h(x)\dot{W}(x). \tag{2.35}$$

This integral inherits the mean-zero property:

$$\mathbb{E}\Big(\int f(x)\dot{W}(dx)\Big) = 0, \quad \text{for any } f \in L^2(\mathbb{R}^d), \tag{2.36}$$

as well as the Ito isometry (2.32): we have

$$\mathbb{E}\Big(\int f_1(x)\dot{W}(dx) - \int f_2(x)\dot{W}(dx)\Big)^2 = \int |f(x) - g(x)|^2 dx, \quad \text{for any } f_1, f_2 \in L^2(\mathbb{R}^d). \tag{2.37}$$

Let us emphasize that the Wiener integral was defined above only for non-random functions. We will later extend this definition to integrals of random functions as well but for now we leave this extension open.

2.3.2 Integration by parts formula for the Wiener integral

The Wiener integral can be re-phrased using the following integration by parts formula.

Proposition 2.4 *Let $B(x)$ be the d-dimensional Brownian sheet constructed from a white noise $\dot{W}(x)$. Then, for all $\phi \in C_c^\infty(\mathbb{R}^d)$ we have*

$$\int \phi(x)\dot{W}(x)dx = (-1)^d \int \frac{\partial^d \phi(x)}{\partial x_1 \dots \partial x_d} B(x)dx, \quad a.s. \tag{2.38}$$

Proof. We will only consider the case $d = 1$ for simplicity of notation. In that case, the one-dimensional Brownian sheet $B(x) = \dot{W}[0, x]$ is a two-sided Brownian motion. We need to show that

$$\int \phi(x)\dot{W}(dx) = -\int \phi'(x)B(x)dx, \quad \text{a.s.}, \tag{2.39}$$

for any $\phi \in C_c^\infty(\mathbb{R})$. Let us assume without loss of generality that $\phi(x)$ is supported inside the interval $[0, 1]$. We denote by $\phi_n(x)$ the piece-wise constant approximation of $\phi(x)$:

$$\phi_n(x) = \phi\Big(\frac{j}{n}\Big), \quad \text{for } x \in \Big[\frac{j}{n}, \frac{j+1}{n}\Big). \tag{2.40}$$

Then, we have, by the definition of the Brownian motion as the one-dimensional Brownian sheet,

$$\begin{aligned} \int \phi_n(x)\dot{W}(dx) &= \sum_{j=0}^{n-1} \phi\Big(\frac{j}{n}\Big)\dot{W}\Big(\Big[\frac{j}{n}, \frac{j+1}{n}\Big)\Big) = \sum_{j=0}^{n-1} \phi\Big(\frac{j}{n}\Big)\Big(B\Big(\frac{j+1}{n}\Big) - B\Big(\frac{j}{n}\Big)\Big) \\ &= -\sum_{j=1}^{n} B\Big(\frac{j}{n}\Big)\Big[\phi\Big(\frac{j}{n}\Big) - \phi\Big(\frac{j-1}{n}\Big)\Big]. \end{aligned} \tag{2.41}$$

The left side above converges in $L^2(\Omega)$, as $n \to +\infty$, to

$$\int \phi(x)\dot{W}(dx).$$

The right side of (2.41) converges as $n \to +\infty$, almost surely, to the Riemann integral

$$-\int \phi'(t)B(t)dt.$$

Now, (2.38) follows. □

Let us comment that we used the continuity of the Brownian motion in the very last step above. It will follow from the Kolmogorov criterion that we will prove in Theorem 3.4 below. The same comment applies to the Brownian sheet in dimensions $d > 1$.

2.3.3 A Fubini theorem for stochastic convolutions

Given $f \in L^2(\mathbb{R}^d)$, we define its stochastic convolution with white noise as

$$(f \star \dot{W})(x) = \int f(x-y)\dot{W}(dy). \tag{2.42}$$

By the Ito isometry, $f \star \dot{W}(x)$ is an $L^2(\Omega)$ random variable for each $x \in \mathbb{R}^d$ fixed, with

$$\mathbb{E}|f \star \dot{W}(x)|^2 = \int |f(x-y)|^2 dy = \|f\|_{L^2}^2. \tag{2.43}$$

Let us comment on the joint measurability of $(f \star w)(x, \omega)$ in $x \in \mathbb{R}^d$ and $\omega \in \Omega$. First, if $f \in C_c^\infty(\mathbb{R}^d)$ is infinitely differentiable then by the Ito isometry we have

$$\mathbb{E}\Big(|f \star \dot{W}(x) - f \star \dot{W}(y)|^2\Big) = \int_{\mathbb{R}^d} |f(x-z) - f(y-z)|^2 dz \le C_f \|x-y\|^2. \tag{2.44}$$

It follows from the Kolmogorov criterion in Theorem 3.4 below that $f \star \dot{W}(x)$ has a continuous modification. Hence, in particular, it is measurable in x and ω. The general case of $f \in L^2(\mathbb{R}^d)$ follows by the density of $C_c^\infty(\mathbb{R}^d)$ in $L^2(\mathbb{R}^d)$ and the Ito isometry of the Wiener integral.

We will extensively use the following Fubini type theorem for stochastic convolutions.

Proposition 2.5 *If $f \in L^2(\mathbb{R}^d)$ and μ is a finite Borel measure on $\mathbb{R}^d$ then*

$$\int (f \star \dot{W})(x)\mu(dx) = \int (\tilde{f} \star \mu)(x)\dot{W}(dx), \quad a.s. \tag{2.45}$$

Here, we have set $\tilde{f}(y) = f(-y)$.

Proof. First, suppose that $f \in C_c^\infty(\mathbb{R}^d)$. In that case, it follows from the integration by parts formula in Proposition 2.4 that

$$\int (f \star \dot{W})(x)\mu(dx) = (-1)^d \int \Big(\int \frac{\partial^d f(x-y)}{\partial y_1 \dots \partial y_d} B(y)dy\Big)\mu(dx), \quad \text{a.s.} \tag{2.46}$$

As $B(y)$ is continuous, we can use the standard Fubini theorem to obtain

$$\begin{aligned} \int (f \star \dot{W})(x)\mu(dx) &= (-1)^d \int \Big(\int \frac{\partial^d f(x-y)}{\partial y_1 \dots \partial y_d} \mu(dx)\Big) B(y)dy \\ &= (-1)^d \int \frac{\partial^d}{\partial y_1 \dots \partial y_d} \Big(\int f(x-y)\mu(dx)\Big) B(y)dy \\ &= (-1)^d \int \frac{\partial^d}{\partial y_1 \dots \partial y_d} (\tilde{f} \star \mu)(y) B(y)dy = \int (\tilde{f} \star \mu)(y)\dot{W}(dy), \end{aligned} \tag{2.47}$$

which is (2.45). The general case of $f \in L^2(\mathbb{R}^d)$ can be done by approximations of $f(x)$ by smooth functions. □

3 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.

3.1 Modulus of continuity from an integral inequality

We first prove a real-analytic result that allows to translate an integral bound on the increments of a function into a point-wise modulus of continuity. In the theorem below, we assume that 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 3.1 *Let f be a measurable function on $R_1 \subset \mathbb{R}^d$ such that*

$$B := \int_{R_1} \int_{R_1} \psi\Big(\frac{f(y) - f(x)}{p(|y-x|/\sqrt{d})}\Big) dx dy < +\infty. \tag{3.1}$$

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

$$|f(y) - f(x)| \le 8 \int_0^{|y-x|} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big) dp(u). \tag{3.2}$$

If f is continuous, then (3.2) holds for all x and y.

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

$$\text{if } x, y \in Q, \text{ then } |y - x| \le \sqrt{d} e(Q). \tag{3.3}$$

The functions p and ψ are increasing for positive arguments and ψ is even. Hence, (3.1) and (3.3) imply

$$\int_Q \int_Q \psi\Big(\frac{f(y) - f(x)}{p(e(Q))}\Big) dx dy \le B, \tag{3.4}$$

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

$$p(e(Q_j)) = \frac{1}{2} p(e(Q_{j-1})), \tag{3.5}$$

and denote

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

Since $p(0) = 0$ and $p(r) > 0$ for $r > 0$, it follows from (3.5) that $r_j \to 0$, and the cubes converge down to a point. As the function ψ is convex, we have, by Jensen's inequality

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

by (3.4). We conclude that

$$|f_j - f_{j-1}| \le p(r_{j-1})\psi^{-1}\Big(\frac{B}{|Q_{j-1}||Q_j|}\Big). \tag{3.6}$$

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 (3.2). The definition (3.5) of Q_j means that

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

hence we may write

$$|f_j - f_{j-1}| \le 4\psi^{-1}\Big(\frac{B}{|Q_{j-1}||Q_j|}\Big)|p(r_{j+1}) - p(r_j)|. \tag{3.7}$$

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

$$|f_j - f_{j-1}| \le 4\psi^{-1}\Big(\frac{B}{u^{2d}}\Big)|p(r_{j+1}) - p(r_j)|, \quad \text{for all } r_{j+1} \le u \le r_j. \tag{3.8}$$

We deduce that

$$|f_j - f_{j-1}| \le 4\int_{r_{j+1}}^{r_j} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big)dp(u). \tag{3.9}$$

Summing over j gives

$$\limsup_{j\to+\infty} |f_j - f_0| \le 4\int_0^{r_0} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big)dp(u). \tag{3.10}$$

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

$$|f(x) - f_0| \le 4\int_0^{|x-y|} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big)dp(u), \tag{3.11}$$

and

$$|f(y) - f_0| \le 4\int_0^{|x-y|} \psi^{-1}\Big(\frac{B}{u^{2d}}\Big)dp(u), \tag{3.12}$$

proving (3.2). □

3.2 The Kolmogorov theorem

We may now apply Theorem 3.1 to various stochastic processes. We begin with the Kolmogorov theorem. Recall that a random field $X'(t)$, $t \in \mathbf{T}$, is a modification of a random field $X(t)$, $t \in \mathbf{T}$ if

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

Exercise 3.2 Construct an example where X' is a modification of X but

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

Theorem 3.3 *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 random field. Suppose there are constants $k > 1$, $C > 0$ and $\varepsilon > 0$ so that for all $s, t \in \mathbf{T}$, we have*

$$\mathbb{E}\big(|X(t) - X(s)|^k\big) \le C|t - s|^{d+\varepsilon}. \tag{3.13}$$

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

$$|X(t) - X(s)| \le Y|t - s|^{\varepsilon/k}\Big(\log\frac{c_1}{|t - s|}\Big)^{2/k}, \tag{3.14}$$

with a deterministic constant $c_1 > 0$, and a random variable Y such that $\mathbb{E}(Y^k) \le C'$.

Proof. Without loss of generality we will assume that $\mathbf{T}$ is the unit cube Q_1. We will use Theorem 3.1 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 3.1 is taken to be $X(t;\omega)$, for a fixed realization ω. This gives

$$\begin{aligned} B(\omega) &= \int_{R_1}\int_{R_1}\psi\Big(\frac{f(y) - f(x)}{p(|y - x|/\sqrt{d})}\Big)dxdy = \int_{Q_1}\int_{Q_1}\frac{|X(t;\omega) - X(s;\omega)|^k}{[p(|t - s|/\sqrt{d})]^k}dtds \\ &= C\int_{Q_1}\int_{Q_1}\frac{|X(t;\omega) - X(s;\omega)|^k}{|t - s|^{2d+\varepsilon}\log^2(c_1/|t - s|)}dtds. \end{aligned} \tag{3.15}$$

Taking the expectation, and using (3.13), we obtain

$$\mathbb{E}(B) \le C\int_{Q_1}\int_{Q_1}\frac{\mathbb{E}|X(t;\omega) - X(s;\omega)|^k}{|t - s|^{2d+\varepsilon}\log^2(c_1/|t - s||)}dtds \le C\int_{Q_1}\int_{Q_1}\frac{1}{|t - s|^d\log^2(c_1/|t - s|)}dtds. \tag{3.16}$$

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

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

Therefore,

$$\mathbb{E}(B) < \infty, \tag{3.17}$$

and $B(\omega)$ is finite almost surely. Going back to (3.2) we get

$$|X(t;\omega) - X(s;\omega)| \le 8B^{1/k}(\omega)\int_0^{|t-s|}\frac{1}{u^{2d/k}}p'(u)du. \tag{3.18}$$

It is now a calculus an exercise to check that (3.14) holds with $Y = CB^{1/k}$. The moment estimate on Y follows immediately from (3.17). □

3.3 Regularity of a Gaussian process

For Gaussian processes we have an improvement of the modulus of continuity estimate in Theorem 3.3.

Theorem 3.4 *Let $X(t)$, $t \in Q_1 \subset \mathbb{R}^d$ be a mean zero Gaussian process, and assume that*

$$\mathbb{E}|X(t) - X(s)|^2 \leq C|t-s|^m. \tag{3.19}$$

Then $X(t)$ has a version that is α-Hölder continuous for any $\alpha < m/2$.

Proof. For any $p \in \mathbb{N}$ there exists a constant C_p so that a Gaussian random variable Z satisfies

$$\mathbb{E}(Z^{2p}) \leq C_p[\mathbb{E}(Z^2)]^p. \tag{3.20}$$

It follows from (3.19) and (3.20) that for any integer $p \geq 1$ we have

$$\mathbb{E}|X(t) - X(s)|^{2p} \leq C|t-s|^{mp}. \tag{3.21}$$

Choosing p sufficiently large so that $mp > d$, we may apply Theorem 3.3 and use (3.14) with $k = 2p$ and $\varepsilon = mp - d$. We deduce that there exists a random variable $Y_p(\omega)$ such that $\mathbb{E}(Y_p^{2p}) < +\infty$ and

$$|X(t,\omega) - X(s,\omega)| \leq Y(\omega)|t-s|^{(mp-d)/(2p)}\Big(\log\frac{c_1}{|t-s|}\Big)^{1/p}. \tag{3.22}$$

Now, given any $\alpha < m/2$, we may take $p > 1$ sufficiently large so that $(mp-d)/(2p) > \alpha$. $\Box$

4 The additive heat equation

One setting where the Wiener integral is sufficient to produce interesting results about a stochastic PDE is the additive linear heat equation that can be formally written as

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

In this section, we will focus on that equation in one spatial dimension, so that $x \in \mathbb{R}$.

4.1 Heat equation forced by a measure

Before considering the random case, let us start with the heat equation forced by a deterministic σ-finite signed measure $\nu(dt,dx)$, in any spatial dimension $d \geq 1$:

$$\frac{\partial u}{\partial t} = \Delta u + \nu(dt,dx). \tag{4.2}$$

We will assume that the initial condition for (4.2) is $u(0,x) = 0$. Then, we expect the weak solution to (4.2) to be given by the Duhamel formula:

$$u(t,x) = \int_{[0,t]\times\mathbb{R}^d} G(t-s,x-y)\nu(ds,dy). \tag{4.3}$$

Here,

$$G(t,x) = \frac{1}{(4\pi t)^{d/2}} e^{-|x|^2/(4t)} \tag{4.4}$$

is the standard heat kernel. The validity of this guess is confirmed by the following exercise.

