Stochastic PDEs

by Boris Rozovsky

Table of contents

Send corrections to `kloeckner@dam.brown.edu`.

Example: Heat Equation. Suppose $\omega \in \Omega$ is part of a probability space. Then chance can come in at any or all of these points:

$$\begin{aligned}
\frac{\partial u(t,x)}{\partial t} &= a(x,\omega)\frac{\partial^2}{\partial x^2}u(t,x) + f(t,x,\omega) \quad x \in (a,b)\\
u(0,x) &= \varphi(x,\omega)\\
u(t,a) &= \psi_1(t,\omega)\\
u(t,b) &= \psi_2(t,\omega)
\end{aligned}$$

1 Basic Facts from Stochastic Processes

Probability Theory	Measure Theory
ω – elementary random event (outcomes)	
$\Omega = \bigcup \omega$ – probability space/space of outcomes	Ω – set
Random events $\leftrightarrow$ subsets of $\Omega \supset A$	Algebra $\mathcal{A} \subset \mathcal{P}(\Omega)$ closed w.r.t. $\cap / \cup / \bar{\cdot}$.
Operations on events: $\cup$, $\cap$, $\bar{A} = \Omega \setminus A$.	
$\varnothing := \Omega \setminus \Omega$	
If A and B are random events, then $A \cup B$, $A \cap B$, $\bar{A}$ are r.e.	
Elementary properties of probability:	Measures (see below)
$P(A) \in [0,1]$, $P(\Omega) = 1$, additive for disjoint events.	

Definition 1.1. *A function $\mu(A)$ on the sets of an algebra $\mathcal{A}$ is called a* measure *if*

a) *the values of μ are non-negative and real,*

b) *μ is an additive function for any finite expression–explicitly, if $A = \bigcup_i A_i$ and $A_i \cap A_j = \varnothing$ iff $i \neq j$, then*

$$\mu(A) = \sum_{i=1}^{n} \mu(A_i).$$

Definition 1.2. *A system $\mathcal{F} \subset \mathcal{P}(\Omega)$ is called a σ-*algebra *if it is an algebra and, in addition, if $(A_i)_{i=1,2,\ldots}$, then also $\bigcup_i A_i \in \mathcal{F}$.*

It is an easy consequence that $\bigcap_i A_i \in \mathcal{F}$.

Definition 1.3. *A measure is called σ-*additive *if*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

if the A_i are mutually disjoint.

The above together form *Kolmogorov's Axioms of Probability:* A tuple $(\Omega, \mathcal{F}, P)$ is called a *probability space* (Ω a set, $\mathcal{F}$ a σ-algebra, P a probability measure).

Lemma 1.4. *Let ε be a set of events. Then there is a smallest σ-algebra $\mathcal{F}$ such that $\varepsilon \subset \mathcal{F}$.*

Definition 1.5. *A function $X: \Omega \to \mathbb{R}^n$ is called a* random variable *if it is $\mathcal{F}$-measurable, i.e. for arbitrary A belonging to the Borel-σ-algebra $\mathcal{B}(\mathbb{R}^n)$, the set $X^{-1}(A) \in \mathcal{F}$.*

Definition 1.6. Completion of $\mathcal{F}$ with respect to P: *For simplicity, $\Omega = (0, 1)$. P is the* Lebesgue measure*, $\mathcal{F}$ the Borel-σ-algebra $\mathcal{B}(0, 1)$ on $\Omega = (0, 1)$. $\mathcal{F}$ is called complete if it contains all subsets B of Ω with the property:*

There are subsets B^- and B^+ from $\mathcal{B}(0,1)$ such that $B^- \subset B \subset B^+$ and $P(B^+ \setminus B^-) = 0$.

This process maps $(\Omega, \mathcal{F}, P)$ to $(\Omega, \bar{\mathcal{F}}^P, P)$, where $\bar{\mathcal{F}}^P$ is the completion *of $\mathcal{F}$ w.r.t. P.*

Now suppose X is a random variable in $(\Omega, \mathcal{F}, P)$ in $\mathbb{R}^n$. $X^{-1}(\mathcal{B}(\mathbb{R}^n)) := \{X^{-1}(A): A \in \mathcal{B}(\mathbb{R}^n)\} = \{\Gamma: X(\Gamma) \in \mathcal{B}(\mathbb{R}^n)\}$. $\mathcal{H}_X$ is called the σ-algebra generated by X.

One reason to use this definition of a random variable is this:

Lemma 1.7. (Doob-Dynkin) *If $\mathcal{F}$ is generated by a random variable Y, then there exists a Borel function g such that $X = g(Y)$.*

1.1 Lebesgue Integral

Definition 1.8. *X on $(\Omega, \mathcal{F}, P)$ is called simple if it is $\mathcal{F}$-measurable and takes a finite number of values: $x_1, x_2, \ldots, x_n$.*

$\Omega_i = \{\omega: X(\omega) = x_i\} = X^{-1}(x_i)$. Then the Lebesuge integral is

$$\int_{\Omega} X \mathrm{d}P = \sum_{i=1}^{n} x_i P(\Omega_i).$$

Definition 1.9. *An arbitrary measurable function X on $(\Omega, \mathcal{F}, P)$ is called P-integrable if there exists a sequence of such simple functions X_n so that $X_n \to X$ a.s. and*

$$\lim_{n,m\to\infty} \int_{\Omega} |X_n - X_m| \mathrm{d}P = 0.$$

Lemma 1.10. *If X is P-integrable, then*

1. *There exists a* finite *limit*

$$\int_{\Omega} X \mathrm{d}P = \lim_{n\to\infty} \int_{\Omega} X_n \mathrm{d}P.$$

2. *This limit does not depend on the choice of the approximating system.*

If X is a random variable $X: \Omega \to \mathbb{R}^n$. Let $\mathcal{B}$ be Borel's σ-algebra on $\mathbb{R}^n$. Then

$$\mu_X(\underbrace{A}_{\in\mathcal{B}}) = P(X^{-1}(A)) = P(\omega: X(\omega) \in A)$$

is called the *distribution function* of X.

Theorem 1.11.

$$\int_{\Omega} f(X)\mathrm{d}P = \int_{\mathbb{R}^n} f(x)\mu_X(\mathrm{d}x).$$

Thus

$$E[X] = \int_{\mathbb{R}^n} X\mu_X(\mathrm{d}X).$$

Example 1.12. Let X have values $x_1, ..., x_n$. $\Omega_i = X^{-1}(x_i)$. $\mu_X(x_i) = P(\Omega_i)$. Then

$$E[X] = \sum x_i \mu_X(x_i) = \sum x_i P(\Omega_i).$$

1.2 Conditional Expectation

ξ and η are are random variables with a joint density $p(x, y)$.

Motivation:

$$E[\xi|\eta = y] = \int x\, p(x|y)\mathrm{d}x.$$

$$P(A|B) = \frac{P(A\cap B)}{P(B)}.$$

Now suppose X is a P-integrable random variable $(\Omega, \mathcal{F}, P)$. G is a σ-algebra on Ω, $\mathcal{G} \subset \mathcal{F}$.

Definition 1.13. *Let η be $\mathcal{F}$-measurable random variable. If there exists a P-integrable $\mathcal{G}$-measurable function ξ such that for any bounded $\mathcal{G}$-measurable function φ*

$$E(\xi\varphi) = E(\eta\varphi),$$

the ξ will be called conditional expectation *of η and denoted $E[\eta|\mathcal{G}]$.*

Properties of conditional expectation:

1. If η is $\mathcal{G}$-measurable, then $E[\eta|\mathcal{G}] = \eta$.

   Proof. (1) By assumption, η is $\mathcal{G}$-measurable. (2) Let φ be an arbitrary $\mathcal{G}$-measurable function. Then

$$E(\eta\varphi) = E(E(\eta|\mathcal{G})\varphi) = E(\eta\varphi).$$ □

2. HW: Prove that the conditional expectation is unique.
3. If f is bounded, $\mathcal{G}$-measurable, then

$$E[f(\omega)X|\mathcal{G}](\omega) = f(\omega)E[X|\mathcal{G}] \quad \text{(a.s.)}$$

4. Let $g(\omega, X)$ be an $\mathcal{F}$-measurable function. Then

$$E[g(\omega, X)|\sigma(X)] = E[g(\omega, c)|\sigma(X)]|_{c=X}.$$

5. Let $\mathcal{G}_1 \subset \mathcal{G}$ be σ-algebras. Then

$$E[E[X|\mathcal{G}]|\mathcal{G}_1] = E[X|\mathcal{G}_1].$$

This property can be memorized as "Small eats big".