Exercise 4.1 Show that $u(t,x)$ given by (4.3) is a weak solution to (4.2) if the following Fubini-type condition holds for the measure ν: for any $\psi \in C_c^\infty((0,+\infty)\times\mathbb{R}^d)$ we have

$$\begin{aligned} &\int_{\mathbb{R}^+\times\mathbb{R}^d} \Big(\int_{[0,t]\times\mathbb{R}^d} G(t-s,x-y)\nu(ds,dy)\Big)\psi(t,x)dxdt \\ &= \int_{\mathbb{R}_+\times\mathbb{R}^d} \Big(\int_{[s,+\infty)\times\mathbb{R}^d} G(t-s,x-y)\psi(t,x)dtdx\Big)\nu(dsdy). \end{aligned} \tag{4.5}$$

Note that condition (4.5) holds for finite measures $\nu(dtdx)$ by the standard Fubini theorem because

$$\begin{aligned} &\int_{\mathbb{R}_+\times\mathbb{R}^d}\Big(\int_{[s,+\infty)\times\mathbb{R}^n} |G(t-s,x-y)\psi(t,x)|dtdx\Big)|\nu|(dsdy) \\ &\qquad\qquad \le |\nu|(\mathbb{R}_+\times\mathbb{R}^d)\int_0^\infty \|\psi(t,\cdot)\|_{L^\infty}dt. \end{aligned} \tag{4.6}$$

As far as the white noise forcing in (4.3) is concerned, going back to Proposition 2.5, and taking $f(t,x) = G(t,x)$ and $\mu(dtdx) = \psi(t,x)dtdx$ in the Fubini formula (2.45) gives, since $G(t,-x) = G(t,x)$:

$$\int\Big(\int G(t-s,x-y)\dot W(dsdy)\Big)\psi(t,x)dtdx = \int\Big(\int G(t-s,x-y)\psi(t,x)dxdt\Big)\dot W(dsdy), \text{ a.s.} \tag{4.7}$$

This is exactly (4.5). Therefore, a weak solution to

$$\frac{\partial u}{\partial t} = \Delta u + \dot W(dt,dx). \tag{4.8}$$

with the initial condition $u(0,x) = 0$ is given by

$$u(t,x) = \int_0^t \int_{\mathbb{R}^d} G(t-s,x-y)\dot W(ds,dy). \tag{4.9}$$

One very small caveat in this construction is that the test functions that we can use to verify that $u(t,x)$ given by (4.9) is a weak solution have to be deterministic and not random. Such solutions are known as mild solutions.

4.2 Temporal regularity of the solution in one dimension

4.2.1 A naive scaling argument

In this section we will look at the mild solution to

$$\frac{\partial Z}{\partial t} = \frac{\partial^2 Z}{\partial x^2} + \dot W(dt,dx), \quad t>0,\ x\in\mathbb{R}, \tag{4.10}$$

in dimension $d = 1$, with the initial condition $Z(0, x) = 0$. It is given by the stochastic convolution

$$Z(t,x) = \int_0^t \int_{\mathbb{R}} G(t-s, x-y)\dot{W}(ds, dy). \tag{4.11}$$

Observe that, as in (1.21) but now in a justified way, for each $t > 0$ and $x \in \mathbb{R}$ fixed, we may compute the variance of $Z(t, x)$ as

$$\begin{aligned}
\mathbb{E}\big[Z^2(t,x)\big] &= \int_0^t \int_{\mathbb{R}} |G(t-s,x-y)|^2 ds dy = \int_0^t \int_{\mathbb{R}} |G(s,y)|^2 ds dy \\
&= \frac{1}{4\pi}\int_0^t \int_{\mathbb{R}} e^{-|y|^2/(2s)} \frac{dyds}{s} = \frac{1}{\sqrt{8\pi}}\int_0^t \int_{\mathbb{R}} \frac{1}{\sqrt{4\pi s}} e^{-|y|^2/(4s)} dy \frac{ds}{\sqrt{s}} \\
&= \frac{1}{\sqrt{2\pi}}\int_0^t \frac{ds}{2\sqrt{s}} = \frac{\sqrt{t}}{\sqrt{2\pi}}.
\end{aligned} \tag{4.12}$$

Therefore, for each $t > 0$ and $x \in \mathbb{R}$ fixed, $Z(t, x)$ is a mean-zero Gaussian random variable with variance

$$\mathbb{E}\big[Z^2(t,x)\big] = \frac{\sqrt{t}}{\sqrt{2\pi}}. \tag{4.13}$$

The goal of this section is not to look at $Z(t, x)$ pointwise but to consider it separately as a function of t, with $x \in \mathbb{R}$ fixed, and also as a function of x, with $t > 0$ fixed.

Our goal will be to show the following two decompositions. First, for each $x \in \mathbb{R}^d$ fixed, we will show that we can decompose $Z(t, x)$ as a sum

$$Z(t,x) = X(t) + R(t), \tag{4.14}$$

where $X(t)$ is a fractional Brownian motion with the Hurst exponent $H = 1/4$, and $R(t)$ is a mean-zero Gaussian process that is almost surely infinitely differentiable, with $R(0) = 0$ a.s.

Next, we will show that for each $t \geq 0$ fixed, we can decompose $Z(t, x)$ as

$$Z(t,x) = Z(t,0) + B(x) + Q(x), \tag{4.15}$$

where $B(x)$ is a two-sided Brownian motion in x, and $Q(x)$ is a mean-zero Gaussian process that is almost surely infinitely differentiable, with $B(0) = 0$ and $Q(0) = 0$ a.s.

In other words, the time trace of $Z(t, x)$ for $x \in \mathbb{R}$ fixed is a smooth perturbation of a fractional Brownian motion $X(t)$, and the spatial trace of $Z(t, x)$ for $t > 0$ fixed is a smooth perturbation of a standard two-sided Brownian motion $B(x)$. We will also see that the variance of the increments of the correctors $R(t)$ and $Q(x)$ decays as $t \to +\infty$.

Let us first explain why this regularity of $Z(t, x)$ in the t and x variables can be expected. In one spatial dimension, an approximation to white noise $\dot{W}(dxdt)$ is given by a random field of the form

$$F_\varepsilon(t,x) = \frac{1}{\varepsilon^{3/2}} F\Big(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\Big),\ t \in \mathbb{R},\ x \in \mathbb{R}, \tag{4.16}$$

with a mean-zero Gaussian random field $F(t, x)$ with the covariance

$$\mathbb{E}(F(t,x)F(t',x')) = R(t-t', x-x'), \tag{4.17}$$

with a rapidly decaying correlation function $R(t, x) \geq 0$ such that $\|R\|_{L^1} = 1$. For example, we can take $R(t, x)$ to be a Gaussian in t and x.

Exercise 4.2 Show that a function $R(t, x)$ is positive-definite if and only if its Fourier transform $\hat{R}(\omega, \xi) \geq 0$ for all $\omega \in \mathbb{R}$ and $\xi\mathbb{R}$.

This convergence can be seen by showing that the Gaussian random variable

$$\int F_\varepsilon(t, x)\phi(t, x)dxdt \tag{4.18}$$

converges in law to the Gaussian random variable

$$\int \phi(t, x)\dot{W}(dtdx), \tag{4.19}$$

for any test function $\phi \in C_c^\infty(\mathbb{R} \times \mathbb{R})$.

Let us then approximate $Z(t, x)$ by the solution $u_\varepsilon(t, x)$ to the heat equation with the forcing $F_\varepsilon(t, x)$:

$$\frac{\partial u_\varepsilon}{\partial t} = \frac{\partial^2 u_\varepsilon}{\partial x^2} + F_\varepsilon(t, x). \tag{4.20}$$

Then, $u_\varepsilon(t, x)$ can be written as

$$u_\varepsilon(t, x) = \varepsilon^{1/2} u\Big(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\Big). \tag{4.21}$$

Here, $u(t, x)$ is the solution to the unscaled problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + F(t, x). \tag{4.22}$$

If the field $F(t, x)$ is sufficiently regular then the function $u(t, x)$ is smooth and, of course, has no dependence on $\varepsilon \in (0, 1)$. Looking at the scaling for $u_\varepsilon(t, x)$ in (4.21), we may, therefore, expect the function $u_\varepsilon(t, x)$ given by (4.21) to obey the uniform bounds

$$|\partial_t^{1/4} u_\varepsilon(t, x)| \leq C, \quad |\partial_x^{1/2} u_\varepsilon(t, x)| \leq C, \tag{4.23}$$

with a constant $C > 0$ that does not depend on $\varepsilon \in (0, 1)$. This is not quite correct because of certain logarithmic factors but nevertheless, this argument indicates that we should expect close to 1/4-Hölder regularity for $Z(t, x)$ in the t variable, and close to 1/2-Hölder regularity for $Z(t, x)$ in the x variable. This is what is behind the decompositions (4.14) and (4.15).

Exercise 4.3 Consider the solution to the additive heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \partial_x F(t, x), \tag{4.24}$$

forced by the derivative of a space-time stationary random field $F(t, x)$ as above. For which $m \in \mathbb{R}$ do you expect that the rescaled solution

$$u_\varepsilon(t, x) = \varepsilon^m u\Big(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\Big), \tag{4.25}$$

would have a reasonable weak limit as $t \to +\infty$? What kind of behavior do you expect for the solution to (4.24)? One way to look at this is to consider this equation with the initial condition $u_T(-T, x) = 0$ for some $T > 1$ and then let $T \to +\infty$. Compare what happens when this is done for the solutions to (4.22) and (4.24).

4.2.2 The fractional Brownian motion

Let us first make a brief detour to recall that a fractional Brownian motion (fBm) with a Hurst exponent $H \in (0,1)$ is a mean-zero Gaussian process $X(t)$, $t \geq 0$, such that

$$X(0) = 0, \quad \mathbb{E}(|X(t_1) - X(t_2)|^2) = |t_1 - t_2|^{2H}, \quad \text{for all } t_1, t_2 \geq 0. \tag{4.26}$$

In other words, $X(t)$ is a mean-zero Gaussian process on $[0, +\infty)$ with the covariance

$$\begin{aligned} \mathbb{E}(X(t_1)X(t_2)) &= \frac{1}{2}\big(\mathbb{E}(X^2(t_1)) + \mathbb{E}(X^2(t_2)) - \mathbb{E}(|X(t_2) - X(t_1)|^2)\big) \\ &= \frac{1}{2}\big(t_1^{2H} + t_2^{2H} - |t_1 - t_2|^{2H}\big). \end{aligned} \tag{4.27}$$

Instead of checking directly that this covariance function is positive-definite, let us define

$$f_H(t,s) = (t-s)_+^{H-1/2} - (-s)_+^{H-1/2} = \begin{cases} 0, & s > t, \\ (t-s)^{H-1/2}, & 0 \leq s \leq t, \\ (t-s)^{H-1/2} - (-s)^{H-1/2}, & s < 0. \end{cases} \tag{4.28}$$

Then, for $0 < t_1 < t_2$ we have

$$f_H(t_2, s) - f_H(t_1, s) = (t_2 - s)_+^{H-1/2} - (t_1 - s)_+^{H-1/2}, \quad \text{for all } s \in \mathbb{R}. \tag{4.29}$$

A key observation is that for $0 < t_1 < t_2$ we have the following representation for the second moment of the increment of a fractional Brownian motion:

$$(t_2 - t_1)^{2H} = \frac{1}{C_H} \int_{\mathbb{R}} |f_H(t_1, s) - f_H(t_2, s)|^2 ds, \tag{4.30}$$

with

$$C_H = \int_0^1 s^{2H-1} ds + \int_0^\infty \big[(1+s)^{H-1/2} - s^{H-1/2}\big]^2 ds. \tag{4.31}$$

Note that C_H is finite precisely when both $2H - 1 > -1$, so that the first integral in the right side above is finite near $s = 1$, and $2(H - 3/2) < -1$, so that the second integral is finite as $s \to +\infty$. That is, we need to have $H \in (0,1)$. The representation (4.30) is verified as follows:

$$\begin{aligned} \int_{\mathbb{R}} |f_H(t_1, s) - f_H(t_2, s)|^2 ds &= \int_{-\infty}^{t_1} \big[(t_2 - s)^{H-1/2} - (t_1 - s)^{H-1/2}\big]^2 ds + \int_{t_1}^{t_2} (t_2 - s)^{2H-1} ds \\ = \int_0^{t_2 - t_1} (t_2 - t_1 - s)^{H-1/2} ds &+ \int_0^\infty \big[(t_2 - t_1 + s)^{H-1/2} - s^{H-1/2}\big]^2 ds = C_H (t_2 - t_1)^{2H}, \end{aligned} \tag{4.32}$$

Next, for each $t > 0$ consider the Wiener integral

$$X(t) = C_H^{-1/2} \int_{\mathbb{R}} f_H(t,s) \dot{W}(ds). \tag{4.33}$$

This integral is well defined because for each $t > 0$ fixed, we have

$$\int_{\mathbb{R}} |f_H(t,s)|^2 ds = \int_0^t |t - s|^{2H-1} ds + \int_0^\infty \big[(t+s)^{H-1/2} - s^{H-1/2}\big]^2 ds < +\infty, \tag{4.34}$$

for all $H \in (0,1)$ for reasons discussed below (4.31).

Then, we can compute the variance of the increments of $X(t)$ using the Ito isometry and (4.30):

$$\mathbb{E}(|X(t_1) - X(t_2)|^2) = C_H^{-1} \int_{\mathbb{R}} [f_H(t_1, s) - f_H(t_2, s)]^2 ds = |t_2 - t_1|^{2H}. \tag{4.35}$$

Therefore, $X(t)$ given by (4.33) does have the correct mean and covariance and is a fractional Brownian motion.

One may be tempted to define the fractional Brownian motion via the Wiener integral, as in (4.33), but with the function $f_H(t,s)$ replaced by $(t-s)^{H-1/2}$, as the contribution of the term $(-s)^{H-1/2}$ to the Wiener integral in (4.33) seems to be "t-independent" and thus irrelevant for the process $X(t)$. This would formally lead to the same covariance as in (4.35). However, the function $(t-s)^{H-1/2}$ is not square integrable as $s \to +\infty$, and the additional term introduced in the definition (4.28) of $f_H(t,s)$ is used precisely to ensure the L^2-integrability in (4.34). This will be a recurring theme in various definitions of Gaussian fields in this section.

The Kolmogorov Theorem 3.4 for Gaussian processes implies that a fractional Brownian motion with a Hurst exponent $H \in (0,1)$ is α-Hölder continuous for any $\alpha < H$. This can be improved to the law of iterated logarithm regularity property of an fBm.

Proposition 4.4 *Let $X(t)$ be an fBm with a Hurst exponent $H \in (0,1)$, then for every $t \geq 0$ we have*

$$\lim_{\varepsilon \downarrow 0} \frac{X(t+\varepsilon) - X(t)}{\varepsilon^H \sqrt{2 \log \log(\varepsilon^{-1})}} = 1, \quad a.s. \tag{4.36}$$

We will not prove this result here. However, we do mention that it implies that fractional Brownian motion is almost surely not Hölder continuous with any exponent $\alpha \geq H$.

4.3 The temporal evolution of the solution at a fixed spatial point

In this section, we characterize the process $Z(t,x)$ for $x \in \mathbb{R}$ fixed.

4.3.1 The temporal corrector at a fixed spatial point

A key observation is that $Z(t,x)$ at a fixed $x \in \mathbb{R}$ can be decomposed as a sum of a fractional Brownian motion $X(t)$ with the Hurst exponent $H = 1/4$ and a corrector $R(t)$ that is defined as follows. Let $\dot{W}_1(dx)$ be a spatial white noise, independent of the space-time white noise $\dot{W}(dtdx)$ and set

$$R(t) = \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} \frac{1}{z} \Big(1 - e^{-tz^2}\Big) \dot{W}_1(dz). \tag{4.37}$$

This is a well-defined $L^2(\Omega)$ random field, for each $t \geq 0$ fixed, because

$$\mathbb{E}(R^2(t)) = \frac{1}{4\pi} \int_{\mathbb{R}} \frac{1}{z^2} \Big(1 - e^{-tz^2}\Big)^2 dz < +\infty. \tag{4.38}$$

The role of the first term in the parentheses in (4.37) is precisely to keep the L^2-norm in (4.38) from blowing up near $z = 0$. It also guarantees that $R(0) = 0$ almost surely.

The main observation is that the corrector process is actually smooth in t, as shown by the following.