Example 1.14. $\Omega = \bigcup_n \Omega_n$, $\Omega_i \cap \Omega_j = \varnothing$. Let $\mathcal{E} = \{\Omega_1, \Omega_2, \ldots\}$. Then $\sigma(\mathcal{E}) = \{\Omega_{i_1} \cup \Omega_{i_2} \cup \ldots\}$. $\Omega_0 = \Omega \setminus \Omega$?. Let ξ be a random variable

$$E[\xi|\sigma(\mathcal{E})] = \sum_i \frac{E[\xi \mathbf{1}_{\Omega_i}]}{P(\Omega_i)} \mathbf{1}_{\Omega_i}. \tag{1.1}$$

Proof of (1.1):

a) The right-hand side is a function of indicators of $\Omega_i \Rightarrow$ it is $\sigma(\mathcal{E})$-measurable.

b) $E[E[\xi|\sigma(\mathcal{E})]g] = E\xi g$ for all g which are $\sigma(\mathcal{E})$-measurable.
Suppose $g = 1_{\Omega_k}$. Then

$$E[\text{rhs}\,\mathbf{1}_{\Omega_k}] = E\left[\frac{E[\xi \mathbf{1}_{\Omega k}]}{P(\Omega_k)} \mathbf{1}_{\Omega_k}\right] = \frac{E[\xi \mathbf{1}_{\Omega_k}]}{P(\not\Omega_k)} P(\not\Omega_k) = E(\xi \mathbf{1}_{\Omega_k}).$$

rhs: $E(\xi \mathbf{1}_{\Omega_k})$. What is a $\sigma(\mathcal{E})$-measurable function? Answer: It is a function of the form

$$\xi = \sum_i y_i \mathbf{1}_{\Omega_i}.$$

What?

1.3 Stochastic Processes

Assume that for all t, we are given a random variable $X_t = X_t(\omega) \in \mathbb{R}^n$. t could be from $\{0, 1, 2, 3, \ldots\}$ or from (a, b), it does not matter. In the former case, X_t is called a sequence of r.v. or a discrete time stochastic process. In the latter, it is called a continuous time stochastic process. If $t \in \mathbb{R}^2$, then X_t is a two-parameter random field.

Motivation: If X is a random variable, $\mu_X(A) = P(\omega: X(\omega) \in A$.

Definition 1.15. *The (finite-dimensional) distribution of the stochastic process $(X_t)_{t \in T}$ are the measures defined on $\mathbb{R}^{nk} = \mathbb{R}^n \otimes \cdots \mathbb{R}^n$ given by*

$$\mu_{t_1, \ldots, t_k}(F_1 \otimes F_2 \otimes \cdots \otimes F_k) = P(\omega: X_{t_1} \in F_1, \ldots, X_{t_k} \in F_k),$$

where the $F_i \in \mathcal{B}(\mathbb{R}^n)$.

1.4 Brownian Motion (Wiener Processes)

Definition 1.16. *A real-valued process X_t is called Gaussian if its finite dimensional distributions are Gaussian $\Leftrightarrow (X_{t_1}, \ldots, X_{t_k}) \sim \mathcal{N}(k)$.*

Remember: A random variable ξ in $\mathbb{R}^k$ is called *normal (multinormal)* if there exists a vector $m \in \mathbb{R}^k$ and a symmetric non-negative $k \times k$-matrix $R = (R_{ij})$ such that

$$\varphi(\lambda) := E[e^{i(\xi, \lambda)}] = e^{i(m,\lambda) - (R\lambda, \lambda)/2}$$

for all $\lambda \in \mathbb{R}^k$, where $(\cdot, \cdot)$ represents an inner product, $m = E[\xi]$ and $R = \text{cov}(\xi_i, \xi_j)$.

Independence: Fact: $Y = (Y_1, \ldots, Y_n)$ are normal vectors in $\mathbb{R}^k$ with (m_i, R_i). Then elements of Y are independent iff

$$\varphi_\lambda(Y) = \prod_{i=1}^n \varphi_{\lambda_i}(Y_i),$$

where $\lambda = (\lambda_1, \ldots, \lambda_n)$, where $\lambda_i \in \mathbb{R}^n$.

Fact 2: $\zeta = (\zeta_1, \ldots, \zeta_m)$ is Gaussian iff for any $\lambda \in \mathbb{R}^m$,

$$(\zeta, \lambda) = \sum \lambda_i \zeta_i$$

is Gaussian in 1D.

Definition 1.17. *Brownian motion W_t is a one-dimensional continuous Gaussian process with*

$$E[W_t]=0, \quad E[W_t W_s]=t\wedge s:=\min(t,s).$$

Alternative Definition:

Definition 1.18. *Brownian motion W_t is a Brownian motion iff*

1. $W_0=0$
2. $\forall t,s\colon W_t-W_s\sim\mathcal{N}(0,t-s)$
3. W_{t_1}, $W_{t_2}-W_{t_1},\ldots$ *are independent for all partitions* $t_1<t_2<t_3<\cdots$.

Yet another:

Definition 1.19. *The property (3) in Definition 1.18 may be replaced by*
3′. $W_{t_n}-W_{t_{n-1}}$ *is independent of* $W_{t_{n-1}}-W_{t_{n-2}},\ldots$

Definition 1.20.

$$\mathcal{F}_t^W:=\sigma(\{W_{s_1},W_{s_2},\ldots\colon s_i\leqslant t\}).$$

Theorem 1.21. *Brownian motion is a* martingale *w.r.t.* $\mathcal{F}_t^W\Leftrightarrow$

$$E[W_t|\mathcal{F}_s^W]=W_s$$

for $s<t$. (This is also the definition of a martingale.)

Remark 1.22. $\sigma(W_{t_1}, W_{t_2}, \ldots, W_{t_n})=\sigma(W_{t_1}, W_{t_2}-W_{t_1}, \ldots, W_{t_n}-W_{t_{n-1}})$ (knowledge of one gives the other–add or subtract). This is important because RHS is independent, but LHS is not.

Corollary 1.23.

1. $E[W_t^2]=t$. *(So W_t grows roughly as $\sqrt{t}$.)*
2. $W_t^2/t\to 0$ *almost surely.*
   Proof: By Chebyshev's inequality, $P(|W_t/t|>c)<E[|W_t/t|^2]/c^2=t/t^2c^2\to 0$ *as* $t\to\infty$.

Law of iterated logarithm:

$$\varphi_t^0=\frac{W_t}{\sqrt{2t\log\log(1/t)}}, \quad \varphi_t^\infty=\frac{W_t}{\sqrt{2t\log\log(t)}},$$
$$\limsup_{t\to 0}\varphi_t^0=1, \quad \limsup_{t\to\infty}\varphi_t^\infty=1,$$
$$\liminf_{t\to 0}\varphi_t^0=-1, \quad \liminf_{t\to\infty}\varphi_t^\infty=-1.$$

Continuity and Differentiability:

- W_t is continuous.
- W_t is nowhere differentiable.

Spectral representation of Brownian motion:

Theorem 1.24.

$$\begin{aligned} W_t &= t\eta_0+\sum_{n=1}^{\infty}\eta_n\sin(n\,t)\approx t\,\eta_0+\sum_{n=1}^{N}\eta_n\sin(n\,t), \quad \textit{where}\\ \eta_n &\sim \mathcal{N}(0,2/\pi n^2) \quad (n\geqslant 1),\\ \eta_0 &\sim \mathcal{N}(0,1/\pi). \end{aligned}$$

Proof. Consider $t \in [0, \pi]$.

$$\tilde{W}_t := W_t - \frac{t}{\pi} W_\pi \quad \text{for } t \in [0, \pi].$$

Then

$$\tilde{W}(t) = \sum_{n=1}^{\infty} \eta_n \sin(n\,t),$$

where

$$\eta_n = \frac{2}{\pi} \int_0^\pi \tilde{W}(t) \sin(n\,t) \mathrm{d}t \quad (n > 0)$$

and

$$\eta_0 = \frac{W(\pi)}{\pi}.$$

First fact: η_n are Gaussian because linear combinations of normal r.v.s. are normal.

$$E\eta_k\eta_n = \frac{4}{\pi^2} \int_0^\pi \int_0^\pi (t \wedge s - t\,s/\pi) \sin(n\,t) \sin(k\,s) = \begin{cases} 0 & k \neq n, \\ \frac{2}{\pi n^2} & k = n > 0. \end{cases}$$

For $n = 0$,

$$E[\eta_0^2] = E\frac{W^2[\pi]}{\pi^2} = \frac{1}{\pi}.$$

□

2 The Itô Integral and Formula

Suppose we have some system described by X_t that has some additive noise ξ_t: $Y_t = X_t + \xi_t$. $(t = 1, 2, 3, \ldots)$ The $\xi_1, \xi_2, \ldots$ are assumed to be

1. iid
2. $\xi_i \sim N(\mu, \sigma^2)$

If the ξ_t satisfy the first property, they are called *white noise*. If they satisfy both, it is *Gaussian white noise*.