Proposition 4.5 *The random process $R(t)$ is a mean-zero Gaussian process with a version that is continuous on $[0, +\infty)$ and infinitely differentiable on $(0, +\infty)$ such that $R(0) = 0$. Its variance is*

$$\mathbb{E}(R^2(t)) = \frac{2 - \sqrt{2}}{\sqrt{4\pi}} t^{1/2}. \tag{4.39}$$

Proof. The basic reason for this regularity is simply that $R(t)$ is a Wiener integral that depends on t as a parameter, and the integrand is a deterministic function of t that has very strong decay properties in the variable of integration z when $t > 0$. This will allow us essentially to differentiate under the integral sign in (4.37).

First, we verify that $R(t)$ is continuous, using the Kolmogorov criterion for the continuity of Gaussian processes in Theorem 3.4. Observe that for $t \geq s \geq 0$ we have

$$\begin{aligned}\mathbb{E}(|R(t) - R(s)|^2) &= \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{1}{z^2} \Big(e^{-sz^2} - e^{-tz^2}\Big)^2 dz \leq \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{1}{z^2} \Big(1 - e^{-(t-s)z^2}\Big)^2 dz \\ &= \frac{(t-s)^{1/2}}{4\pi} \int_{-\infty}^{\infty} \frac{1}{z^2} \Big(1 - e^{-z^2}\Big)^2 dz.\end{aligned} \tag{4.40}$$

It follows from the aforementioned theorem that $R(t)$ is almost surely α-Hölder continuous with any exponent $\alpha \in (0, 1/4)$.

We will only show that $R(t)$ is differentiable as higher order differentiability is proved in an identical fashion. Consider the Wiener integral

$$D(t) = \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} \frac{1}{z} \partial_t \Big(1 - e^{-tz^2}\Big) \dot{W}_1(dz) = \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} z e^{-tz^2} \dot{W}_1(dz). \tag{4.41}$$

We will show that

$$D(t) = \partial_t R(t) \qquad \text{a.s.} \tag{4.42}$$

First, we estimate the second moment of the increments of $D(t)$ for $t \geq s > 0$ as

$$\begin{aligned}\mathbb{E}(|D(t) - D(s)|^2) &= \frac{1}{4\pi} \int_{-\infty}^{\infty} z^2 \Big(e^{-sz^2} - e^{-tz^2}\Big)^2 dz = \frac{1}{4\pi} \int_{-\infty}^{\infty} z^2 e^{-2sz^2} \Big(1 - e^{-(t-s)z^2}\Big)^2 dz \\ &\leq \frac{(t-s)^2}{4\pi} \int_{-\infty}^{\infty} z^4 e^{-2sz^2} dz.\end{aligned} \tag{4.43}$$

It follows that $D(t)$ is continuous on $(0, +\infty)$ a.s. Next, to prove (4.42), let $\phi(t) \in C_c^\infty(0, +\infty)$, then by the stochastic Fubini theorem (Proposition 2.5) we have

$$
\begin{aligned}
\int_0^\infty D(t)\phi(t)dt &= \frac{1}{\sqrt{4\pi}}\int_0^\infty \phi(t)\Big(\int_{\mathbb{R}} \frac{1}{z}\partial_t\Big(1-e^{-tz^2}\Big)\dot W_1(dz)\Big)dt \\
&= \frac{1}{\sqrt{4\pi}}\int_{\mathbb{R}} \frac{1}{z}\Big(\int_0^{+\infty} \phi(t)\partial_t\Big(1-e^{-tz^2}\Big)dt\Big)\dot W_1(dz) \\
&= -\frac{1}{\sqrt{4\pi}}\int_{\mathbb{R}} \frac{1}{z}\Big(\int_0^{+\infty} \Big(1-e^{-tz^2}\Big)\partial_t\phi(t)dt\Big)\dot W_1(dz) \\
&= -\frac{1}{\sqrt{4\pi}}\int_0^{+\infty} (\partial_t\phi(t))\Big(\int_{\mathbb{R}} \frac{1}{z}\Big(1-e^{-tz^2}\Big)\dot W_1(dz)\Big)dt = -\int_0^{+\infty} R(t)\partial_t\phi(t)dt.
\end{aligned}
\tag{4.44}
$$

Therefore, $D(t)$ is the weak derivative of $R(t)$. Since $D(t)$ is actually continuous, it follows that it is the ordinary derivative of $R(t)$.

It remains to compute the variance of $R(t)$. As in (4.40), we have

$$
\mathbb{E}(|R(t)|^2) = \frac{1}{4\pi}\int_{-\infty}^\infty \frac{1}{z^2}\Big(1-e^{-tz^2}\Big)^2 dz = \frac{t^{1/2}}{2\pi}\int_0^\infty \frac{1}{z^2}\Big(1-e^{-z^2}\Big)^2 dz. \tag{4.45}
$$

We may now use the change of variable $z=\sqrt{s}$ to re-write the integral in the right side as

$$
\int_0^\infty \Big(1-e^{-z^2}\Big)^2\frac{dz}{z^2} = \frac{1}{2}\int_0^\infty \frac{(1-e^{-s})^2}{s^{3/2}}ds = \frac{1}{2}\int_0^\infty \frac{1-e^{-s}}{s^{3/2}}ds - \frac{1}{2}\int_0^\infty \frac{e^{-s}-e^{-2s}}{s^{3/2}}ds. \tag{4.46}
$$

Observe that

$$
\begin{aligned}
\int_0^\infty \frac{1-e^{-s}}{s^{3/2}}ds &= \int_0^\infty \frac{1}{\sqrt{s}}\int_0^1 e^{-rs}dr = \int_0^1\int_0^\infty \frac{1}{\sqrt{s}}e^{-rs}dsdr \\
&= \int_0^1 \frac{1}{\sqrt{r}}\int_0^\infty \frac{1}{\sqrt{s}}e^{-s}dsdr = 2\Gamma(1/2),
\end{aligned}
\tag{4.47}
$$

and similarly we can compute

$$
\begin{aligned}
\int_0^\infty \frac{e^{-s}-e^{-2s}}{s^{3/2}}ds &= \int_0^\infty \frac{1}{\sqrt{s}}\int_1^2 e^{-rs}dr = \int_1^2\int_0^\infty \frac{1}{\sqrt{s}}e^{-rs}dsdr \\
&= \int_1^2 \frac{1}{\sqrt{r}}\int_0^\infty \frac{1}{\sqrt{s}}e^{-s}dsdr = 2(\sqrt{2}-1)\Gamma(1/2).
\end{aligned}
\tag{4.48}
$$

As $\Gamma(1/2)=\sqrt{\pi}$, we obtain

$$
\int_0^\infty \Big(1-e^{-z^2}\Big)^2\frac{dz}{z^2} = \frac{1}{2}\sqrt{\pi}[2-2(\sqrt{2}-1)] = (2-\sqrt{2})\sqrt{\pi}. \tag{4.49}
$$

Using this identity in (4.45), we conclude that

$$
\mathbb{E}(R^2(t)) = \frac{(2-\sqrt{2})\sqrt{\pi}t^{1/2}}{2\pi} = \frac{2-\sqrt{2}}{\sqrt{4\pi}}t^{1/2}, \tag{4.50}
$$

which is (4.39). □

The main reason for the introduction of the corrector $R(t)$ is the specific form of the variance of its increments.

Proposition 4.6 *Given any $t > 0$ and $\varepsilon > 0$ variance of the increment $R(t+\varepsilon) - R(t)$ is*

$$\mathbb{E}\big[(R(t+\varepsilon) - R(t))^2\big] = \int_t^\infty \int_{\mathbb{R}} |G(s+\varepsilon, y) - G(s,y)|^2 dyds. \tag{4.51}$$

Proof. Let us compute

$$\begin{aligned}\mathbb{E}\big[(R(t+\varepsilon) - R(t))^2\big] &= \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{1}{z^2}\Big(e^{-tz^2} - e^{-(t+\varepsilon)z^2}\Big)^2 dz = \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{1}{z^2} e^{-2tz^2}\Big(1 - e^{-\varepsilon z^2}\Big)^2 dz \\ &= \frac{1}{2\pi}\int_t^\infty \int_{-\infty}^{\infty} e^{-2sz^2}\Big(1 - e^{-\varepsilon z^2}\Big)^2 dzds = \frac{1}{2\pi}\int_t^\infty \int_{-\infty}^{\infty} \Big(e^{-sz^2} - e^{-(s+\varepsilon)z^2}\Big)^2 dzds.\end{aligned} \tag{4.52}$$

The integral in the very right side above can be computed by noting that the Fourier transform of the heat kernel is

$$\hat{G}(t,\xi) = \int_{\mathbb{R}} G(t,y) e^{-2\pi i y\xi} dy = e^{-4\pi^2|\xi|^2 t}, \tag{4.53}$$

and using the Plancherel identity to write

$$\begin{aligned}&\frac{1}{2\pi}\int_{\mathbb{R}} \Big|e^{-(s+\varepsilon)z^2} - e^{-sz^2}\Big|^2 dz = \int_{\mathbb{R}} \Big|e^{-4\pi^2|\xi|^2(s+\varepsilon)} - e^{-4\pi^2|\xi|^2 s}\Big|^2 d\xi \\ &= \int_{\mathbb{R}} |\hat{G}(s+\varepsilon,\xi) - \hat{G}(s,\xi)|^2 d\xi = \int_{\mathbb{R}} [G(s+\varepsilon,y) - G(s,y)]^2 dy.\end{aligned} \tag{4.54}$$

Using this identity in (4.52), we arrive at

$$\mathbb{E}\big[(R(t+\varepsilon) - R(t))^2\big] = \int_t^\infty \int_{\mathbb{R}} |G(s+\varepsilon, y) - G(s,y)|^2 dyds, \tag{4.55}$$

which is (4.51). □

4.3.2 The decomposition into a fractional Brownian motion and the corrector

We now prove the following decomposition originally proved in a paper by Lei and Nualart in 2009.

Theorem 4.7 *Let $R(t)$ be defined by (4.37). For every $x \in \mathbb{R}$ fixed there exists a fractional Brownian motion $X(t)$, $t \geq 0$, with the Hurst exponent $H = 1/4$ such that*

$$Z(t,x) = \pi^{-1/4} X(t) + R(t),\ t > 0. \tag{4.56}$$

In particular, the non-differentiable component of the time trace of $Z(t,x)$ is a fractional Brownian motion.

We also note that at the level of variances the constant $\pi^{-1/4}$ is correct because $Z(t,x)$ and $R(t)$ are independent and recalling (4.13) and (4.39) we have

$$\mathbb{E}\big[Z^2(t,x)\big] + \mathbb{E}\big[R^2(t)\big] = \frac{1}{\sqrt{2\pi}} t^{1/2} + \frac{2-\sqrt{2}}{\sqrt{4\pi}} t^{1/2} = \frac{\sqrt{t}}{\sqrt{\pi}}. \tag{4.57}$$

This agrees with the idea that $Z(t,x)-R(t)$ could be the corresponding multiple of a fractional Brownian motion with the Hurst exponent $H=1/4$.

Proof. Let us take $t\geq 0$, $\varepsilon>0$ and $x\in\mathbb{R}$. We know that the corresponding increment

$$Z(t+\varepsilon,x)-Z(t,x)$$

is a mean-zero Gaussian. To prove the claim of Theorem 4.7 we need to understand the variance of such increments. More precisely, as $Z(t,x)$ and $R(t)$ are independent mean-zero Gaussian fields with

$$Z(0,x)=R(0)=0,\quad \text{a.s.} \tag{4.58}$$

it suffices to show that for any $t>0$, $\varepsilon>0$ and $x\in\mathbb{R}$ we have

$$\mathbb{E}(|Z(t+\varepsilon,x)-Z(t,x)|^2)+\mathbb{E}(|R(t+\varepsilon)-R(t)|^2)=\frac{\sqrt{\varepsilon}}{\sqrt{\pi}}, \tag{4.59}$$

as this, together with (4.58), would imply that

$$\pi^{1/4}(Z(t,x)-R(t)) \tag{4.60}$$

is a fractional Brownian motion with the Hurst exponent $H=1/4$. According to Proposition 4.6, identity (4.59) is equivalent to

$$\mathbb{E}(|Z(t+\varepsilon,x)-Z(t,x)|^2)=\frac{\sqrt{\varepsilon}}{\sqrt{\pi}}-\int_t^\infty\int_{\mathbb{R}}[G(s+\varepsilon,y)-G(s,y)]^2dsdy. \tag{4.61}$$

To prove (4.61), let us split the increment of $Z(t,x)$ as

$$Z(t+\varepsilon,x)-Z(t,x)=J_1(t,x)+J_2(t,x), \tag{4.62}$$

with

$$\begin{aligned}J_1(t,t+\varepsilon,x)&=\int_0^t[G(t+\varepsilon-s,x-y)-G(t-s,x-y)]\dot{W}(dsdy),\\ J_2(t,t+\varepsilon,x)&=\int_t^{t+\varepsilon}G(t+\varepsilon-s,x-y)\dot{W}(dsdy).\end{aligned} \tag{4.63}$$

Then, $J_1(t,t+\varepsilon,x)$ and $J_2(t,t+\varepsilon,x)$ are independent mean-zero Gaussians, and the variance of an increment of $Z(t,x)$ has the form

$$\mathbb{E}(|Z(t+\varepsilon,x)-Z(t,x)|^2)=\mathbb{E}(J_1^2(t,t+\varepsilon,x))+\mathbb{E}(J_2^2(t,t+\varepsilon,x)). \tag{4.64}$$

By the Ito isometry and the semi-group property of the heat kernel, we have

$$\begin{aligned}\mathbb{E}(J_2^2(t,t+\varepsilon,x))&=\int_t^{t+\varepsilon}\int_{\mathbb{R}}G^2(t+\varepsilon-s,x-y)dyds=\int_0^\varepsilon\int_{\mathbb{R}}G^2(s,y)dyds\\ &=\int_0^\varepsilon\int_{\mathbb{R}}G(s,0-y)G(s,y)dyds=\int_0^\varepsilon G(2s,0)ds=\int_0^\varepsilon\frac{ds}{\sqrt{8\pi s}}=\sqrt{\frac{\varepsilon}{2\pi}}.\end{aligned} \tag{4.65}$$

The computation of the variance of $J_1(t,t+\varepsilon,x)$ is a bit longer. By the Ito isometry we have

$$\begin{aligned}\mathbb{E}\big(J_1^2(t,t+\varepsilon,x)\big) &= \int_0^t\int_{\mathbb{R}}[G(t+\varepsilon-s,x-y)-G(t-s,x-y)]^2dsdy\\ &= \int_0^t\int_{\mathbb{R}}[G(s+\varepsilon,y)-G(s,y)]^2dsdy = \int_0^\infty\int_{\mathbb{R}}[G(s+\varepsilon,y)-G(s,y)]^2dsdy\\ &-\int_t^\infty\int_{\mathbb{R}}[G(s+\varepsilon,y)-G(s,y)]^2dsdy.\end{aligned} \tag{4.66}$$

The second integral in the very right side above is precisely the integral that appears in the right side of (4.61). The first integral in the very right side above can be computed by noting that the Fourier transform of the heat kernel is

$$\hat{G}(t,\xi) = \int_{\mathbb{R}} G(t,y)e^{-2\pi iy\xi}dy = e^{-4\pi^2|\xi|^2t}, \tag{4.67}$$

and using the Plancherel identity to write

$$\begin{aligned}\int_{\mathbb{R}}[G(s+\varepsilon,y)-G(s,y)]^2dy &= \int_{\mathbb{R}}|\hat{G}(s+\varepsilon,\xi)-\hat{G}(s,\xi)|^2d\xi\\ &= \int_{\mathbb{R}}\Big|e^{-4\pi^2|\xi|^2(s+\varepsilon)}-e^{-4\pi^2|\xi|^2s}\Big|^2d\xi = \int_{\mathbb{R}}e^{-8\pi^2|\xi|^2s}\Big|e^{-4\pi^2\varepsilon|\xi|^2}-1\Big|^2d\xi.\end{aligned} \tag{4.68}$$