If we now consider W_t $\xi_0 = W_0 = 0$, $\xi_1 = W_{t_1} - W_0$, $\xi_2 = W_{t_2} - W_{t_1}$, ..., then

1. holds
2. holds

A popular model in dynamics is

$$X_{t+\Delta} = A\,X_t + B + \xi_{t+1}$$

for, say, the dynamics of an "aircraft". Another possibility is modeling the price of a risky asset

$$X_{t+\Delta} = X_t + \mu X_t \Delta + \sigma X_t (W_{t+1} - W_t),$$

where μ is the individual trend of the stock, while σ is market-introduced volatility. Equivalently, we might write

$$\frac{X_{t+\Delta} - X_t}{\Delta} = \mu X_t - \sigma X_t \frac{W_{t+1} - W_t}{\Delta}$$

and then let $\Delta t \downarrow 0$, such that we obtain

$$\dot{X}_t = \mu X_t + \sigma X_t \dot{W}_t,$$

which is all nice and well except that the derivative of white noise does not exist. But note that there is less of a problem defining the same equation in integral terms.

Step 1: Suppose we have a function $f(s)$, which might be random. Then define

$$I_n(f) = \sum_k f(s_k^*)(W_{s_{k+1}} - W_{s_k}).$$

But what happens if $f(s)=W_s$. We get the term

$$W_{s_k}(W_{s_{k+1}}-W_{s_k}).$$

Or is it

$$W_{s_{k+1}}(W_{s_{k+1}}-W_{s_k})?$$

Or even

$$W_{\frac{s_{k+1}+s_k}{2}}(W_{s_{k+1}}-W_{s_k})?$$

In the Riemann integral, it does not matter where you evaluate the integrand–it all converges to the same value. But here, we run into trouble. Consider

$$\begin{array}{rcl} E|W_{s_k}(W_{s_{k+1}}-W_{s_k})|^2 & \neq & E|W_{s_{k+1}}(W_{s_{k+1}}-W_{s_k})|^2 \\ \| & & \nparallel \\ E|W_{s_k}|^2 E|W_{s_{k+1}}-W_{s_k}|^2 & & E|W_{s_{k-1}}(W_{s_{k+1}}-W_{s_k})|^2 \\ \| & & \nparallel \\ s_k(s_{k+1}-s_k) & & E|W_{\frac{s_{k+1}+s_k}{2}}(W_{s_{k+1}}-W_{s_k})|^2. \end{array}$$

Problem: Compute each of the above expectations, and show they are not equal.

2.1 The Itô Construction

The idea here is to use simple functions:

$$f(s)=\sum_{i=0}^{n} e_i(\omega)\mathbf{1}_{(t_i,t_{i+1})}(s),$$

where e_i is $\mathcal{F}_{t_i}^W$-measurable, where $\mathcal{F}_{t_i}^W=\sigma(W_{s_1},\ldots,W_{s_k}\colon s_i\leqslant s)$

$$\Leftrightarrow$$
$$e_i=e_i(W_r, r\in[0,t_i])$$
$$\Leftrightarrow$$
$$e_i \text{ is “adapted” to } \mathcal{F}_{t_i}^W.$$

Definition 2.1.

$$I(f)=\sum_{i=0}^{n} e_i(W_{t_{i+1}}-W_{t_i}).$$

Properties:

1. $E[I(f)]=0$

   Proof:

$$\begin{aligned} E[I(f)] &= \sum_{i=0}^{n} E\, e_i(W_{t_{i+1}}-W_{t_i}) \\ &= \sum_{i=0}^{n} E\big(E\big(e_i(W_{t_{i+1}}-W_{t_i})\big|\mathcal{F}_{t_i}^W\big)\big) \\ &= \sum_{i=0}^{n} E(e_i E[(W_{t_{i+1}}-W_{t_i})|\mathcal{F}_{t_i}^W]) \\ &= \sum_{i=0}^{n} E(e_i E[W_{t_{i+1}}-W_{t_i}]) \\ &= \sum_{i=0}^{n} E(e_i\, 0)=0. \end{aligned}$$

2.

$$\begin{aligned}
E|I(f)|^2 &= \sum_{i=1}^{N} E|e_i|^2(t_{i+1}-t_i) = \int_0^T E|f(s)|^2 \mathrm{d}s \\
&= E\Big[\sum e_i(W_{t_{i+1}}-W_{t_i})\Big]^2 \\
&= \sum E\big[e_i^2(W_{t_{i+1}}-W_{t_i})^2\big] - E\big[e_i e_j(W_{t_{i+1}}-W_{t_i})(W_{t_{j+1}}-W_{t_j})\big] \\
&= E\big(E(e_i^2(W_{t_{i+1}}-W_{t_i})|\mathcal{F}_{t_i}^W)\big) \\
&= E[e_i^2](t_{i+1}-t_i).
\end{aligned}$$

3. $I(f)$ is linear.

Next: If $f(s)$ is only $\mathcal{F}_s^W$-measurable (but not a step function), and if $\int_0^T E\, f^2(s)\mathrm{d}s < \infty \Rightarrow$ could be approximated by a sequence of step functions $f_n(s) \to f(s)$.

[Insert `lecture6.pdf` here, courtesy of Mario.]

2.2 Itô's Formula

Suppose we have a partition of a time interval $(0, T)$ as $t_0 = 0, t_1, t_2, \ldots, t_n = T$. $\Delta t_i = t_{i+1} - t_i$. We assume $\max \Delta t_i \to 0$. Also, we assume we have a function

$$f = f(t), \quad \Delta f_i = f(t_{i+1}) + f(t_i).$$

a) If $f = f(t)$, continuous, bounded variation. Then

$$\lim_{\max \Delta t_i \to 0} \sum_{i=0}^{n-1} |\Delta f_i|^2 = \lim_{\max \Delta t_i \to 0} \max \underbrace{|\Delta f_i|}_{\to 0} \underbrace{\sum_{i=0}^{n-1} |\Delta f_i|}_{\text{variation} \to \text{bounded}} = 0.$$

b) If $W = W(t)$ is Standard Brownian Motion, then

$$\lim_{\max \Delta t_i} \sum_{i=0}^{n-1} |\Delta W_i|^2 = T \quad (\text{in } L^2 \text{ and in probability}).$$

Proof. We need $E|\sum |\Delta W_i|^2| - T|^2 \to 0$. So

$$\begin{aligned}
& E\left(\left(\sum_i (\Delta W_i)^2\right)^2 - 2\sum_i (\Delta W_i)^2 T + T^2\right) \\
= & E\left[\sum_{i,j} |\Delta W_i|^2 |\Delta W_j|^2 - 2T^2 + T^2\right] \\
= & E\left[\sum_{i=0}^{n-1} |\Delta W_i|^4 + \sum_{i \neq j} |\Delta W_i|^2 |\Delta W_j|^2 - T^2\right] \\
= & 3\sum_i |\Delta t_i|^2 + \sum_{i \neq j} \Delta t_i \Delta t_j - T^2 \\
= & 2\sum_i |\Delta t_i|^2 + \underbrace{\left(\sum_i |\Delta t_i|\right)^2}_{T^2} - T^2 \\
= & 2\sum_i |\Delta t_i|^2 \leqslant 2\max\{\Delta t_i\} \cdot T \to 0.
\end{aligned}$$

□

So we essentially showed:

$$\begin{aligned}
\sum_{i=0}^{n-1} |\Delta W_i|^2 & \to T, \\
(\mathrm{d}W)^2 & \to \mathrm{d}t, \\
\mathrm{d}W & \to \sqrt{\mathrm{d}t}. \quad (\text{not rigorous})
\end{aligned}$$

2.2.1 Deriving from the Chain Rule

if $x = x(t) \in C^1$ and $F = F(y) \in C^1$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} F(x(t)) = F'(x(t)) x'(t).$$