We deduce that

$$\begin{aligned}&\int_0^\infty\int_{\mathbb{R}}[G(s+\varepsilon,y)-G(s,y)]^2dsdy = \int_{\mathbb{R}}\Big|e^{-4\pi^2\varepsilon|\xi|^2}-1\Big|^2\Big(\int_0^\infty e^{-8\pi^2|\xi|^2s}ds\Big)d\xi\\ &= \int_{\mathbb{R}}\frac{1}{8\pi^2|\xi|^2}\Big|e^{-4\pi^2\varepsilon|\xi|^2}-1\Big|^2d\xi = \frac{\sqrt{4\pi^2\varepsilon}}{8\pi^2}\int_{\mathbb{R}}\frac{1}{|z|^2}\Big|e^{-|z|^2}-1\Big|^2dz\\ &= \frac{\sqrt{\varepsilon}}{2\pi}\int_0^\infty\frac{1}{z^2}\Big|e^{-z^2/2}-1\Big|^2dz.\end{aligned} \tag{4.69}$$

We used above the change of variable $z=\sqrt{4\pi^2\varepsilon}\xi$. We may now recall (4.49):

$$\int_0^\infty\Big(1-e^{-z^2}\Big)^2\frac{dz}{z^2} = (2-\sqrt{2})\sqrt{\pi}. \tag{4.70}$$

We obtain

$$\int_0^\infty\int_{\mathbb{R}}[G(s+\varepsilon,y)-G(s,y)]^2dsdy = \frac{\sqrt{\varepsilon}}{2\pi}\sqrt{\pi}[2-\sqrt{2}]. \tag{4.71}$$

Going back to (4.66), we obtain

$$\mathbb{E}\big(J_1^2(t,t+\varepsilon,x)\big) = \frac{\sqrt{\varepsilon}(2-\sqrt{2})}{\sqrt{4\pi}} - \int_t^\infty\int_{\mathbb{R}}[G(s+\varepsilon,y)-G(s,y)]^2dsdy. \tag{4.72}$$

Using (4.65) and (4.72) in (4.64) gives

$$\mathbb{E}(|Z(t+\varepsilon,x)-Z(t,x)|^2) = \frac{\sqrt{\varepsilon}}{\sqrt{\pi}} - \int_t^\infty\int_{\mathbb{R}}[G(s+\varepsilon,y)-G(s,y)]^2dsdy, \tag{4.73}$$

which is (4.61). □

4.4 The spatial variation of the solution at a fixed time

We now view the process

$$Z(t,x)=\int_0^t\int_{\mathbb{R}} G(t-s,x-y)\dot W(ds,dy). \tag{4.74}$$

as a function of $x\in\mathbb{R}$ with some $t>0$ fixed.

4.4.1 The spatial corrector at a fixed time

We first construct a spatial corrector at a fixed time $t>0$. Let us take a space-time white-noise $\dot W_1(dtdx)$ that is independent of the space-time white noise $\dot W(dtdx)$ that appears in the definition (4.74) of $Z(t,x)$, and define the process

$$S(t,x)=Z(t,0)+\int_{[t,+\infty)\times\mathbb{R}}[G(s,y-x)-G(s,y)]\dot W_1(dsdy). \tag{4.75}$$

Here, we view $t>0$ as simply a parameter and consider $S(t,x)$ as a Gaussian random process in x.

First, note that, due to the term $G(s,y)$ inside the parentheses in (4.75), we have

$$S(t,0)=Z(t,0) \text{ a.s., for any } t\ge 0 \text{ fixed.} \tag{4.76}$$

This term also ensures the square integrability of the kernel that appears in (4.75) that is needed to make sure that the Wiener integral in the definition of $S(t,x)$ is well-defined. Let us verify that property. The Fourier transform of the function $G(s,\cdot-x)$ is

$$\int e^{-2\pi i\xi y}G(s,y-x)dy=\int e^{-2\pi i\xi(x+y)}G(s,y)dy=e^{-2\pi i\xi x}\hat G(s,\xi)=e^{-2\pi i\xi x-4\pi^2|\xi|^2 s}. \tag{4.77}$$

The Plancherel identity implies that

$$\begin{aligned}&\int_t^{+\infty}\int_{\mathbb{R}}|G(s,y-x)-G(s,y)|^2dyds=\int_t^{\infty}\int_{\mathbb{R}}|\hat G(s,\xi)|^2|1-e^{-2\pi i\xi x}|^2d\xi ds\\ &=\int_t^{\infty}\int_{\mathbb{R}}e^{-8\pi^2|\xi|^2 s}\Big((1-\cos(2\pi\xi x))^2+\sin^2(2\pi\xi x)\Big)d\xi ds\\ &=2\int_t^{\infty}\int_{\mathbb{R}}e^{-8\pi^2|\xi|^2 s}\big(1-\cos(2\pi\xi x)\big)d\xi ds=\frac{1}{4\pi^2}\int_{\mathbb{R}}\frac{1}{|\xi|^2}e^{-8\pi^2|\xi|^2 t}\big(1-\cos(2\pi\xi x)\big)d\xi<+\infty,\end{aligned} \tag{4.78}$$

for all $t\ge 0$ and $x\in\mathbb{R}$. As a consequence, we have

$$\mathbb{E}(S^2(t,x))=\mathbb{E}(Z^2(t,0))+\int_t^{+\infty}\int_{\mathbb{R}}|G(s,y-x)-G(s,y)|^2dyds \tag{4.79}$$

We have the following analog of Proposition 4.5.

Proposition 4.8 *For each $t>0$ fixed, the random process $S(t,x)$ is a mean-zero Gaussian process in x with a version that is continuous on $[0,+\infty)$ and infinitely differentiable on $(0,+\infty)$.*

Proof, Let us first check the continuity of $S(t,x)$ in x. Given any $x \in \mathbb{R}$ and $\varepsilon > 0$, we have, by the Ito isometry formula and (4.78):

$$\begin{aligned}\mathbb{E}(|S(t,x+\varepsilon) - S(t,x)|^2) &= \int_t^\infty \int_{\mathbb{R}} |G(s,y-x-\varepsilon) - G(s,y-x)|^2 dsdx \\ = \int_t^\infty \int_{\mathbb{R}} |G(s,y+\varepsilon) - G(s,y)|^2 dyds &= \frac{1}{4\pi^2}\int_{\mathbb{R}} \frac{1}{|\xi|^2} e^{-8\pi^2|\xi|^2 t}\big(1-\cos(2\pi\xi\varepsilon)\big)d\xi.\end{aligned} \tag{4.80}$$

Making a change of variable $\xi = \zeta/(2\pi\varepsilon)$ gives

$$\begin{aligned}\mathbb{E}(|S(t,x+\varepsilon) - S(t,x)|^2) &= \frac{1}{4\pi^2}\int_{\mathbb{R}} \frac{1}{|\xi|^2} e^{-8\pi^2|\xi|^2 t}\big(1-\cos(2\pi\xi\varepsilon)\big)d\xi \\ = \frac{\varepsilon}{2\pi}\int_{\mathbb{R}} \frac{1}{|\zeta|^2} e^{-2|\zeta|^2 t/\varepsilon^2}\big(1-\cos(\zeta)\big)d\zeta &\le \frac{\varepsilon}{2\pi}\int_{\mathbb{R}} \frac{1}{|\zeta|^2}\big(1-\cos(\zeta)\big)d\zeta = \frac{\varepsilon}{2}.\end{aligned} \tag{4.81}$$

We used above the remarkable fact that

$$\int_{-\infty}^\infty (1-\cos\zeta)\frac{d\zeta}{\zeta^2} = 2\int_0^\infty (1-\cos\zeta)\frac{d\zeta}{\zeta^2} = 2\int_0^\infty \frac{\sin\zeta}{\zeta}d\zeta = \pi, \tag{4.82}$$

as can be easily computed by the method of residues. Theorem 3.4 and (4.81) imply that the random process $S(x)$ is almost surely α-Hölder continuous with any exponent $\alpha \in (0,1/2)$.

Differentiability can be verified exactly as for the process $R(t)$ in Proposition 4.5, by showing that

$$\partial_x S(t,x) = \int_{[t,+\infty)\times\mathbb{R}} \partial_x[G(s,y-x)]\dot{W}_1(dsdy) \;\; \text{a.s.} \tag{4.83}$$

The details are essentially identical to that proof and we omit the details. □

4.4.2 The decomposition into a spatial Brownian motion and a corrector at a fixed time

The next theorem characterizes the process $Z(t,x)$ for $t \ge 0$ fixed. It comes from a paper by Foondun, Khoshnevisan and Mahboubi in 2015.

Theorem 4.9 *Fix $t \ge 0$, and let $S(t,x)$ be defined by (4.75). There exists a standard two-sided Brownian motion $B(x)$, $x \in \mathbb{R}$, such that*

$$Z(t,x) = \frac{1}{\sqrt{2}}B(x) + S(t,x), \;\; t > 0. \tag{4.84}$$

Proof. Note that

$$B(x) = \sqrt{2}(Z(t,x) - S(t,x)) = \sqrt{2}(Z(t,x) - Z(t,0)) + \sqrt{2}(Z(t,0) - S(t,x)). \tag{4.85}$$

is a mean-zero Gaussian field in x such that $B(0) = 0$ a.s. The two terms in the right side of (4.85) are independent. Hence, the variance of the increments of $B(x)$ is

$$\mathbb{E}\big[(B(x+\varepsilon) - B(x))^2\big] = 2\mathbb{E}\big[(Z(t,x+\varepsilon) - Z(t,x))^2\big] + 2\mathbb{E}\big[(S(t,x+\varepsilon) - S(t,x))^2\big]. \tag{4.86}$$

We have already computed the variance of the increments of the second term in the right side above in (4.81). Thus, we only need to compute the variance of the increments of

$$Z(t,x) = \int_0^t \int_{\mathbb{R}} G(t-s, x-y)\dot{W}(ds, dy), \tag{4.87}$$

in order to show that

$$\mathbb{E}\big[(B(x+\varepsilon) - B(x))^2\big] = \varepsilon. \tag{4.88}$$

We proceed as in (4.77)-(4.78):

$$\begin{aligned}\mathbb{E}|Z(t,x+\varepsilon) - Z(t,x)|^2 &= \int_0^t \int_{\mathbb{R}} |G(t-s, x+\varepsilon-y) - G(t-s,x-y)|^2 dyds \\ &= \int_0^t \int_{\mathbb{R}} |G(s, y+\varepsilon) - G(s,y)|^2 dyds.\end{aligned} \tag{4.89}$$

The Fourier transform of the function $G(s, \cdot + \varepsilon)$ is

$$\int e^{-2\pi i\xi x} G(s, x+\varepsilon)dx = e^{2\pi i\varepsilon\xi}\hat{G}(s,\xi) = e^{2\pi i\varepsilon\xi - 4\pi^2|\xi|^2 s}.$$

The Plancherel identity implies that

$$\begin{aligned}&\mathbb{E}|Z(t,x+\varepsilon) - Z(t,x)|^2 = \int_0^t \int_{\mathbb{R}} |\hat{G}(s,\xi)|^2 |1 - e^{2\pi i\varepsilon\xi}|^2 d\xi ds \\ &= \int_0^t \int_{\mathbb{R}} e^{-8\pi^2|\xi|^2 s}\Big((1-\cos(2\pi\varepsilon\xi))^2 + \sin^2(2\pi\varepsilon\xi)\Big) d\xi ds \\ &= 2\int_0^t \int_{\mathbb{R}} e^{-8\pi^2|\xi|^2 s}\big(1-\cos(2\pi\varepsilon\xi)\big) d\xi ds = \frac{1}{4\pi^2}\int_{\mathbb{R}} \frac{1}{|\xi|^2}\Big(1 - e^{-8\pi^2|\xi|^2 t}\Big)\big(1-\cos(2\pi\varepsilon\xi)\big)d\xi \\ &= \frac{\varepsilon}{2\pi}\int_{\mathbb{R}} \frac{1}{|\xi|^2}\Big(1 - e^{-2|\xi|^2 t/\varepsilon^2}\Big)\big(1-\cos\xi\big)d\xi.\end{aligned} \tag{4.90}$$

Using (4.82) in (4.90) gives

$$\mathbb{E}|Z(t,x+\varepsilon) - Z(t,x)|^2 = \frac{\varepsilon}{2} - \frac{\varepsilon}{2\pi}\int_{\mathbb{R}} \frac{1}{|\xi|^2} e^{-2|\xi|^2 t/\varepsilon^2}\big(1-\cos\xi\big)d\xi. \tag{4.91}$$

Recalling the last expression in (4.81) for the variance of $S(t,x+\varepsilon) - S(t,x)$ we deduce that

$$\mathbb{E}|Z(t,x+\varepsilon) - Z(t,x)|^2 + \mathbb{E}|S(t,x+\varepsilon) - S(t,x)|^2 = \frac{\varepsilon}{2}. \tag{4.92}$$

This gives (4.88). As we already know that $B(0) = 0$ a.s. and $B(x)$ is a Gaussian random field, we conclude that $B(x)$ is a standard Brownian motion. □

4.4.3 Flattening of the correctors as $t \to +\infty$

Let us now discuss what happens to the correctors that we have constructed above, in the long time limit $t \to +\infty$.

First, Theorem 4.7 says that for any $x \in \mathbb{R}$ fixed, we may decompose the solution

$$Z(t,x) = \int_0^t \int_{\mathbb{R}} G(t-s, x-y)\dot{W}(dsdy), \tag{4.93}$$

to the additive heat equation

$$\frac{\partial Z}{\partial t} = \frac{\partial^2 Z}{\partial x^2} + \dot{W}(t,x), \tag{4.94}$$

as

$$Z(t,x) = \frac{1}{\pi^{1/4}} X(t) + R(t). \tag{4.95}$$

Here, $X(t)$ is a fractional Brownian motion with the Hurst exponent $H = 1/4$. The variance of the increments of the corrector $R(t)$ are given by (4.52):

$$\begin{aligned} \mathbb{E}\big[(R(t_2) - R(t_1))^2\big] &= \frac{1}{2\pi}\int_0^\infty \frac{1}{z^2} e^{-2t_1 z^2}\Big(1 - e^{-(t_2-t_1)z^2}\Big)^2 dz \\ &\le C(t_2-t_1)^2 \int_0^\infty z^2 e^{-2t_1 z^2} dz \le \frac{C(t_2-t_1)^2}{t_1^{3/2}}. \end{aligned} \tag{4.96}$$

Thus, at large times the increments of $Z(t,x)$ are well approximated by those of a fractional Brownian motion with the exponent $H = 1/4$. The corrector $R(t)$ does not become small as it variance is of the order $O(\sqrt{t}$, as seen from (4.39). However, the variance of the increments of $R(t)$ becomes small: the corrector is "nearly flat".

A similar phenomenon happens with the spatial corrector. Theorem 4.9 says that at each time $t > 0$ the solution $Z(t,x)$ can be decomposed as

$$Z(t,x) = \frac{1}{\sqrt{2}} B(x) + S(t,x). \tag{4.97}$$

Here, $B(x)$ is a two-sided Brownian motion while $S(t,x)$ is defined by (4.75)

$$S(t,x) = Z(t,0) + \int_{[t,+\infty)\times\mathbb{R}} [G(s,y-x) - G(s,y)]\dot{W}_1(dsdy). \tag{4.98}$$

The corrector itself does not become small as $t \to +\infty$ because of the term $Z(t,0)$ whose variance, once again, grows as $O(\sqrt{t})$ due to (4.13). However, for any $x_1 \in \mathbb{R}$ and $x_2\mathbb{R}$ fixed the increment

$$I_S(t,x_1,x_2) := S(t,x_1) - S(t,x_2), \tag{4.99}$$

has variance that decreases to 0 as $t \to +\infty$: according to (4.81), for any $x_1, x_2 \in \mathbb{R}$ we have

$$\mathbb{E}|S(t,x_1) - S(t,x_2)|^2 = \frac{|x_1 - x_2|}{2\pi}\int_{\mathbb{R}} \frac{1}{|\zeta|^2} e^{-2|\zeta|^2 t/|x_1-x_2|^2}\big(1 - \cos(\zeta)\big)d\zeta \to 0, \quad \text{as } t \to +\infty. \tag{4.100}$$

Thus, the spatial increments of $Z(t,x)$ look more and more like the increments of the Brownian motion as $t \to +\infty$.

5 Stochastic Integrals

5.1 The Ito integral in one dimension

Before talking about stochastic integrals with respect to space-time white noises $\dot{W}(t,x)$, let us very briefly recall 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

$$f(t,\omega) = X(\omega)\mathbb{1}_{[a,b]}(t),$$

with some $b > a$. Here, $X(\omega)$ needs to be 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

$$\int_0^t f(s,\omega)dB_s = \begin{cases} 0, & \text{if } 0 < t < a, \\ X(\omega)(B_t - B_a), & \text{if } a < t < b, \\ X(\omega)(B_b - B_a), & \text{if } b < t. \end{cases} \tag{5.1}$$

This may be written more succinctly as

$$\int_0^t f(s,\omega)dB_s = X(\omega)[B_{t\wedge b} - B_{t\wedge a}]. \tag{5.2}$$

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

$$f(t,\omega) = \sum_{j=1}^{n} c_j f_j(t,\omega), \tag{5.3}$$

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 an elementary function $f(t,\omega)$ we have

$$\mathbb{E}\Big(\int_0^t f(s,\omega)dB_s\Big)^2 = \mathbb{E}(X^2(\omega))(t\wedge b - t\wedge a) = \int_0^t \mathbb{E}(f^2(s,\omega))ds. \tag{5.4}$$

The same identity holds for simple functions of the form (5.3).

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 how we will construct the stochastic integral with respect to space-time white noises.

Another important aspect of the Ito integral is that, the integral

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

is a martingale. This is easy to see for simple functions, and then for a general $\mathcal{F}_t$-adapted function $f(t,\omega)$, this follows from the Ito isometry and density of simple functions in $L^2(\Omega)$. The quadratic variation of I_t is

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

This follows from the Ito formula:

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

which shows that

$$I_t^2 - \int_0^t f^2(s,\omega)ds \tag{5.5}$$

is a martingale.

5.2 Space-time white noise as a martingale measure

In order to construct the stochastic integral with respect to a white noise, we split out one variable in the noise, and call it $t \geq 0$, while keeping all other variables as ”spatial variables”. Let $\mathcal{B}$ be the collection of the Borel sets on $\mathbb{R}^d$. 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

$$W_t(A) = \dot{W}([0,t] \times A), \tag{5.6}$$

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. 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

$$\begin{aligned}\mathbb{E}[(W_t(A) - W_s(A))W_u(A)] &= \mathbb{E}[(\dot{W}([0,t] \times A) - \dot{W}([0,s] \times A))\dot{W}([0,u] \times A)] \\ &= |([0,t] \times A) \cap ([0,u] \times A)| - |([0,t] \times A) \cap ([0,u] \times A)| \\ &= |([0,u] \times A)| - |([0,u] \times A)| = 0.\end{aligned} \tag{5.7}$$

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 $0 \le r \le s$, $A \in \mathcal{B}$. 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.

We now compute the quadratic variation of the martingale $W_t(A)$, for a given Borel set $A \in \mathcal{B}$. Let us recall that for a martingale N_t its quadratic variation $\langle N \rangle_t$ is an increasing process such that

$$N_t^2 - \langle N \rangle_t \tag{5.8}$$

is a martingale. Similarly, if N_t and M_t are martingales, their covariance is

$$\langle N, M \rangle_t = \frac{1}{4}\Big[\langle N + M, N + M \rangle_t - \langle N - M, N - M \rangle_t\Big]. \tag{5.9}$$

Note that

$$N_t M_t - \langle N, M \rangle_t \tag{5.10}$$

is a martingale.

We claim that given two Borel sets $A, B \in \mathcal{B}$, the white noise covariance is

$$Q_t(A, B) := \langle W(A), W(B) \rangle_t = t|A \cap B|. \tag{5.11}$$

Indeed, let $X(\omega)$ be $\mathcal{F}_s$ measurable. As the increment $W_t(A) - W_s(A)$ is independent of $\mathcal{F}_s$ for $t > s$, we can compute

$$\begin{aligned}\mathbb{E}((W_t^2(A) - t|A|)X) &= \mathbb{E}[(W_s^2(A) - s|A|)X] + \mathbb{E}[((W_t(A) - W_s(A))^2 - (t - s)|A|)X] \\ &\quad + 2\mathbb{E}[W_s(A)(W_t(A) - W_s(A))X)] = \mathbb{E}[(W_s^2(A) - s|A|)X].\end{aligned} \tag{5.12}$$

It follows that $W_t^2(A) - t|A|$ is a martingale and hence

$$Q_t(A, A) = t|A|. \tag{5.13}$$

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

$$Q_t(A, B) = 0, \tag{5.14}$$

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

$$\begin{aligned}Q_t(A, B) &= \langle W(A), W(B) \rangle_t = \langle (W(A \setminus B) + W(A \cap B)), (W(B \setminus A) + W(A \cap B) \rangle_t \\ &= \langle W(A \cap B), W(A \cap B) \rangle_t = t|A \cap B|,\end{aligned} \tag{5.15}$$

which is (5.11).

5.3 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

$$f(t,x,\omega) = X(\omega)\mathbb{1}_{(a,b]}(t)\mathbb{1}_A(x). \tag{5.16}$$

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 of the form (5.16), we define the stochastic-integral process of f, with respect to the martingale measure W_t defined in (5.6), as

$$(f\cdot W)_t(B)(\omega) = X(\omega)[W_{t\wedge b}(A\cap B) - W_{t\wedge a}(A\cap B)](\omega), \quad \text{for } B\in\mathcal{B}. \tag{5.17}$$

On the informal level, this agrees with

$$\begin{aligned}\int_0^t\int_B f(s,x,\omega)\dot W(dsdx) &= X(\omega)\int_0^t\int_B \mathbb{1}_{(a,b]}(s)\mathbb{1}_A(x)\dot W(dsdx)\\ &= \begin{cases} 0, & \text{if } 0<t<a<b,\\ X(\omega)[W_t(A\cap B)-W_a(A\cap B)], & \text{if } 0<a<t<b,\\ X(\omega)[W_b(A\cap B)-W_a(A\cap B)], & \text{if } 0<a<b<t.\end{cases}\end{aligned} \tag{5.18}$$

This is a direct generalization of (5.1)-(5.2) for the Ito integral. We can extend the definition (5.17) to simple functions in a straightforward way as linear combinations. Note that if f is a simple function then $f\cdot W_t$ defined by (5.17) is a martingale measure as well.

Let us now compute the second moment of the stochastic integral process of an elementary function $f(t,x,\omega)$ of the form (5.16). We take a Borel set B and find

$$\begin{aligned}&\mathbb{E}\Big[\big[(f\cdot W_t)(B)\big]^2\Big] = \mathbb{E}\Big[X^2[W_{t\wedge b}(A\cap B) - W_{t\wedge a}(A\cap B)]^2\Big]\\ &= \mathbb{E}[X^2W^2_{t\wedge b}(A\cap B)] + \mathbb{E}[X^2W^2_{t\wedge a}(A\cap B)] - 2\mathbb{E}[X^2W_{t\wedge b}(A\cap B)W_{t\wedge a}(A\cap B)].\end{aligned} \tag{5.19}$$

Let us recall that X is $\mathcal{F}_a$-measurable. Hence, by the definition of the quadratic variation of the martingale $W_t(A\cap B)$, we have

$$\begin{aligned}&\mathbb{E}\Big[X^2\big(W^2_{t\wedge b}(A\cap B) - \langle W(A\cap B), W(A\cap B)\rangle_{t\wedge b}\big)\Big]\\ &= \mathbb{E}\Big[X^2\big(W^2_{t\wedge a}(A\cap B) - \langle W(A\cap B), W(A\cap B)\rangle_{t\wedge a}\big)\Big],\end{aligned} \tag{5.20}$$

and, since $W_t(A\cap B)$ is a martingale, we also have

$$\mathbb{E}\Big[X^2W_{t\wedge b}(A\cap B)W_{t\wedge a}(A\cap B)\Big] = \mathbb{E}\Big[X^2W^2_{t\wedge a}(A\cap B))\Big]. \tag{5.21}$$

Using this in (5.19) gives

$$
\begin{aligned}
\mathbb{E}\Big[((f\cdot W_t)(B))^2\Big] &= \mathbb{E}\Big[X^2\big(W^2_{t\wedge a}(A\cap B)-\langle W(A\cap B),W(A\cap B)\rangle_{t\wedge a}\big)\Big]\\
&+\mathbb{E}\Big[X^2\langle W(A\cap B),W(A\cap B)\rangle_{t\wedge b})\Big]+\mathbb{E}\Big[X^2W^2_{t\wedge a}(A\cap B)\Big]-2\mathbb{E}\Big[X^2(W^2_{t\wedge a}(A\cap B))\Big]\\
&=\mathbb{E}\Big[X^2(\langle W(A\cap B),W(A\cap B)\rangle_{t\wedge b}-\langle W(A\cap B),W(A\cap B)\rangle_{t\wedge a})\Big].
\end{aligned}
\tag{5.22}
$$

Recalling expression (5.11) for the quadratic variation of the white noise, we obtain

$$
\mathbb{E}\Big[((f\cdot W_t)(B))^2\Big]=\mathbb{E}(X^2)\big[t\wedge b-t\wedge a\big]|A\cap B|. \tag{5.23}
$$

In other words, for an elementary function we have

$$
\mathbb{E}\Big[\int_0^t\int_B f(s,x,\omega)\dot{W}(dsdx)\Big]^2=\mathbb{E}\Big[\int_0^t\int_B f^2(s,x,\omega)dsdx\Big]. \tag{5.24}
$$

It is a straightforward exercise to extend the Ito isometry (5.24) to simple functions, using the independence of the increments of the white noise.

5.4 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 as

$$
\|f\|^2=\mathbb{E}\Big(\int_0^T\int_{\mathbb{R}^d}|f(t,x,\omega)|^2dxdt\Big). \tag{5.25}
$$

We will denote by P_2 the space of predictable function of a finite norm (5.25). 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 the Ito isometry (5.24)

$$
\mathbb{E}\Big[\int_0^t\int_B f(s,x,\omega)W(ds,dx)\Big]^2=\mathbb{E}\Big[\int_0^t\int_{B\times B}|f(s,x,\omega)|^2dxds\Big]. \tag{5.26}
$$

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 (5.26) shows that the sequence

$$
\int_0^t\int_B f(s,x,\omega)W(ds,dx) \tag{5.27}
$$

is Cauchy in $L^2(P)$. Hence, for any function $f\in P_2$ we may define the stochastic integral (5.27) as the limit in $L^2(P)$ of the

$$
\int_0^t\int_B f_n(s,x,\omega)W(ds,dx), \tag{5.28}
$$

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.

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

We now consider a very basic example of a parabolic SPDE

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{6.1}$$

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:

$$|f(u_1) - f(u_2)| \le K|u_1 - u_2|. \tag{6.2}$$

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. We will avoid this last assumption for the moment, as we would like to include the standard stochastic heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u\dot{W}(t, x). \tag{6.3}$$

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 $u_0(x) \ge 0$, 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 $d > 1$.

6.1 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 (6.1) would have the form

$$u(t, x) = \int_{\mathbb{R}} G(t, x - y)u_0(y)dy + \int_0^t \int_{\mathbb{R}} G(t - s, x - y)f(u(s, y))\dot{W}(dsdy), \tag{6.4}$$

and this will be our starting point. That is, a solution to (6.1) is a solution to (6.4) 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:

$$G(t, x) = \frac{1}{(4\pi t)^{1/2}} e^{-|x|^2/(4t)}. \tag{6.5}$$

These are also known as mild solutions.

Theorem 6.1 *The stochastic heat equation (6.1) 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*

$$\sup_{x\in\mathbb{R}} \sup_{0\le t\le T} \mathbb{E}(|u(t,x)|^2) < +\infty. \tag{6.6}$$

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

$$\|u\|^2_{P_{2,\infty}} = \sup_{x\in\mathbb{R}} \sup_{0\le t\le T} \mathbb{E}(|u(t,x)|^2). \tag{6.7}$$

Note that the stochastic integral in the right side of (6.4) makes sense for all functions in $P_{2,\infty}[0,T]$.

6.1.1 Uniqueness of the solution

We first prove uniqueness. Suppose that u and v are two mild solutions of (6.1) – or, equivalently, of (6.4) 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

$$z(t,x) = \int_0^t \int_{\mathbb{R}} G(t-s, x-y)[f(u(s,y)) - f(v(s,y))]\dot{W}(dyds). \tag{6.8}$$

The Ito isometry and the Lipschitz property of $f(u)$ implies that

$$\begin{aligned}
\mathbb{E}(|z(t,x)|^2) &= \int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)\mathbb{E}\Big[|f(u(s,y)) - f(v(s,y))|^2\Big]dxds \\
&\le K\int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)\mathbb{E}\Big[|u(s,y) - v(s,y)|^2\Big]dxds \\
&= K\int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)\mathbb{E}\Big[|z(s,y)|^2\Big]dxds.
\end{aligned} \tag{6.9}$$

We set

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

and get from (6.9) that

$$H(t) \le K\int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)H(s)dxds. \tag{6.10}$$

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 (6.10) that

$$H(t) \le K' \int_0^t \frac{H(s)}{|t-s|^{1/2}}ds. \tag{6.11}$$

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

$$\frac{1}{p} + \frac{1}{q} = 1, \tag{6.12}$$

implies that

$$H(t)^q \leq K'' \int_0^t H(s)^q ds. \tag{6.13}$$

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

6.1.2 Existence of a mild solution

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

$$u_{n+1}(t, x) = \int_{\mathbb{R}} G(t, x - y)u_0(y)dy + \int_0^t \int_{\mathbb{R}} G(t - s, x - y)f(u_n(s, y))\dot{W}(dsdy). \tag{6.14}$$

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

$$q_n(t, x) = \int_0^t \int_{\mathbb{R}} G(t - s, x - y)(f(u_n(s, y)) - f(u_{n-1}(s, y))\dot{W}(dsdy). \tag{6.15}$$

As f is Lipschitz, it follows that

$$\begin{aligned} \mathbb{E}(|q_n(t, x)|^2) &= \int_0^t \int_{\mathbb{R}} G^2(t - s, x - y)\mathbb{E}(f(u_n(s, y)) - f(u_{n-1}(s, y))^2 dyds \\ &\leq K^2 \int_0^t \int_{\mathbb{R}} G^2(t - s, x - y)\mathbb{E}|q_{n-1}(s, y)|^2 dyds. \end{aligned} \tag{6.16}$$