Alternatively,

$$x(t) = x(0) + \int_0^t \underbrace{f(s)}_{x'(s)} \mathrm{d}s.$$

Then

$$F(x(t)) = F(x(0)) + \int_0^t F'(x(s)) f(s) \mathrm{d}s.$$

First of all, there is no "Stratonovich Formula". Suppose $W^n \rightrightarrows W$ (double arrows: uniformly), then

$$\begin{aligned}
X^n(t) &= X(0) + \underbrace{\int_0^t A(s)\mathrm{d}s + \int_0^t B(s)\dot{W^n}(s)\mathrm{d}s}_{=\int_0^t (X^n)'(s)\mathrm{d}s}, \\
X(t) &= X(0) + \int_0^t A(s)\mathrm{d}s + \underbrace{\int_0^t B(s) \circ \mathrm{d}W(s)}_{\text{Stratonovich Int.}}. \\
F(X^n(t)) &= F(X(0)) + \int_0^t F'(X^n(s))A(s)\mathrm{d}s + \int_0^t F'(X^n(s))B(s)\dot{W^n}(s)\mathrm{d}s \\
F(X(t)) &= F(X(0)) + \int_0^t F'(X(s))A(s)\mathrm{d}s + \int_0^t F'(X(s))B(s) \circ (s)\mathrm{d}W(s).
\end{aligned}$$

In particular,

$$X = W(t) = \int_0^t 1 \circ \mathrm{d}W(s), \quad F(y) = y^2, \quad \int_0^t W(s) \circ \mathrm{d}W(s) = \frac{1}{2}W^2(t).$$

Remark 2.2. Itô integral is a martingale, Stratonovich is *not*. Also: there is no connection between the two in the non-smooth case.

Now, let's see what happens for Itô, again starting from a process $X(t)$ given as

$$X(t) = X(0) + \int_0^t A(s)\mathrm{d}s + \int_0^t B(s)\mathrm{d}W(s).$$

Now, what is $F(X(t))$? Let's look at a Taylor expansion of

$$F(X(t_{i+1})) - F(X(t_i)) = F'(X(t_i))\Delta x_i + \frac{1}{2}F''(X(t_i))(\Delta x_i)^2 + (\cdots)\underbrace{(\Delta x_i)^3}_{\sim(\Delta t)^{3/2}}$$

So, in continuous time

$$\begin{aligned}
F(X(t)) &= \sum \Delta F \\
&= F(X(0)) + \int_0^t F'(X(s))\mathrm{d}X(s) + \frac{1}{2}\int_0^t F''(X(s))(\mathrm{d}X(s))^2 \\
&= F(X(0)) + \int_0^t F'(X(s))A(s)\mathrm{d}s + \int_0^t F'(X(s))B(s)\mathrm{d}W(s) + \frac{1}{2}\int_0^t F''(X(s))B^2(s)\mathrm{d}s
\end{aligned}$$

Theorem 2.3. *If*

$$X(t) = X(0) + \int_0^t A(s)\mathrm{d}s + \int_0^t B(s)\mathrm{d}W(s)$$

and $F \in C^3$, *then*

$$F(X(t)) = F(X(0)) + \int_0^t F'(X(s))A(s)\mathrm{d}s + \int_0^t F'(X(s))B(s)\mathrm{d}W(s) + \frac{1}{2}\int_0^t F''(X(s))B^2(s)\mathrm{d}s.$$

Now if $F \in C^3(\mathbb{R}^n, \mathbb{R}^n)$,then

$$X(t) = X(0) + \int_0^t A(s)\mathrm{d}s + \int_0^t B(s)\mathrm{d}W(s) \in \mathbb{R}^n,$$

where we recall that $W \in \mathbb{R}^p$ with all p components independent. Itô's Formula in multiple dimensions takes the form

$$F_k(X(t)) = F_k(X(0)) + \int_0^t \sum_i \frac{\partial F_k}{\partial x_i} A_i \mathrm{d}s + \sum_{i,l} \int_0^t \frac{\partial F_k}{\partial x_i} B_{i,l} \mathrm{d}W_l + \frac{1}{2}\int_0^t \sum_{i,j} \frac{\partial F_k}{\partial x_i \partial x_j} \sum_l B_{il}B_{jl}\mathrm{d}s.$$

Example 2.4. If $F(x)=x^2$ and

$$X=\int_0^t \mathrm{d}W(s),$$

then

$$\int_0^t W(s)\mathrm{d}W(s)=\frac{1}{2}(W^2(t)-t).$$

Example 2.5. If $\Delta F=0$ (i.e. F is harmonic), then $F(W(t))$ is a martingale.

2.2.2 SODEs

$$X(t)=X(0)+\int_0^t b(s,X(s))\mathrm{d}s+\int_0^t \sigma(s,X(s))\mathrm{d}W(s)$$

or, equivalently

$$\mathrm{d}X=b(t,X)\mathrm{d}t+\sigma(t,X)\mathrm{d}W(t).$$

Example 2.6. *Ornstein-Uhlenbeck Process*. The equation

$$\mathrm{d}X(t)=a\,X(t)\mathrm{d}t+b\,\mathrm{d}W(t)$$

has the solution

$$X(t)=e^{at}X(0)+b\int_0^t e^{a(t-s)}\mathrm{d}W(s).$$

Proof. Consider

$$\begin{aligned} X_t &= e^{at}X(0)+b\int_0^t e^{a(t-s)}\mathrm{d}W_s \\ &= e^{at}X(0)+b\,e^{at}\int_0^t e^{-as}\mathrm{d}W_s \\ &= e^{at}X(0)+b\,e^{at}Z_t \\ &= g(t,Z_t) \quad \text{with} \quad \mathrm{d}Z_t=e^{-at}\mathrm{d}W_t. \end{aligned}$$

Ito's Formula then gives

$$\begin{aligned} \mathrm{d}X_t &= \frac{\partial g}{\partial t}\mathrm{d}t+\frac{\partial g}{\partial x}\mathrm{d}Z_t+\frac{1}{2}\frac{\partial^2 g}{\partial x^2}(\mathrm{d}Z_t)^2 \\ &= X(0)a\,e^{at}+b\,e^{at}\mathrm{d}Z_t+0 \\ &= X(0)a\,e^{at}+b\,e^{at}e^{-at}\mathrm{d}W_t \\ &= X(0)a\,e^{at}+b\,\mathrm{d}W_t. \end{aligned}$$

□

Example 2.7. (*Geometric Brownian Motion*)

$$\mathrm{d}X(t)=a\,X(t)\mathrm{d}t+b\,X(t)\mathrm{d}W(t)$$

is solved by

$$X(t)=X(0)e^{(a-b^2/2)t+bW(t)}.$$

(Check this by Itô.)

Homework: Solve

$$\mathrm{d}X(t)=(a_1+a_2X(t))\mathrm{d}t+(b_1+b_2X)\mathrm{d}W(t).$$

Theorem 2.8. *If* $|b_i(s,x)-b_i(s,y)|+|\sigma_{i,k}(s,x)-\sigma_{i,k}(s,y)\leqslant C|x-y|$ *(a Lipschitz condition) and* $|b_i(s,x)|+|\sigma_{i,k}(s,x)|\leqslant C(1+|X|)$ *(a linear growth condition) and* $X(0)$ *is independent of* W *and* $E|X(0)|^2<\infty$*, then there exists a solution* $X(t)$ *that is continuous in time.* $X(t)$ *is measurable w.r.t* $\sigma(X(0), W(s), s\leqslant t)$ *and*

$$E\Big[\sup_{t\leqslant T}|X(t)|^2\Big]<\infty.$$

3 Some SPDEs

$$u_t = a\, u_{xx}, \quad u(0,x) = u_0(x).$$

($a > 0$–ellipticity: if it holds, then the equation is called parabolic) General solution:

$$u(t,x) = \frac{1}{\sqrt{4\pi\, a\, t}} \int_R \exp\biggl(- \frac{2|x-y|^2}{4\, a\, t} \biggr) u_0(y) \mathrm{d}y = E[u_0(x + \sqrt{2\pi} W(t)]$$

(*Feynman-Kac* formula–averaging over characteristics)

Monte-Carlo simulation:

$$\text{area}(A) = \frac{\#\text{hits in a set } A}{\#\text{hits in a surrounding square}}.$$

More general parabolic equation:

$$u_t(x,t) = a_{i\,j} D_i D_j u(x,t) + b_i D_i u(x,t) + c\, u + f \quad (t > 0, x \in \mathbb{R}^d) \quad u(0,x) = u_0(x)$$