Hence, the function

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

satisfies

$$Z_n(t) \leq K^2 \int_0^t \int_{\mathbb{R}} G^2(t - s, x - y)Z_{n-1}(s)dyds, \tag{6.17}$$

and thus

$$Z_n(t) \leq C \int_0^t \frac{Z_{n-1}(s)}{|t - s|^{1/2}} ds. \tag{6.18}$$

Once again, with $p \in (1,2)$ and q as in (6.12) we get

$$Z_n(t)^q \le C \int_0^t Z_{n-1}(s)^q ds. \qquad (6.19)$$

Hence, Gronwall's lemma implies that

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

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

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

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

$$u(t,x) = \int_{\mathbb{R}} G(t,x-y) u_0(y) dy + \int_0^t \int_{\mathbb{R}} G(t-s,x-y) f(u(s,y)) \dot{W}(dsdy), \qquad (6.20)$$

finishing the existence proof. □

6.2 Higher moments of the solutions

One may ask if the solutions of the stochastic heat equation (6.1) that we have constructed lie in better spaces, such as $P_{s,\infty}$ with the norm

$$\|u\|_{P_{s,\infty}}^s = \sup_{x \in \mathbb{R},\, 0 \le t \le T} \sup \mathbb{E}(|u(t,x)|^s), \qquad (6.21)$$

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

$$\mathbb{E}|N_t|^p \le c_p \mathbb{E}(\langle N, N \rangle_t)^{p/2}, \qquad (6.22)$$

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

$$\mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} f(s,x) \dot{W}(dsdx)\Big]^p \le c_p \mathbb{E}\Big[\int_0^t \int_{\mathbb{R}^d} |f(s,x)|^2 dsdx\Big]^{p/2}. \qquad (6.23)$$

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

$$u(t,x)=\int_{\mathbb{R}}G(t,x-y)u_0(y)dy+\int_0^t\int_{\mathbb{R}}G(t-s,x-y)f(u(s,y))dW(s,y) \tag{6.24}$$

can be estimated as

$$\begin{aligned} M_s(t,x):=\mathbb{E}|u(t,x)|^s &\le C_0+C\mathbb{E}\Big[\int_0^t\int_{\mathbb{R}}G^2(t-s,x-y)f^2(u(s,y))dsdy\Big]^{s/2} \\ &\le C_0+C\mathbb{E}\Big[\int_0^t\int_{\mathbb{R}}G^2(t-s,x-y)u^2(s,y)dsdy\Big]^{s/2}. \end{aligned} \tag{6.25}$$

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

$$\begin{aligned} &M_4(t,x)\le C_0+C\mathbb{E}\Big[\int_0^t\int_{\mathbb{R}}G^2(t-s,x-y)u^2(s,y)dsdy\Big]^2 \\ &=C_0+C\int_0^t\int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}}G^2(t-s,x-y)G^2(t-s',x-y')\mathbb{E}\big[u^2(s,y)u^2(s',y')\big]dsdyds'dy'. \end{aligned} \tag{6.26}$$

Set

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

then we have, using the Cauchy inequality

$$\begin{aligned} \bar{M}_4(t)&\le C_0+C\int_0^t\int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}}G^2(t-s,y)G^2(t-s',y')\bar{M}_4^{1/2}(s)\bar{M}_4^{1/2}(s')dsdyds'dy' \\ &\le C_0+C\Big(\int_0^t\frac{\bar{M}_4(s)^{1/2}ds}{|t-s|^{1/2}}\Big)^2\le C_0+C\Big(\int_0^t\bar{M}_4^{q/2}(s)ds\Big)^{2/q}, \end{aligned} \tag{6.27}$$

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

$$\sup_{0\le t\le T}\bar{M}_4(t)\le C_T. \tag{6.28}$$

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 6.3 *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*

$$\sup_{x\in\mathbb{R}}\mathbb{E}(|u(t,x)|^k)\le C^ke^{Ck^3t}. \tag{6.29}$$

6.3 The spatial L^2-bound

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

$$u(t,x) = \int_{\mathbb{R}} G(t,x-y)u_0(y)dy + \int_0^t \int_{\mathbb{R}} G(t-s,x-y)f(u(s,y))dW(s,y), \tag{6.30}$$

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

$$|f(u)| \le K|u|. \tag{6.31}$$

The first integral in the right side of (6.30) is obviously in any $L^p(\mathbb{R})$, $1 \le p \le +\infty$, 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

$$\begin{aligned} &\mathbb{E}\int_{\mathbb{R}} u^2(t,x)dx \le 2\int |u_0(x)|^2dx + 2\int_{\mathbb{R}}\int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)\mathbb{E}[f^2(u(s,y))]dydsdx \\ &\le 2\int |u_0(x)|^2dx + 2K\int_{\mathbb{R}}\int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)\mathbb{E}[|u(s,y)|^2]dydsdx \\ &\le 2\|u_0\|_2^2 + C\int_0^t \int_{\mathbb{R}} \mathbb{E}[|u(s,y)|^2]dy\frac{ds}{|t-s|^{1/2}}. \end{aligned} \tag{6.32}$$

Thus,

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

satisfies

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

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

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

Gronwall's lemma implies that

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

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

$$\sup_{0\le t\le T} \mathbb{E}\int_{\mathbb{R}} |u(t,x)|^2 dx < +\infty. \tag{6.33}$$

In the special case of the stochastic heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u\dot{W}, \tag{6.34}$$

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

$$u(t,x) = v(t,x) + \int_0^t \int_{\mathbb{R}} G(t-s,x-y)u(s,y)W(dsdy). \tag{6.35}$$

It follows that

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

satisfies a closed equation

$$Z(t,x) = v^2(t,x) + \int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)Z(s,y)dy. \tag{6.36}$$

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

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2}, \tag{6.37}$$

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

$$\bar{Z}(t) = \|v(t)\|_{L^2}^2 + b\int_0^t \frac{\bar{Z}(s)}{\sqrt{t-s}} ds, \tag{6.38}$$

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

$$Z_\gamma(t) = a(t) + \int_0^t g(t-s)Z_\gamma(s)ds, \tag{6.39}$$

with

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

Let us choose

$$\gamma = \pi b^2, \tag{6.40}$$

so that

$$\int_0^\infty g(s)ds = 1. \tag{6.41}$$

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

$$\bar{Z}_\gamma = \lim_{t\to+\infty} Z_\gamma(t) \tag{6.42}$$

exists.

In order to find the limit, let us write

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

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

$$\bar{Z}_\gamma + \beta(t) = a(t) + \bar{Z}_\gamma \int_0^t g(t-s)ds + \int_0^t g(t-s)\beta(s)dy, \tag{6.43}$$

so that

$$\beta(t) = a(t) - \bar{Z}_\gamma \int_t^\infty g(s)ds + \int_0^t g(t-s)\beta(s)dy. \tag{6.44}$$

Integrating (6.44) gives

$$\int_0^t \beta(s)ds = \int_0^t a(s)ds - \bar{Z}_\gamma \int_0^t \int_s^\infty g(s')ds'ds + \int_0^t \int_0^s g(s-s')\beta(s')ds'ds. \tag{6.45}$$

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

$$\int_0^t \int_s^\infty g(s')ds'ds = \int_0^t s'g(s')ds' + t\int_t^\infty g(s')ds' \to \int_0^\infty sg(s)ds, \text{ as } t \to +\infty, \tag{6.46}$$

while for the last integral in the right side of (6.45) we have

$$\begin{aligned} &\int_0^t \int_0^s g(s-s')\beta(s')ds'ds \to \int_0^\infty \int_0^s g(s-s')\beta(s')ds'ds \\ &= \int_0^\infty \beta(s') \int_{s'}^\infty g(s-s')dsds' = \int_0^\infty \beta(s)ds, \text{ as } t \to +\infty. \end{aligned} \tag{6.47}$$

Going back to (6.45) we conclude that

$$\bar{Z}_\gamma = \Big(\int_0^\infty sg(s)ds\Big)^{-1} \int_0^\infty a(s)ds. \tag{6.48}$$

Therefore, the solution of the stochastic heat equation (6.34) satisfies

$$\mathbb{E}\|u(t)\|_{L^2}^2 \sim \bar{Z}_\gamma e^{\gamma t}, \text{ as } t \to +\infty, \tag{6.49}$$

with $\gamma > 0$ given by (6.40).

Exercise 6.5 Generalize this argument to the solutions of equations of the form

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{6.50}$$

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 (6.49).

The exponential growth of the second moment in (6.49) should be contrasted with the simple bound on the integral:

$$\mathbb{E}\int_{\mathbb{R}} u(t,x)dx = \int_{\mathbb{R}} u_0(x)dx. \tag{6.51}$$

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 (6.51) 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:

$$\sup_{u\in\mathbb{R}} |f(u)| \leq K. \tag{6.52}$$

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 6.6 *There exists a modification of the solution of (6.4) 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 3.3.

Theorem 6.7 *Let X_t, $t \in \mathbf{T} = [a_1,b_1]\times\dots,[a_d,b_d] \subset \mathbb{R}^d$ be a real-valued stochastic process. Suppose there are constants $k > 1$, $C > 0$ and $\alpha_i > 0$, $i = 1,\dots,d$, so that*

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

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

$$\mathbb{E}\big(|X(t) - X(s)|^k\big) \leq C\sum_{i=1}^d |t_i - s_i|^{\alpha_i}. \tag{6.53}$$

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 3.3. We will leave the proof as an exercise.

The proof of Theorem 6.6

Let us consider

$$U(t,x)=\int_0^t\int_{\mathbb{R}} G(t-s,x-y)f(u(s,y))dW(s,y). \tag{6.54}$$

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

$$\begin{aligned}U(t',x)-U(t,x)&=\int_0^t\int_{\mathbb{R}}[G(t'-s,x-y)-G(t-s,x-y)]f(u(s,y))dW(s,y)\\&+\int_t^{t'}\int_{\mathbb{R}} G(t'-s,x-y)f(u(s,y))dW(s,y).\end{aligned} \tag{6.55}$$

Young's and Burkholder's inequalities imply that

$$\begin{aligned}\mathbb{E}|U(t',x)-U(t,x)|^p&\le C_p\mathbb{E}\Big[\int_0^t\int_{\mathbb{R}}|G(t'-s,x-y)-G(t-s,x-y)|^2f^2(u(s,y))dyds\Big]^{p/2}\\&+C_p\mathbb{E}\Big[\int_t^{t'}\int_{\mathbb{R}} G^2(t'-s,x-y)f^2(u(s,y))dyds\Big]^{p/2}=I+II.\end{aligned} \tag{6.56}$$

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

$$II\le C\Big[\int_t^{t'}\int_{\mathbb{R}} G^2(t'-s,x-y)dyds\Big]^{p/2}\le C\Big[\int_t^{t'}\frac{ds}{|t'-s|^{1/2}}\Big]^{p/2}\le C|t'-t|^{p/4}. \tag{6.57}$$

For the first term in the right side of (6.56) we write, using the Plancherel identity

$$\begin{aligned}\int_{\mathbb{R}}|G(t'-s,x-y)-G(t-s,x-y)|^2dy&=\int_{\mathbb{R}}|G(t'-s,y)-G(t-s,y)|^2dy \quad (6.58)\\=C\int_{\mathbb{R}}\left|e^{-(t'-s)|\xi|^2}-e^{-(t-s)|\xi|^2}\right|^2d\xi&=C\int_{\mathbb{R}}e^{-2(t-s)|\xi|^2}\left[1-e^{-(t'-t)|\xi|^2}\right]^2d\xi.\end{aligned}$$

It follows that the first integral in there right side of (6.56) can be bounded as

$$I^{2/p}\le C\int_0^t\int_{\mathbb{R}}e^{-2(t-s)|\xi|^2}\left[1-e^{-(t'-t)|\xi|^2}\right]^2d\xi ds=C\int_{\mathbb{R}}\frac{1}{|\xi|^2}\left(1-e^{-2t|\xi|^2}\right)\left[1-e^{-(t'-t)|\xi|^2}\right]^2d\xi. \tag{6.59}$$

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

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

and, second,

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

Using these estimates in (6.59) 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

$$\mathbb{E}|U(t',x)-U(t,x)|^p \le C_T|t'-t|^{p/4}. \tag{6.60}$$

Exercise 6.8 Show that

$$\mathbb{E}|U(t,x)-U(t,x')|^p \le C_p\Big(\int_0^t\int_{\mathbb{R}} |G(t-s,y)-G(t-s,x-x'-y)|^2 dyds\Big)^{p/2}, \tag{6.61}$$

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

$$\mathbb{E}|U(t,x)-U(t,x')|^p \le C_T|x-x'|^{p/2}. \tag{6.62}$$

Summarizing, we have

$$\mathbb{E}|U(t',x')-U(t,x)|^p \le C_T\Big(|t'-t|^{p/4}+|x-x'|^{p/2}\Big). \tag{6.63}$$

Now, we use Theorem 6.7 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 6.9 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 (6.52).

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

$$\frac{\partial u}{\partial t}=\Delta u+f(u), \tag{6.64}$$

satisfy the comparison principle. That is, if $u(t,x)$ and $v(t,x)$ are two solutions of (6.64) 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

$$\frac{\partial u}{\partial t} = \Delta u + g(t,x)f(u), \tag{6.65}$$

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

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

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{6.66}$$

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 6.10 *Let $u(t,x)$ and $v(t,x)$ be two solutions of (6.66) 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 6.10 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

$$\frac{du}{dt} = (A+B)u, \tag{6.67}$$

is given by

$$u(t) = e^{(A+B)t}u_0. \tag{6.68}$$

If the linear operators A and B commute then we have

$$u(t) = e^{At}v(t), \quad v(t) = e^{Bt}u_0. \tag{6.69}$$

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

$$e^{(A+B)t} = \lim_{n\to+\infty} (e^{A/n} e^{B/n})^n. \tag{6.70}$$

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

$$T_{nj} = \Big\{\frac{j}{n^2} \le t < \frac{j+1}{n^2}\Big\}. \tag{6.71}$$

On each time interval T_{nj} we first solve

$$\frac{\partial v}{\partial t} = Bv, \quad v(\frac{j}{n^2}) = u(\frac{j}{n^2}), \quad \frac{j}{n^2} \le t \le \frac{j+1}{n^2}, \tag{6.72}$$

followed by

$$\frac{\partial u}{\partial t} = Av, \quad u(\frac{j}{n^2}) = v(\frac{j+1}{n^2}), \quad \frac{j}{n^2} \le t \le \frac{j+1}{n^2}, \tag{6.73}$$

which gives us $u((j+1)/n^2)$, and we can solve (6.72) on the time interval $T_{n,j+1}$, and so on. Convergence of $u(t)$ to the solution of (6.67) 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

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{6.74}$$

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

$$\frac{\partial u_{n,j+1/2}}{\partial t} = f(u_{n,j+1/2})\dot{W}, \quad \frac{j}{n^2} \le t < \frac{j+1}{n^2}, \tag{6.75}$$

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

$$\frac{\partial u_{n,j+1}}{\partial t} = \frac{\partial^2 u_{n,j+1}}{\partial x^2}, \quad \frac{j}{n^2} \le t < \frac{j+1}{n^2}, \tag{6.76}$$

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 (6.75) on the time interval $T_{n,j+1}$, and we would be able to re-start.

One difficulty is making sense of (6.75) 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

$$u_{n,j+1/2}(t, \frac{k}{n}) = u_{nj}(\frac{j}{n^2}, \frac{k}{n}) + n\int_{j/n^2}^{t}\int_{I_{nk}} f(u_{n,j+1/2}(s, \frac{k}{n}))W(dsdy). \tag{6.77}$$

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

$$du_n(t, \frac{k}{n}) = f(u_n(t, \frac{k}{n}))dB_k, \tag{6.78}$$

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

$$\frac{\partial u_{n,j+1}}{\partial t} = \Delta_n u_{n,j+1}, \quad t \in T_{nj}, \tag{6.79}$$

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

$$\frac{\partial G_n}{\partial t} = \Delta_n G_n, \tag{6.80}$$

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

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

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)

$$u_n(t,x) = \int_{\mathbb{R}} \bar{G}_n(t,0,x,y)u_n(0,y)dy + \int_0^t \int_{\mathbb{R}} \bar{G}_n(t,s,x,y)f(u_n(s,y))W(dsdy), \tag{6.81}$$

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 (6.81) as

$$u_n(t,x) = u_n(\frac{j}{n^2},x) + \int_{j/n^2}^t \int_{\mathbb{R}} \bar{G}_n(t,s,x,y) f(u_n(s,y)) W(dsdy). \tag{6.82}$$

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, (6.82) says

$$u_n(t,x) = u_n(\frac{j}{n^2},x) + n\int_{j/n^2}^t \int_{I_{nk}} f(u_n(s,\frac{k}{n})) W(dsdy). \tag{6.83}$$

Therefore, on the time interval T_{nj}, the piece-wise constant (in space) function $u_n(t,x)$ satisfies the SDE (6.78). At the time $t = (j+1)/n^2$ the function $u_n(t,x)$ experiences a jump. To describe it, we go back to (6.81): the solution after the jump is given by

$$\begin{aligned} u_n(\frac{j+1}{n^2}+,x) = &\int_{\mathbb{R}} \bar{G}_n(\frac{j+1}{n^2},0,x,y) u_n(0,y) dy \\ &+ \int_0^{(j+1)/n^2} \int_{\mathbb{R}} \bar{G}_n(\frac{j+1}{n^2},s,x,y) f(u_n(s,y)) W(dsdy), \end{aligned} \tag{6.84}$$

while just before the jump we have

$$\begin{aligned} u_n(\frac{j+1}{n^2}-,x) = &\int_{\mathbb{R}} \bar{G}_n(\frac{j}{n^2},0,x,y) u_n(0,y) dy \\ &+ \int_0^{(j+1)/n^2} \int_{\mathbb{R}} \bar{G}_n(\frac{j}{n^2},s,x,y) f(u_n(s,y)) W(dsdy). \end{aligned} \tag{6.85}$$

The semi-group property for (6.80) implies that

$$G_n(t,\frac{k}{n},\frac{p}{n}) = \frac{1}{n}\sum_m G(t-s,\frac{k}{n},\frac{m}{n}) G(s,\frac{m}{n},\frac{p}{n}), \tag{6.86}$$

for all $0 \leq s \leq t$, k and p. The continuous version of (6.86) is

$$G_n(t,x,y) = \int_{\mathbb{R}} G_n(t-s,x,z) G_n(s,z,y) dz. \tag{6.87}$$

It follows from (6.87) 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 (6.84), together with the semi-group property gives

$$\begin{aligned} u_n(\frac{j+1}{n^2}+, x) &= \int_{\mathbb{R}}\int_{\mathbb{R}} G_n(\frac{1}{n^2}, x, z)\bar{G}_n(\frac{j}{n^2}, z, y)u_n(0, y)dydz \\ &+ \int_0^{(j+1)/n^2}\int_{\mathbb{R}}\int_{\mathbb{R}} G_n(\frac{1}{n^2}, x, z)\bar{G}_n(\frac{j}{n^2}, s, z, y)f(u_n(s, y))W(dsdy)dz \\ &= \int_{\mathbb{R}} G_n(\frac{1}{n^2}, x, z)u_n(\frac{j+1}{n^2}-, z)dz. \end{aligned} \tag{6.88}$$

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 (6.79).

Convergence of the approximation

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

Exercise 6.11 Show that the following two bounds hold: first,

$$\int_0^t\int_{\mathbb{R}}[\bar{G}_n(t, s, x, y) - G(t-s, x-y)]^2 dsdy \le \frac{c}{n}, \tag{6.89}$$

and, second,

$$\sup_{0\le s\le t}\sup_{x\in\mathbb{R}}\Big[\int_{\mathbb{R}}[\bar{G}_n(t, s, x, y) - G(t-s, x-y)]u(0, y)dy\Big]^2 \to 0 \text{ as } n\to+\infty. \tag{6.90}$$

We now show that

$$M(t) := \sup_{0\le s\le t}\sup_{x\in\mathbb{R}}\mathbb{E}|u_n(s, x) - u(s, x)|^2 \to 0 \text{ as } n\to+\infty. \tag{6.91}$$

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

$$\bar{M}(s, x) = \mathbb{E}|u_n(s, x) - u(s, x)|^2 \le C(I + II), \tag{6.92}$$

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

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

and can be bounded as

$$\begin{aligned}II&\le C\int_0^t\int_{\mathbb{R}}\mathbb{E}\Big|(G(t-s,x-y)-\bar{G}_n(t,s,x,y))f(u_n(s,y))\Big|^2dsdy\\&+C\int_0^t\int_{\mathbb{R}}\mathbb{E}\Big|G(t-s,x-y)(f(u(s,y))-f(u_n(s,y)))\Big|^2dsdy=II_1+II_2.\end{aligned}\tag{6.93}$$

It is straightforward to verify that there exists C_T so that

$$\sup_{0\le s\le T}\sup_{x\in\mathbb{R}}\mathbb{E}|u_n(s,x)|^2\le C_T.\tag{6.94}$$

This, together with (6.89) and the Lipschitz bond on f means that

$$II_1\le\frac{C_T}{n}.\tag{6.95}$$

The last term in right side of (6.93) is bounded as

$$II_2\le C\int_0^t\frac{M(s)ds}{\sqrt{t-s}}.\tag{6.96}$$

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

$$M(t)\le\alpha(n)+\int_0^t\frac{M(s)ds}{\sqrt{t-s}},\tag{6.97}$$

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

$$M(t)\to 0 \text{ as } n\to+\infty.\tag{6.98}$$

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 6.10.

The strong maximum principle

The heat equation

$$\frac{\partial u}{\partial t} = \Delta u, \tag{6.99}$$

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

$$\frac{\partial u}{\partial t} = \Delta u - \sqrt{u}, \tag{6.100}$$

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 (6.100) as

$$\frac{\partial u}{\partial t} = \Delta u - g(t,x)u, \tag{6.101}$$

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

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u\dot{W}, \tag{6.102}$$

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 (6.102) 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 (6.101) with the Dirichlet boundary conditions at $x = \pm 2R$:

$$\begin{aligned} &\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + u\dot{W}, \quad t > 0, \quad |x| < 2R, \\ &v(t,-2R) = v(t,2R) = 0, \end{aligned} \tag{6.103}$$

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

$$v(t,x) = \int_{|y|\le 2R} G_R(t,x,y)u_0(y)dy + \int_0^t \int_{|y|\le 2R} G_R(t-s,x,y)f(u(s,y))W(dyds). \tag{6.104}$$

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

$$\begin{aligned} &\frac{\partial G_R}{\partial t} = \frac{\partial^2 G_R(t,x,y)}{\partial x^2}, \quad t > 0, \quad |x| < 2R, \\ &G_R(t,-2R,y) = G_R(t,2R,y) = 0, \\ &G_R(0,x,y) = \delta(x-y). \end{aligned} \tag{6.105}$$

Exercise 6.13 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

$$v(t=1,x) > 0 \text{ for all } |x| \le R. \tag{6.106}$$

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

$$\mathbb{P}\Big(\mathcal{A}_{k+1}^c \Big| \mathcal{A}_1 \cap \cdots \cap \mathcal{A}_k\Big) < \frac{\varepsilon}{N_\varepsilon}. \tag{6.107}$$

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 (6.107) in hand, we would have

$$\mathbb{P}(\mathcal{A}_{N_\varepsilon}^c) \leq \sum_{k=0}^{N_\varepsilon-1} \mathbb{P}\Big(\mathcal{A}_{k+1}^c \Big| \mathcal{A}_1 \cap \cdots \cap \mathcal{A}_k\Big) < \varepsilon. \tag{6.108}$$

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

$$\mathbb{P}(v(t=1,x) > 0 \text{ for all } x \in [-R,R]) \geq \mathbb{P}(\mathcal{A}_{N_\varepsilon}) \geq 1-\varepsilon. \tag{6.109}$$

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

$$\mathbb{P}(v(t=1,x) > 0 \text{ for all } x \in [-R,R]) = 1, \tag{6.110}$$

finishing the proof. Thus, it suffices to verify (6.107) to finish the proof of Theorem 6.12.

To prove (6.107), let us assume that $\mathcal{A}_k$ occurs, so that $v(t_k,x) \geq \delta_k I_k(x)$. By the comparison principle, it is enough then to show that

$$\mathbb{P}\Big[\text{there exists } \delta_{k+1} > 0 \text{ so that } v(t_{k+1},x) \geq \delta_{k+1} I_{k+1}(x) \text{ for all } |x| \leq 2R\Big] \geq 1 - \frac{\varepsilon}{N}. \tag{6.111}$$

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

$$\begin{aligned}
&\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + v\dot{W}, \quad t > t_k, \quad |x| < 2R, \\
&v(t,-2R) = v(t,2R) = 0, \\
&v(t_k,x) = I_k(x).
\end{aligned} \tag{6.112}$$

Let us then write

$$\begin{aligned}
v(t_{k+1},x) &= \int_{-2R}^{2R} G_R(t_{k+1}-t_k,x,y) I_k(y)dy \\
&+ \int_{t_k}^{t_{k+1}} \int_{-2R}^{2R} G_R(t_{k+1}-s,x,y) v(s,y) W(dsdy) = v_1(t_{k+1},x) + v_2(t_{k+1},x).
\end{aligned} \tag{6.113}$$

Exercise 6.14 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 (6.107) we need to show that

$$\mathbb{P}\Big[\sup_{|x|\leq 2R} |v_2(t_{k+1},x)| \geq \frac{1}{20}\Big] < \frac{\varepsilon}{N}. \tag{6.114}$$

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:

$$\mathbb{P}\Big[\sup_{0\le t\le T}\sup_{|x|\le 2R}|N(t,x)|>\lambda\Big]\le C_R\exp\Big(-\frac{\lambda^2}{C_RT^{1/2}K^2}\Big). \tag{6.115}$$

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 (6.115) to v_2. To deal with this, we can consider instead the solution of a modified equation

$$\begin{aligned}&\frac{\partial\tilde v}{\partial t}=\frac{\partial^2\tilde v}{\partial x^2}+\chi(\tilde v)\tilde v\dot W,\quad t>t_k,\quad x\in[-2R,2R]\\&\tilde v(t,-2R)=\tilde v(t,2R)=0,\\&\tilde v(t_k,x)=I_k(x).\end{aligned} \tag{6.116}$$

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 (6.115) for $\tilde v(t,x)$ to see that

$$\mathbb{P}(\tau<1/N))=\mathbb{P}\Big[\sup_{t_k\le t\le t_{k+1}}\sup_{|x|\le 2R}\tilde v(t,x)>5\Big]\le C\exp\Big(-\frac{C\cdot 25}{(1/N)^{1/2}10^2}\Big)\le C\exp(-CN^{1/2}). \tag{6.117}$$

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 (6.115)

$$\mathbb{P}\Big[\sup_{t_k\le t\le t_{k+1}}\sup_{|x|\le 2R}|\tilde N(t,x)|>\frac{1}{20}\Big]\le C\exp\Big(-CN^{1/2}\Big). \tag{6.118}$$

Therefore, we can estimate

$$\begin{aligned}\mathbb{P}\Big[\sup_{|x|\le 2R}|v_2(t_{k+1},x)|\ge\frac{1}{20}\Big]&\le\mathbb{P}(\tau<1/N)+\mathbb{P}\Big[\sup_{0\le t\le T}\sup_{|x|\le 2R}|\tilde N(t,x)|>\frac{1}{20}\Big]\\&\le C\exp(-CN^{1/2}).\end{aligned} \tag{6.119}$$

This proves (6.114), hence (6.107), finishing the proof of Theorem 6.12 except for the large deviations estimate (6.115).

Exercise 6.15 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 (6.115) 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 (6.115):

$$\mathbb{P}\Big[\sup_{0\leq t\leq T}\sup_{|x|\leq R}|N(t,x)| > \lambda\Big] \leq C_R \exp\Big(-\frac{\lambda^2}{C_R T^{1/2} K^2}\Big). \tag{6.120}$$

Exercise 6.16 Use the scaling of both $G(t,x)$ and of the white noise to verify that it suffices to prove (6.120) 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

$$\mathbb{P}(|N(t,x) - N(t,y)| > \lambda) = \mathbb{P}(B_{\langle M\rangle_t} > \lambda) \le C\mathbb{P}(B_{CK^2|x-y|} > \lambda) \le Ce^{-c\lambda^2/K^2|x-y|}. \quad (6.121)$$

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

$$\mathbb{P}(|N(t,x) - N(s,x)| > \lambda) \le Ce^{-c\lambda^2/K^2|t-s|^{1/2}}. \quad (6.122)$$

Now, the proof of (6.115):

$$\mathbb{P}\Big[\sup_{0\le t\le T} \sup_{|x|\le R} |N(t,x)| > \lambda\Big] \le C_R \exp\Big(-\frac{\lambda^2}{C_R T^{1/2} K^2}\Big) \quad (6.123)$$

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 (6.121) and (6.122). Summing up all the differences leads to (6.123).

Exercise 6.17 Fill in the details in the last step in the proof.

7 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

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u, \tag{7.1}$$

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

$$\limsup_{t\to+\infty} \sup_{|x|\geq c't} u(t,x) = 0, \tag{7.2}$$

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

$$\liminf_{t\to+\infty} \inf_{|x|\leq c't} u(t,x) = 0, \tag{7.3}$$

Of course, this definition can be applied to other problems than (7.1).

Solutions with compactly supported initial conditions

Solutions of (7.1) 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

$$u(t,x) = e^t \int_{-1}^{1} e^{-|x-y|^2/(4t)} dy. \tag{7.4}$$

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

$$u(t,ct) \leq e^t \int_{-1}^{1} e^{-|ct-1|^2/(4t)} dy = 2\exp\Big\{(1-\frac{c^2}{4})t + \frac{c}{2} - \frac{1}{4t}\Big\} \to 0 \text{ as } t \to +\infty. \tag{7.5}$$

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

$$u(t,x) \geq 2e^t \int_{-1}^{1} e^{-|ct+1|^2/(4t)} dy = 2\exp\Big\{(1-\frac{c^2}{4})t - \frac{c}{2} - \frac{1}{4t}\Big\}, \tag{7.6}$$

hence (7.3) 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

$$u(t, 2t - \frac{1}{2}\log t + x_0 + \xi) \to \exp(-\xi), \tag{7.7}$$

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

$$X(t) = 2t - \frac{1}{2}\log t + x_0, \tag{7.8}$$

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 (7.1): it satisfies an approximate equation

$$\frac{\partial \tilde{u}}{\partial t} + \frac{1}{2t}\frac{\partial \tilde{u}}{\partial x} = \frac{\partial^2 \tilde{u}}{\partial x^2} + \tilde{u}. \tag{7.9}$$

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 (7.1) 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 (7.1) gives

$$c\lambda = \lambda^2 + 1. \tag{7.10}$$

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

$$u_0(x) \sim Ce^{-\lambda x}, \text{ as } x \to +\infty, \tag{7.11}$$

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

$$C_1 e^{-\lambda x} \le u_0(x) \le C_2 e^{-\lambda x}, \text{ for all } x > x_0. \tag{7.12}$$

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 (7.10). 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

$$u(t,x) \le C_0 e^{-\lambda(x-ct)}, \tag{7.13}$$

hence $u(t,x)$ spreads at most with the speed c. On the other hand, given $\lambda < 1$ we can find c from (7.10) 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

$$u(t,x) \ge Ce^{-\lambda(x-ct)} - C'e^{-\lambda'(x-ct)}, \quad \text{for all } t > 0 \text{ and } x \in \mathbb{R}. \tag{7.14}$$

In particular, we have

$$u(t, ct+y) > \alpha_0 \text{ for all } t > 0 \text{ and } y \in (a,b). \tag{7.15}$$

Exercise 7.1 Use (7.15) to show that (7.3) holds for all $c' \in [0,c)$.