This equation is parabolic iff $a_{\,i\,j} y_i\, y_j \geqslant a|y|^2$ for all $y \in \mathbb{R}^d$ (the ellipticity property). If the highest order partial differential operator in the equation is elliptic, then the equation is parabolic. (The elliptic equation would be

$$a_{i\,j} D_i D_j u(x,t) + b_i D_i u(x,t) + c\, u + f = 0.)$$

Now, onwards to *Stochastic* PDEs. A model equation is

$$\mathrm{d}u(t,x) = a\, u_{x\,x}(t,x) \mathrm{d}t + \sigma u_x(t,x) \mathrm{d}W_t.$$

Recall from geometric Brownian motion:

$$\mathrm{d}u(t) = a\, u(t) \mathrm{d}t + \sigma u(t) \mathrm{d}W_t, \quad u(0) = 0.$$

The solution is

$$u(t) = u_0 \exp\biggl(\biggl(a - \frac{\sigma^2}{2} \biggr) + \sigma W_t \biggr)$$

and

$$E[u^2(t)] = u_0^2 \exp\{u\, t\} E[\exp\{2\sigma W_t - \sigma^2 t^2\}].$$

Now consider

$$E\Biggl[\underbrace{\exp\biggl(b\, W_t - \frac{1}{2} b^2 t \biggr)}_{\rho(t)} \Biggr] = 1,$$

which is an example of an *exponential martingale*, which satisfies the general property

$$E[\rho(t) | \mathcal{F}_s^W] = \rho(s) \quad \text{for} \quad s < t, \quad \rho(0) = 1.$$

We find

$$E(\rho(t)] = E[\rho(s)] = E[\rho(0)] = 1.$$

Proof. By Ito's formula,

$$\mathrm{d}\rho(t) = b \rho(t) \mathrm{d}W_t \quad \Rightarrow \quad \rho(t) = 1 + b \int_0^t \rho(s) \mathrm{d}W_s.$$

□

Here's a crude analogy: In stochastic analysis, $\rho(t)$ plays the role of $\exp(t)$ in "regular" real analysis. Going back to our above computation, we find

$$E[u^2(t)] = u_0^2 \exp\{u\, t\} E[\exp\{2\sigma W_t - \sigma^2 t^2\}] = u_0 \exp\{2a\, t\}.$$

So we find for geometric Brownian motion that it remains square-integrable for all time. (Consider that this is also the case for the regular heat equation.) Now, let's return to our SPDE,

$$\mathrm{d}u(t,x) = a\, u_{x\,x}(t,x) \mathrm{d}t + \sigma u_x(t,x) \mathrm{d}W_t.$$

We begin by applying the Fourier transform to u, yielding $\hat{u}$.

$$\mathrm{d}\hat{u} = -a\,y^2\hat{u} + i\sigma y\hat{u}(t,y)\mathrm{d}W_t$$

$$\hat{u} = \hat{u}(0,y)\exp\left(-\left(a-\sigma^2/2\right)y^2t\right) + i\,y\sigma W_t).$$

Parseval's equality tells us

$$\int |u(t,x)|^2 = \int |\hat{u}(t,y)|^2\mathrm{d}y < \infty$$

iff $a - \sigma^2/2 > 0$. In SPDEs, first order derivatives in stochastic terms has the same strength as the second derivative in deterministic terms. The above condition is also called *super-ellipticity*, and the whole evolution equation is then called *super-parabolic*.

There's another example of SPDE in the lecture notes:

$$\mathrm{d}u(t,x) = a\,u_{xx}(t,x)\mathrm{d}t + \sigma u(t,x)\mathrm{d}W_t.$$

Here, the superellipticity equation is

$$a - \frac{0^2}{2} > 0 \quad \Leftrightarrow \quad a > 0.$$

For the homework, see the notes as well. One of these problems is to consider the more general equation

$$\mathrm{d}u = a_{ij}D_iD_ju + b_iD_iu + c\,u\mathrm{d}t + (\sigma_{ik}D_iu + \nu_k)\mathrm{d}W_k(t) \quad i,j = 1,\ldots,d, \quad k = 1,2,3,\ldots$$

where we have

$$\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots \\ \vdots & & \\ \sigma_{d1} & \sigma_{d2} & \cdots \end{pmatrix}.$$

We have to assume

$$\sigma\sigma^* = \sum_{k=1}^{\infty} \sigma_{ik}\sigma_{jk} < \infty, \qquad \sum_k \nu_k^2 < \infty.$$

A substitution that sometimes helps in the deterministic case is illustrated below:

$$\frac{\partial u}{\partial t} = a(t,x)u_{xx} + c\,u$$

Then we set $v(t,x) = e^{-ct}u(t,x)$ and obtain

$$\mathrm{d}v(t,x) = -c\,e^{-ct}u(t,x) + a(t,x)\,e^{-ct}u_{xx} + c\,e^{-ct}u = a(t,x)u\,v_{xx}.$$

For the stochastic case, note:

$$\mathrm{d}\rho(t) = \rho(t)\sigma\mathrm{d}W_t.$$

Then, let

$$\begin{aligned}
\eta(t) &:= e^{-\sigma W(t) - (\sigma^2/2)t} \\
\mathrm{d}\eta(t) &= -\eta(t)\sigma\mathrm{d}\Omega_t \\
\rho^{-1}(t) &= \eta(t)\exp(\sigma^2 t) \\
\mathrm{d}\rho^{-1}(t) &= -\eta(t)\sigma\mathrm{d}W_t\exp(\sigma^2 t) + \eta(t)\exp(\sigma^2 t)\sigma^2\mathrm{d}t = -\rho^{-1}\sigma\mathrm{d}W_t + \sigma^2\rho^{-1}(t)\mathrm{d}t
\end{aligned}$$

Applied to an SPDE, we get

$$\begin{aligned}
\mathrm{d}u(t,x) &= a\,u(t,x)\mathrm{d}t + \sigma u(t,x)\mathrm{d}W_t \\
u(0,x) &= u_0 \\
v(t,x) &= \underbrace{e^{-\sigma W(t) + (\sigma^2/2)t}}_{\rho^{-1}(t)}u(t,x) \\
\mathrm{d}(u(t,x)\rho^{-1}(t)) &= a\,v_{xx}\mathrm{d}t + \sigma v\mathrm{d}W_t - v\sigma\mathrm{d}W_t + \sigma^2 v\mathrm{d}t - \sigma^2 v\mathrm{d}t \\
&= a\,v_{xx}\mathrm{d}t.
\end{aligned}$$

Let $\tilde{W}(t)$ be a Wiener process independent of W.

$$v(t,x) = E\Big[u_0\Big(t+\sqrt{2a}\tilde{W}_t\Big)\Big].$$

Then

$$\begin{aligned} u(t,x) &= E\Big[u_0\Big(x+\sqrt{2a}\tilde{W}_t\Big)\Big]\exp(\sigma^2 W_t-(\sigma^2/2)t \\ &= E\Big[u_0\Big(x+\sqrt{2a}\tilde{W}_t\Big)\exp(\sigma^2 W_t-(\sigma^2/2)t\Big|\mathcal{F}_t^W\Big]. \end{aligned}$$

Example 3.1. Now consider

$$\mathrm{d}u(t,x) = a\,u_{xx}(t,x)+\sigma u_x(t,x)\mathrm{d}W_t \quad \Leftrightarrow \quad 2a-\sigma^2>0.$$

(Remark: There is not a chance to reduce to $\partial_t\tilde{u} = a\tilde{u}_{xx}$.)

$$\begin{aligned} \frac{\partial v}{\partial t} &= (a-\sigma^2/2)v_{xx}(t) \\ u(t,x) &= v(t,x+\sigma W(t)) \quad \text{then} \quad u \text{ verifies equation.} \end{aligned}$$

$$\begin{aligned} v(t,x) &= E\Big[u_0\Big(x+\sqrt{2a-\sigma^2}\tilde{W}(t)\Big)\Big] \\ &\Downarrow \\ u(t,x) &= E\Big[u_0\Big(x+W_t+\sqrt{2a-\sigma^2}\tilde{W}_t\Big)\Big|\mathcal{F}_t^W\Big]. \end{aligned}$$

(Note that, as above, the point of the conditional expectation is not measurability w.r.t. time (...), but with respect to W and not w.r.t. $\tilde{W}$.) By a naive application of Ito's formula, we would get

$$\begin{aligned} u(t,x) &= v(t,x+\sigma W_t) \\ v(t,x) &= u(t,x-\sigma W_t) \\ \mathrm{d}u(t,x-\sigma W_t) &= \sigma^2/2u_{xx}(t,x-\sigma W_t) \\ -\sigma u_x(t,x-\sigma W_t)\mathrm{d}W_t &= \frac{\sigma^2}{2}v_{xx}\mathrm{d}t-\sigma v_x\mathrm{d}W_t. \end{aligned}$$

But this is wrong because Ito's formula only applies to *deterministic* functions of brownian motion. The function u itself is random, though, so it does not work. To the rescue, the Ito-Wentzell formula.