Hence, solutions with an initial condition that has an exponential decay as in (7.11) with $\lambda < 1$ propagate with the speed $c > 2$ given by (7.10).

Exercise 7.2 Show that if the initial condition has an exponential decay with a rate faster than $\lambda_* = 1$, that is, if (7.11) 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

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u), \tag{7.16}$$

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 7.3 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 (7.12). Show that the spreading rate of the solutions of (7.1) is the same as for the solutions of

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + f'(0)v, \tag{7.17}$$

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

Spreading in the stochastic case

Let us now consider solutions of the stochastic heat equation

$$\frac{\partial u}{\partial t} = \nu\frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}, \tag{7.18}$$

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:

$$|f(u) - f(v)| \leq \bar{L}|u - v|, \tag{7.19}$$

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

$$f(u) \geq \beta u \text{ for all } u \geq 0. \tag{7.20}$$

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

$$\mathbb{E}|u(t,x)|^2 \sim \exp(-\lambda_*(x - c_*t), \quad \text{for } x > 0. \tag{7.21}$$

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

$$\bar{c}_2 = \inf\{c > 0 : \ \bar{S}(c) < 0\}, \tag{7.22}$$

and

$$\underline{c}_2 = \inf\{c > 0 : \ \underline{S}(c) > 0\}. \tag{7.23}$$

If

$$\bar{c}_2 = \underline{c}_2, \tag{7.24}$$

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 (7.1) we know that $\bar{c}_2 = \underline{c}_2 = 2\sqrt{\nu}$. In the stochastic case, it is known that (7.24) 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 (7.20).

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 7.4 *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 $\bar{S}(c) < 0$.*

There is a natural open question whether $c_0 = c_*$. As we have mentioned, this has been resolved recently affirmatively in [1] in the case $f(u) = u$. We will not address it at the moment.

Unlike in the deterministic equation

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

where the propagation speed is $c_* = 2\sqrt{\nu}$, the speed $c_* = \bar{L}^2/2$ in the stochastic case does not depend on the diffusivity ν – this is very natural since ν can be scaled out by the transformation

$$u(t,x) = v(\frac{t}{\nu}, \frac{x}{\nu}). \tag{7.25}$$

The function $v(s,y)$ satisfies

$$\frac{\partial v}{\partial s} = \frac{\partial^2 v}{\partial y^2} + f(v(s,y))\nu \dot{W}(\nu s, \nu y).$$

The white noise (in one spatial dimension) has the scaling (with the equality in law)

$$\dot{W}(\varepsilon t, \varepsilon x) = \frac{1}{\varepsilon}\dot{W}(t,x).$$

Hence, we have

$$\frac{\partial v}{\partial s} = \frac{\partial^2 v}{\partial y^2} + f(v(s,y))\dot{W}(s,y), \tag{7.26}$$

which is nothing but (7.18) with $\nu = 1$. The change of variable (7.25) does not change the speed of propagation since time and space are scaled identically, and the speed of propagation in (7.26) has no memory of ν. This is not immediately obvious since the initial condition for (7.26) is $v(s,x) = u_0(\nu x)$ and does apparently depend on ν. However, it is natural to expect that the speed of propagation is universal and does not depend on the initial condition, making it independent of ν. Hence, we will set $\nu = 1$ with no fear, and our starting point will be

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(u)\dot{W}. \tag{7.27}$$

The exponential solutions and guessing the speed

Let us now use a generalized version of the exponential solutions to try to guess the value of the minimal speed $c_* = \bar{L}^2/2$ in Theorem 7.4. We will assume $\bar{L} = 1$ and $f(u) = \bar{L}u$, so that $u(t,x)$ satisfies

$$u(t,x) = \int_{\mathbb{R}} G(t, x-y)u_0(y)dy + \int_0^t \int_{\mathbb{R}} G(t-s, x-y)u(s,y)W(dsdy). \tag{7.28}$$

However, one also look for global in time solutions, defined also for $t < 0$. To do this, one would consider a sequence of initial value problems starting at $t = -n$:

$$u_n(t,x) = \int_{\mathbb{R}} G(t+n, x-y)u_{0,n}(y)dy + \int_{-n}^t \int_{\mathbb{R}} G(t-s, x-y)u_n(s,y)W(dsdy), \quad t > -n. \tag{7.29}$$

Let us choose $u_{0,n}(x)$ as

$$u_{0,n} = e^{-\lambda(x+\lambda n)},$$

so that

$$\int_{\mathbb{R}} G(t+n, x-y) u_{0,n}(y) dy = e^{-\lambda(x-\lambda t)},$$

for all $t > -n$ and all n. Let us assume that the initial condition is taken at $t = -\infty$ and is

$$u_0(y) = e^{-\lambda x},$$

so that

$$\bar{u}(t, x) = e^{\lambda^2 t - \lambda x}.$$

We also introduce

$$v_n(t, x) = e^{-\lambda^2 t + \lambda x} u_n(t, x),$$

and (7.29) becomes

$$v_n(t, x) = 1 + \int_{-n}^{t} \int_{\mathbb{R}} G(t-s, x-y) e^{\lambda(x-y) - \lambda^2(t-s)} v_n(s, y) W(dsdy). \tag{7.30}$$

Exercise 7.5 Show that for $\lambda > 1/2$, the sequence $v_n(t, x)$ converges, as $n \to +\infty$, to the solution of

$$v(t, x) = 1 + \int_{-\infty}^{t} \int_{\mathbb{R}} G(t-s, x-y) e^{\lambda(x-y) - \lambda^2(t-s)} v(s, y) W(dsdy), \tag{7.31}$$

and that (7.31) has a well-defined solution in $P_{2,\infty}[(-\infty, +\infty)]$.

The function $v(t, x)$ is stationary in time and space. Thus, for $\lambda > 1/2$, (7.28) has special solutions of the form

$$u(t, x) = e^{-\lambda(x-\lambda t)} v(t, x), \tag{7.32}$$

where $v(t, x)$ is a space-time stationary field. Let us see for why the restriction $\lambda > 1/2$ appears. Let us denote

$$q = \mathbb{E}|v(t, x)|^2.$$

As $v(t, x)$ is time-space stationary, q must be a positive constant. Starting with (7.30), squaring and taking the expectation gives

$$q = 1 + q \int_0^{\infty} \int_{\mathbb{R}} G^2(s, y) e^{2\lambda y - 2\lambda^2 s} dsdy. \tag{7.33}$$

An explicit computation shows that

$$\int_0^{\infty} \int_{\mathbb{R}} G^2(s, y) e^{2\lambda y - 2\lambda^2 s} dsdy = \frac{1}{2\lambda}. \tag{7.34}$$

Thus, for $q > 0$ to exist, we must have $\lambda > 1/2$. As λ serves both as the exponential decay and the propagation speed in (7.32), the minimal propagation speed is $c_* = 1/2$. While the exponential solutions are special, they give a good indication of what one should expect.

The upper bound

We now prove Theorem 7.4. The upper bound follows from the following lemma.

Lemma 7.6 *If $c > c_* = \bar{L}^2/2$ then $u(t,x)$ satisfies*

$$\mathbb{E}(|u(t,x)|^2) \le A_c \exp\Big[-c(|x| - ct)\Big], \quad \text{for all } x \in \mathbb{R} \text{ and } t \ge 0. \tag{7.35}$$

The constant A_c depends only on c and u_0.

This lemma immediately implies the upper bound in Theorem 7.4. Indeed, let us take any speed $c > \bar{L}^2/2$, and write, for any $c' > 0$:

$$S(c') = \limsup_{t\to\infty} \frac{1}{t} \sup_{|x|\ge c't} \log \mathbb{E}(|u(t,x)|^2) \le \limsup_{t\to\infty} \frac{1}{t}\left[c^2 t - cc't\right] = c(c - c'). \tag{7.36}$$

Thus, for any $c' > \bar{L}^2/2$, we can choose $c = (c' + \bar{L}^2/2)/2 > \bar{L}^2/2$ and conclude that $S(c') < 0$. This proves the second statement in Theorem 7.4 modulo the proof of Lemma 7.6.

The proof of Lemma 7.6

We first need an auxiliary proposition. Given a function $\phi \in P_{2,\infty}[0,+\infty)$, a positive number $r > 0$ and $\lambda \in \mathbb{R}$, we set

$$N_{\lambda,r}(\phi) = \sup_{t\ge 0}\sup_{x\in\mathbb{R}} \Big[e^{\lambda x - rt}\mathbb{E}(|\phi(t,x)|^2)\Big]^{1/2}.$$

A typical example to keep in mind is to take a function $\phi(t,x)$ such that

$$\mathbb{E}(|\phi(t,x)|^2) \sim e^{-\omega x + \gamma t}\mathbb{1}_{[0,+\infty)}(x) = e^{-\omega(x - (\gamma/\omega)t)}\mathbb{1}_{[0,+\infty)}(x)$$

In that case, we have

$$N_{\lambda,r}(\phi) = \sup_{t\ge 0}\sup_{x\in\mathbb{R}} \Big[e^{\lambda x - rt}e^{-\omega x + \gamma t}\mathbb{1}_{[0,+\infty)}(x)\Big]^{1/2} < +\infty,$$

provided that $\lambda < \omega$ and $r > \gamma$. Thus, $N_{r,\lambda}(\phi)$ measures the rate of the exponential decay of the function ϕ. The next proposition shows how the decay translates under the stochastic heat kernel convolution.

Proposition 7.7 *Let $\phi(t,x) \in P_{2,\infty}[0,+\infty)$, and set*

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

Then we have

$$N_{\lambda,r}(v) \le \frac{N_{\lambda,r}(\phi)}{(4(2r - \lambda^2))^{1/4}}, \text{ for all } r > \frac{\lambda^2}{2}. \tag{7.37}$$

Proof. The Ito isometry implies

$$
\begin{aligned}
&e^{\lambda x - rt}\mathbb{E}(|v(t,x)|^2) = e^{\lambda x - rt}\int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)\mathbb{E}|\phi(s,y)|^2 dyds \\
&\le [N_{\lambda,r}(\phi)]^2 \int_0^t \int_{\mathbb{R}} G^2(t-s,x-y)e^{\lambda x - rt}e^{-\lambda y + rs} dyds \\
&= [N_{\lambda,r}(\phi)]^2 \int_0^t \int_{\mathbb{R}} G^2(s,y)e^{\lambda y - rs} dyds.
\end{aligned}
$$

It follows that

$$
[N_{\lambda,r}(v)]^2 \le [N_{\lambda,r}(\phi)]^2 \int_0^\infty \int_{\mathbb{R}} G^2(s,y)e^{\lambda y - rs} dyds,
$$

and (7.37) follows from the explicit computation of the integral in the right side. □

We now prove Lemma 7.6. Writing

$$
u(t,x) = \int G(t-s,x-y)u_0(y)dy + \int_0^t \int G(t-s,x-y)f(u(s,y))W(dsdy),
$$

we see that

$$
N_{\lambda,r}(u) \le N_{\lambda,r}(u_0) + N_{\lambda,r}(v), \tag{7.38}
$$

with

$$
v(t,x) = \int_0^t \int G(t-s,x-y)f(u(s,y))W(dsdy).
$$

Proposition 7.7 implies that

$$
N_{\lambda,r}(v) \le \frac{N_{\lambda,r}(f(u))}{(4(2r-\lambda^2))^{1/4}},
$$

for $r > \lambda^2/2$. The Lipschitz bound on f implies that

$$
N_{\lambda,r}(v) \le \bar{L}\frac{N_{\lambda,r}(u)}{(4(2r-\lambda^2))^{1/4}},
$$

Using this in (7.38) gives

$$
N_{\lambda,r}(u) \le N_{\lambda,r}(u_0) + \bar{L}\frac{N_{\lambda,r}(u)}{(4(2r-\lambda^2))^{1/4}}. \tag{7.39}
$$

Hence, $N_{\lambda,r}(u) < +\infty$ provided that $r > \lambda^2/2$ and

$$
\bar{L}^4 < 4(2r-\lambda^2). \tag{7.40}
$$

Hence, we may take $r = c^2$ and $\lambda = \pm c$, provided that $c > \bar{L}^2/2$, finishing the proof of Lemma 7.6.

The lower bound

We now prove the lower bound in Theorem 7.4. As $f(u) \geq \beta u$, we have the following inequality for $m_2(t,x) = \mathbb{E}|u(t,x)|^2$:

$$m_2(t,x) \geq |\bar{u}(t,x)|^2 + \beta^2 \int_0^t \int_{\mathbb{R}} G^2(t-s,x-y) m_2(s,y) dy ds. \tag{7.41}$$

Here, $\bar{u}(t,x)$ is the solution of the heat equation:

$$\begin{aligned} \frac{\partial \bar{u}}{\partial t} &= \frac{\partial^2 \bar{u}}{\partial x^2}, \\ \bar{u}(0,x) &= u_0(x). \end{aligned} \tag{7.42}$$

Let us take $\alpha > 0$ and set

$$M_\alpha(t) = \int_{\alpha t}^{\infty} m_2(t,x) dx.$$

Integrating (7.41) form $x = \alpha t$ to infinity gives

$$M_\alpha(t) \geq \int_{\alpha t}^{\infty} |\bar{u}(t,x)|^2 + \beta^2 \ . \tag{7.43}$$

Note that if $x - y \geq \alpha(t-s)$ and $y \geq \alpha s$ then $x \geq \alpha t$. It follows that

$$\int_{\alpha t}^{\infty} \int_0^t \int_{\mathbb{R}} G^2(t-s,x-y) m_2(s,y) dy ds dx \tag{7.44}$$

$$\geq \int_0^t ds \int_{\alpha s}^{\infty} dy \int_{y+\alpha(t-s)}^{\infty} dx G^2(t-s,x-y) m_2(s,y)$$

$$= \int_0^t ds \Big(\int_{\alpha s}^{\infty} m_2(s,y) dy \Big) \Big(\int_{\alpha(t-s)}^{\infty} G^2(t-s,x) dx \Big). \tag{7.45}$$

Hence, the function $M_\alpha(t)$ satisfies the inequality

$$M_\alpha(t) \geq \int_{\alpha t}^{\infty} |\bar{u}(t,x)|^2 dx + \beta^2 \int_0^t Q_\alpha(t-s) M_\alpha(s) ds, \tag{7.46}$$

with

$$Q_\alpha(t) = \int_{\alpha t}^{\infty} G^2(t,x) dx.$$

An identical argument shows that the function

$$M'_\alpha(t) = \int_{-\infty} -\alpha t m_2(t,x) dx$$

satisfies

$$M'_\alpha(t) \geq \int_{-\infty}^{\alpha t} |\bar{u}(t,x)|^2 dx + \beta^2 \int_0^t Q_\alpha(t-s) M'_\alpha(s) ds, \tag{7.47}$$

thus the sum

$$\bar{M}_\alpha(t) = \int_{|x| > \alpha t} m_2(t,x) dx$$

obeys the inequality

$$\bar{M}_\alpha(t) \geq \int_{|x| \geq \alpha t} |\bar{u}(t,x)|^2 dx + \beta^2 \int_0^t Q_\alpha(t-s) \bar{M}_\alpha(s) ds. \tag{7.48}$$

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. 67

[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. 67

[3] J. Huang, K. Lê, D. Nualart, Large time asymptotic for the parabolic Anderson model driven by spatially correlated noise, Preprint, 2015. 67

[4] C. Mueller, On the support of solutions to the heat equation with noise, Stochastics and Stochastic Reports, **37**, 1991, 225–245. 55