Theorem 3.2. (Ito-Wentzell) *Suppose*

$$\mathrm{d}F(t,x) = J(t,x)\mathrm{d}t+H(t,x)\mathrm{d}W_t$$

and

$$\mathrm{d}Y(t) = b(t)\mathrm{d}t+\sigma(\tau)\mathrm{d}W_t.$$

Then

$$\mathrm{d}F(t,Y(t)) = \underbrace{J(Y(t))\mathrm{d}t+H(Y(t))\mathrm{d}W_t}_{\mathrm{d}_tF}+F_x(Y(t))b\mathrm{d}t+\frac{\sigma^2}{2}F_{xx}(Y(t))\mathrm{d}t+\sigma F_x(Y(t))\mathrm{d}W_t+H_x(t,Y(t))\sigma(t)\mathrm{d}t$$

For comparison, if we suppose $\mathrm{d}G(t,x) = J(t,x)\mathrm{d}t$ *and work out the regular Ito formula, we would find*

$$\mathrm{d}G(t,Y(t)) = \underbrace{J(t,Y(t))\mathrm{d}t}_{\mathrm{d}_tG}+G_x(Y(t))b(t)\mathrm{d}t+\frac{1}{2}G_{xx}\sigma^2\mathrm{d}t+G_x(Y)\mathrm{d}W_t.$$

4 PDE/Sobolev Recap

- Spaces: $H_2^\gamma = H_2^\gamma(\mathbb{R}^d)$
- Heat equation: H_2^γ, $L_2(\mathbb{R}^d)$, H_2^{-1}.
- an SPDE: H_2^γ, $L_2(\mathbb{R}^d)$, H_2^{-1}.

We will need:

- Gronwall Inequality: ...
- BDG Inequality ($p=1$)
$$E\left|\sup_{t\leqslant T}\int_0^t g(s)\mathrm{d}W_s\right|\leqslant C\,E\left|\int_0^T g^2(t)\mathrm{d}t\right|^{1/2}.$$
- ε-inequality
$$|a\,b|\leqslant \varepsilon a^2+\frac{1}{\varepsilon}b^2.$$
- Itô-Wentzell formula.

4.1 Sobolev Spaces H_2^γ

Definition 4.1. *Suppose $f\in C_0^\infty(\mathbb{R}^d)$. Then*
$$\hat{f}(y)=\frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} e^{-ixy}f(x)\mathrm{d}x.$$
Then we have Parseval's Inequality
$$\int_{\mathbb{R}^d}|f|^2\mathrm{d}x=\int_{\mathbb{R}^d}|\hat{f}|^2\mathrm{d}y$$
and define
$$\|f\|_\gamma:=\sqrt{\int_{\mathbb{R}^d}(1+|y|^2)^\gamma|\hat{f}(y)|^2\mathrm{d}y},$$
a norm. Then H_γ^2 is the closure of C_0^∞ in the norm $\|\cdot\|_y$.

$\delta(x)$, $\hat{\delta}(x)=\text{const}$, $\delta\in H_\gamma^2$ for what γ? ($\gamma<-d/2$?)

$H_2^0=L_2$, $H_2^{\gamma_1}\subset H_2^{\gamma_2}$ if $\gamma_1>\gamma_2$.

Sobolev embeddings: $H_2^{\gamma+d/2}\subset C^{0,\gamma}$ if $0<\gamma<1$. Alternative (but equivalent) definition:
$$H_2^n=\{f\colon f,Df,\ldots,D^nf\in L^2\}$$
with
$$\|f\|_n\sim\|f\|_{L^2}+\sum_{k=1}^n\left\|D^kf\right\|_{L^2}.$$
H_2^γ is a Hilbert space with
$$(f,g)_\gamma=\int_{\mathbb{R}^d}(1+|y|^2)^\gamma\hat{f}(y)\overline{\hat{g}(y)}\mathrm{d}y.$$
H_2^γ is dual to $H_2^{-\gamma}$ relative to L^2. ($\gamma>0$) Because if $f\in H_2^\gamma$ and $g\in H_2^{-\gamma}$. Then
$$(f,g)_0=\int_{\mathbb{R}^d}(1+|y|^2)^{\gamma/2}\hat{f}(y)\frac{\overline{\hat{g}(y)}}{(1+|\gamma|^2)^{\gamma/2}}\mathrm{d}y\leqslant\|f\|_\gamma\|g\|_{-\gamma}.$$
All this by S.L. Sobolev (1908-1989). Derived Sobolev spaces & generalized derivatives in the 1930s.

4.2 SPDEs in Sobolev Spaces

4.2.1 Classical Theory

Let's consider the heat equation in (H_2^1,L_2,H_2^{-1}), namely
$$u_t=u_{xx}+f,\quad u|_{t=0}=u_0.$$

Theorem 4.2. *If u is a classical solution and $u(t,\cdot)$ and u_0 are in $C_0^\infty(\mathbb{R})$, then*
$$\sup_{t\leqslant T}\|u(t)\|_0^2+\int_0^T\|u(t)\|_1^2\mathrm{d}t\leqslant C(T)\left(\|u_0\|_0^2+\int_0^T\|f(t)\|_{-1}^2\mathrm{d}t\right).$$

(Note the slight abuse of notation with $\|u(t)\|_\gamma$.)

Proof.

$$\begin{aligned}
\int u\frac{\partial u}{\partial t}\mathrm{d}x &= \int u\,u_{xx}\mathrm{d}x+\int u\,f\mathrm{d}x \quad \Big|\int\cdot\,u\mathrm{d}x\\
\| \\
\frac{\mathrm{d}v}{\mathrm{d}t} &= \|u_x\|_0^2+(u,f_0)\pm 2v(t)\\
v(t) &= v(0)-\int_0^t\Big(\|u(s)\|_0^2+\|u_x(s)\|_0^2\Big)\mathrm{d}s+\int_0^t(u,f)_0\mathrm{d}s+2\int_0^t v(s)\mathrm{d}s\\
v(t)+C\int_0^t\|u(s)\|_1^2\mathrm{d}s &\leqslant v(0)+\int_0^t\|u\|_1\|f\|_{-1}\mathrm{d}s+2\int v(s)\mathrm{d}s+\frac{C}{2}\int_0^t\|u\|_1^2\mathrm{d}s+C_1\int_0^t\|f\|_{-1}^2\mathrm{d}s\\
v(t)+\frac{C}{2}\int_0^t\|u(s)\|_1^2\mathrm{d}s &\leqslant F+2\int_0^t v(s)\mathrm{d}s\\
v(t) &\leqslant F+2\int_0^t v(s)\mathrm{d}s\\
\sup v(t) &\leqslant F.
\end{aligned}$$

where $v(t)=\frac{1}{2}\|u(t)\|_0^2$ and all the constant-tweaking is done with the ε-inequality. □

4.2.2 Stochastic Theory

$$\mathrm{d}u=(a(t)u_{xx}+f)\mathrm{d}t+(\sigma(t)u_x+g)\mathrm{d}W_t,$$

where $0<\delta<a(t)-\sigma^2(t)/2<C^*$. f,g adapted to $\mathcal{F}_t^W$, $u,f,g\in C_0^\infty$, $u|_{t=0}=u_0$ independent of W. Then

$$E\big[\sup\|u(t)\|_0\big]^2+E\int_0^T\|u(t)\|_1^2\mathrm{d}t\leqslant E\bigg(\|u_0\|_0^2+\int_0^T\|f\|_{-1}^2\mathrm{d}t+\int_0^T\|g\|_0^2\mathrm{d}t\bigg).$$

Step 1: WLOG, $\sigma=0$ (check at home!). Use the substitution

$$v(t,x)=u\bigg(t,x-\int_0^t\sigma(s)\mathrm{d}W_s\bigg).$$

Step 2: Ito formula for $|u(t,x)|^2$.

$$u^2=u_0^2+\underbrace{2\int_0^t a\,u_{xx}\,u\mathrm{d}s}_{-\|u\|_1^2}+\underbrace{\int_0^t f\,u\mathrm{d}s}_{\varepsilon\|u\|_1^2+C\|f\|_{-1}^2}+\int_0^t g\,u\mathrm{d}W_s+\int_0^t g^2\mathrm{d}s.$$

Step 3: Take expectation, which kills the $\mathrm{d}W_s$ term, giving a bound on

$$E\int_0^T\|u\|_1^2\mathrm{d}s\quad\text{and}\quad E\|u(t)\|_0^2.$$

Step 4: Take care of the sup, which is outside of the expectation, but needs to be inside.

$$E\bigg|\sup_t\int_0^{t_1}g\,u\mathrm{d}W\bigg|\leqslant C\,E\bigg(\int_0^T(g,u)_0^2\mathrm{d}t\bigg)^{1/2}\leqslant C\,E\bigg[\sup_t\int_0^T\|g\|_0^2\mathrm{d}t\bigg]\leqslant\varepsilon\mathrm{E}\sup_t\|u\|^2+C(\varepsilon)\int_0^t\|g\|_0^2\mathrm{d}s.$$

5 Nonlinear Filtering ("Hidden Markov Models")

State/signal X_t: Markov process/chain. Observation $Y_t=h(X_t)+g\dot{V}(t)$. State is not observed directly. The inf about X_t comes "only" from Y_s, $s\leqslant t$. Find the best mean-squares estimate of $f(X_t)$ given Y_s, $s\leqslant t$, where f is a known function. *Claim:* This estimator is given by

$$\hat{f}_t\ :=\ E\big[f(X_t)|\mathcal{F}_t^Y\big].$$

Proof. Let g_t be an $\mathcal{F}_t^Y$-measurable square-integable function $\Leftrightarrow E[g_t^2] < \infty$, $g_t = g(Y_0^t)$.

$$\begin{aligned} E[f_t - g_t]^2 &= E[f(X_t) - \hat{f}_t + \hat{f}_t - g_t]^2 \\ &= E[f(Y_t) - \hat{f}(X_t)]^2 + E[\hat{f}_t - g_t]^2 \\ &\geqslant E[f(X_t) - \hat{f}(X_t)]^2 + 2E[(f(Y_t) - \hat{f}_t)(\hat{f}_t - g_t)] \\ &= E[E[(f(X_t) - \hat{f}_t)(\hat{f}_t - g_t)|\mathcal{F}_t^Y]] = 0. \end{aligned}$$

Geometric interpretation: conditional expectation, with respect ot the σ-algebra $\mathcal{G}$ is an orthogonal projection on a space of $\mathcal{G}$-measurable functions.

$$\begin{aligned} \hat{f}_t &:= E[f(X_t)|\mathcal{F}_t^Y] \\ &= \int f(x) P(X_t \in \mathrm{d}x | \mathcal{F}_t^Y). \end{aligned}$$

□

State:

$$\begin{aligned} \mathrm{d}X_t &= b(X_t)\mathrm{d}t + \sigma(X(t))\mathrm{d}W_t \\ \mathrm{d}Y_t &= A(X(t))\mathrm{d}t + g(Y_t)\mathrm{d}V_t, \end{aligned}$$

We assume W_t and V_t are independent Wiener processes. $X(0) = x_0$, $Y(0) = 0$. Further $f = f(x)$, with $\sup_t E[f(X_t)^2] < \infty$.

$$\hat{f}_t = E[f(X_t)|\mathcal{F}_t^Y].$$

Zakai Equation of nonlinear filtering:

$$\hat{f}_t = \frac{\int f(x) u(t,x)\mathrm{d}x}{\int u(t,x)\mathrm{d}x},$$

where $u(t,x)$ is a solution of the SPDE

$$\mathrm{d}u(t,x) = \left[\frac{1}{2}\sigma^2(x) u(t,x)_{xx} - (b(x)u(t,x))_x\right]\mathrm{d}t + h(x)u(t,x)\mathrm{d}Y_t,$$

where $h = g^{-1}A$.

$$\begin{aligned} \tilde{P}(A) &= \int_A \exp\left\{-\int_0^T h\mathrm{d}s - \frac{1}{2}\int_0^T h^2 \mathrm{d}V\right\}\mathrm{d}P \\ \mathrm{d}Y_t &= \mathrm{d}V_t. \end{aligned}$$

If we add another term to the state process,

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X(t))\mathrm{d}W_t + f(X(t))\mathrm{d}V_t,$$

then we get

$$\mathrm{d}u(t,x) = \left[\left[\frac{1}{2}\sigma^2(x) + \rho^2\right] u(t,x)_{xx} - (b(x)u(t,x))_x\right]\mathrm{d}t - (\rho u(t,x))_x \mathrm{d}Y_t + h(x)u(t,x)\mathrm{d}Y_t$$

as the corresponding Zakai equation. (not sure about this last equation)

6 Solutions of PDEs and SPDEs

6.1 Classical Solutions

Here, we assume that u is twice continuously differentiable in x and once in t.

$$\dot{u}(t,x) = a(x)u_{xx}, \quad u(0,x) = u_0(x). \tag{6.1}$$

6.2 Generalized Solutions

First, let us talk about generalized functions. Suppose we wanted to find a derivative of $f(x) = \mathrm{sign}(x)$. Classically, $f'(0)$ does not exist. Let g be a differentiable function and φ very smooth with compact support. Then

$$\int f\varphi'(x)\mathrm{d}x = -\int f(x)\varphi(x)\mathrm{d}x.$$

If f is not differentiable,

$$\int f'(x)\varphi(x)\mathrm{d}x = -\int \varphi(x)\varphi'(x)\mathrm{d}x$$

for all $\varphi \in C_0^\infty(\mathbb{R}^n)$.

Now reconsider the heat equation in a different form, namely

$$\dot{u}(t,x) = (a(x)u_x)_x, \quad u(0,x) = u_0(x). \tag{6.2}$$

A weak general solution of (6.2) is a function $u \in H_2^1(\mathbb{R})$ such that for all $t > 0$

$$(u(t), \varphi) = (u_0, \varphi) - \int_0^t (u_x, \varphi_x)\mathrm{d}s$$

for every function $\varphi \in C_0^\infty(\mathbb{R})$.

Going back to (6.1), we find that a generalized solution is also a function from H_2^1 so that

$$(u(t), \varphi) = (u_0, \varphi) - \int_0^t (u_x, (a\varphi)_x)\mathrm{d}s$$

for all $\varphi \in C_0^\infty(\mathbb{R})$.

This definition is equivalent to saying that

$$u(t) = u_0 + \int a\, u_{xx}\mathrm{d}s$$

as an equality in H^{-1}.

6.3 Mild Solutions

Let us now consider yet another different equation, namely

$$\dot{u}(t,x) = u_{xx}(t,x) + \sin(u(t,x)), \quad u(t,x) = u_0(x). \tag{6.3}$$

Direct differentiation shows

$$u(t,x) = \int_{\mathbb{R}} k(t, x-y)u_0(y)\mathrm{d}y + \int_0^t \int_{\mathbb{R}} k(t-s, x-y)\sin(u(s,y))\mathrm{d}y\mathrm{d}s,$$

where k is the heat kernel

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

Write this now in SPDE form

$$\mathrm{d}u(t,x) = a\, u_{xx} + f(u(t,x)).$$

A *mild solution* is a solution u that satisfies

$$u(t,x) = \int_{\mathbb{R}} k(t, x-y)u_0(y)\mathrm{d}y + \int_0^t \int_{\mathbb{R}} k(t-s, x-y) f(u(s,y))\mathrm{d}y\mathrm{d}s.$$

6.4 Generalization of the notion of a "solution" in SDE

OSDE

$$\mathrm{d}X_t = b(X(t))\mathrm{d}t + \sigma(X(t))\mathrm{d}W_t, \quad X_0 = x_0.$$

Given b, σ, x_0, (Ω, P), W. If b and σ are Lipschitz-continuous and

$$|b(x)| \leqslant K(1+|x|), \quad |\sigma(x)| \leqslant K(1+|x|) \quad \Rightarrow \quad \exists! u.$$

Tanaka's Example shows an OSDE that can't be solved in this way:

$$\mathrm{d}X_t = \mathrm{sign}(X_t)\mathrm{d}W_t.$$

This equation has no solution for fixed $(\Omega,\ P)$, W. One could find $(\tilde{\Omega},\ \tilde{P})$, $\tilde{W}$ such that $\mathrm{d}X_t = \mathrm{sign}(X_t)\mathrm{d}\tilde{W}_t$. The mechanism for this is Girsanov's theorem, by which you can kill the drift and obtain a different equation.

If you specify the measure space and the Wiener process, you are looking for a *probabilistically strong soltuion*. If you allow yourself the freedom of choosing these as part of your solution, your solution is *probabilistically weak*.

7 Existence and Uniqueness

7.1 Scales of Sobolev Spaces

Simple Example: $x \in (0, b)$, $\Delta := \partial_x^2$, $\Lambda := 1 - \Delta$. $H := L^2(0, b)$. For smooth functions f, clearly

$$(\Lambda f, f)_H = ((1 - \Delta) f, f)_H = \int_0^b f^2(X)\mathrm{d}x + \int_0^b f_x^2 \mathrm{d}x =: \|f\|_{H_2^1}^2.$$

Let us consider the basis

$$\left\{ m_k(x) = \sqrt{\frac{2}{b}} \sin\frac{\pi(k-1)x}{b} \right\},$$

which is an ONS in H. Observe

$$\Lambda m_k = (1 - \Delta) m_k = m_k + \left[\frac{\pi(k-1)}{b}\right]^2 m_k = \left(1 + \left[\frac{\pi(k-1)}{b}\right]^2\right) m_k.$$

Define

$$\lambda_k := 1 + \left[\frac{\pi(k-1)}{b}\right]^2$$

as the eigenvalues of Λ w.r.t. the eigenbasis m_k. For $s \in (-\infty, \infty)$, we can construct an arbitrary power of the operator by defining its effect on the eigenbasis m_k by $\Lambda^s m_k := \lambda_k^s m_k$. Further, we may observe

$$(\Lambda^s f, f)_H = \sum_k \lambda_s^k f_k = \left(\Lambda^{s/2} f, \Lambda^{s/2} f\right) = \left\|\Lambda^{s/2}\right\|_H,$$

where

$$f_k = (f, m_k)_H$$

are the Fourier coefficients. Then the *Sobolev Space*

$$H_2^s(0, b) := \left\{ f \in H: \|f\|_s^2 := \left\|\Lambda^{s/2} f\right\|_H^2 < \infty \right\}.$$

For $s < 0$, define

$$H_2^s(0, b) := \Lambda^{-s} H.$$

We may also define

$$\|f\|_s := \sqrt{\sum_{k \geqslant 1} \left(\lambda_k^{s/2} f_k, \lambda_k^{s/2} f_k\right)}.$$ It was $\sum_{k \geqslant 1} \left(\lambda_k^{s/2} f_k, \lambda_k^s f_k\right)$ on the board, but that seemed wrong.

The spaces $\{H_2^s(0, b), s \in \mathbb{R}\}$ form the scale of spaces $H_2^{s_1} \subset H_2^{s_2}$ if $s_1 > s_2$.

Properties: Let $s_1 > s_2$. Then

1. H^{s_1} is dense in H^{s_2} in the norm $\|\cdot\|_{s_2}$.
2. H^s is a Hilbert space $(f, g)_s = \left(\Lambda^{s/2} f, \Lambda^{s/2} g\right)_0$.

3. For $s \geqslant 0$, $v \in H^{-s}(0, b)$, $u \in H^s(0, b)$, denote

$$[u, v] := \Big(\underbrace{\Lambda^s v}_{\in H}, \underbrace{\Lambda^{-s} u}_{\in H}\Big).$$

   a. If v also belongs to H, then $[u, v] = (v, u)_H$. Proof: Λ^s is self-adjoint in H.

Remark 7.1. We will typically work with three elements of the Sobolev scale–the middle, e.g. L^2, then the space where the solution lives and finally the space that the solution gets mapped to by the operator.

Important mnemonic rule:

$$\underbrace{\frac{\partial^n}{\partial x^n}}_{\Lambda^{n/2}}: H^s \to H^{s-n}.$$

7.2 Normal triples/Rigged Hilbert space/Gelfand's triple

Definition 7.2. *The triple of Hilbert spaces (V, H, V') is called a normal triple if the following conditions hold:*

1. $V \subset H \subset V'$.
2. *The imbeddings $V \to H \to V'$ are dense and continuous.*
3. *V' is the space dual to V with respect to the scalar product in H.*

Note that we always assume that H is identified with its dual.

Example 7.3. Any triple $H_2^{s+\gamma}, H^s, H^{s-\gamma}$ for $\gamma \geqslant 0$ is a normal triple.

7.3 Actual SPDEs

$$\mathrm{d}u(t) = (A\, u(t) = f(t))\mathrm{d}t + \sum_{k=1}^{\infty} (M_k u(t) + g_k(t))\mathrm{d}W_k^t, \quad u(0) = u_0 \in H.$$

We will assume that $A: V \to V'$ and $M_k: V \to H$, and further $f \in L^2(0, T; V')$ and $g_k \in L^2(0, T; H)$. We further assume $f(t)$ and $g_k(t)$ are $\mathcal{F}_t^W$-measurable, and $V = H_2^1(\mathbb{R}^d)$, $H = L_2(\mathbb{R}^d)$, $V' = H^{-1}(\mathbb{R}^d)$.

$$A\, u = \sum_{i,j} (a^{i,j}(t, x) u_{x_i})_{x_j} + \sum_i b^i(t, x) u_{x_i} + c.$$

$$M_k u = \sum_i \sigma^{i,k}(t, x) u_{x_i} + h^k(t, x) u.$$

We might also want to consider

$$A\, u = \sum_{|\alpha| \leqslant 2n} a_\alpha \partial^\alpha u, \quad M_k u = \sum_{|\alpha| \leqslant n} \sigma_\alpha \partial^\alpha u.$$

8 Existence and Uniqueness

We assume we have a normal triple $V \subset H \subset V'$. Consider

$$\mathrm{d}u(t) = (A\, u(t) + f(t))\mathrm{d}t + (\mu_k u(t) + g_k(t))\mathrm{d}W_k(t), \tag{8.1}$$

where we assume that W_k are infinitely many independent Brownian motions, $u(0) = u_0$, A: $A(t): V \to V'$, μ_k: $\mu_k(t): V \to H$,

$$\sum_k E \int_0^T \|\mu_k \varphi\|_H^2 \mathrm{d}t < \infty,$$

$f \in L^2(\Omega \times (0,T)); V')$, i.e.

$$E \int_0^T \|f(t)\|_{V'}^2 \mathrm{d}t < \infty,$$

$g_k \in L^2(\Omega \times (0,T); H)$ and

$$\sum_{k=1}^{\infty} E \int_0^T \|g_k(t)\|_H^2 \mathrm{d}t < \infty.$$

If A is $A(t,\omega)$, then $A(t)\varphi$ is $\mathcal{F}_t^W$-adapted, and likewise for μ_k.

$$\begin{aligned} Au &= a(t,x)u(t,x)_{xx}, \\ \mu_k u &= \sigma_k(t,x)u(t,x)_x, \\ V &= H^1(\mathbb{R}^d), \\ H &= L^2(\mathbb{R}^d), \\ V' &= H^{-1}(\mathbb{R}^d). \end{aligned}$$

Saying that $A(t)\varphi \in V'$ is $\mathcal{F}_t^W$-adapted means that $\forall \psi \in V$, $[A(t)\varphi, \psi]$ is an $\mathcal{F}_t^W$-adapted random variable. Consider *Pettis' Theorem*, which states that

> Suppose we have a measure space $(\Omega, \mathcal{F}, P)$. Suppose X and Y are Hilbert spaces. Then
>
> - $f(\omega): \Omega \to X$ is $\mathcal{F}$-measurable iff $\{\omega: f(\omega) \in A \subset X\} \in \mathcal{F}$
>
> is equivalent to
>
> - $(g, f(\omega))_X$ is $\mathcal{F}$-measurable for all $g \in \tilde{X}$ where $\tilde{X}$ is a dense subset of X.

u is a solution of (8.1) iff for all t

$$u(t) = u_0 + \int_0^t (Au(s) + f(s))\mathrm{d}s + \sum_k \int_0^t (\mu_k u(s) + g_k(s))\mathrm{d}W_k(s)$$

with probability 1 in V', that is

$$[u(t), \varphi] = [u_0, \varphi] + \int_0^t [Au(s) + f(s), \varphi]\mathrm{d}s + \sum_k \int_0^t [\mu_k u + g_k, \varphi]\mathrm{d}W_k(s).$$

If $u \in V$, we would have

$$(u(t), \varphi)_H = (u_0, \varphi)_H + \int_0^t [Au(s) + f(s), \varphi]\mathrm{d}s + \sum_k \int_0^t (\mu_k u + g_k, \varphi)\mathrm{d}W_k(s).$$

Theorem 8.1. *In addition to the assumptions we already made, assume*

(A1). $\exists \delta > 0$ *and* $C_0 \geqslant 0$*, so that*

$$\exists \delta > 0, C_0 \geqslant 0: 2[A\varphi(t), \varphi] + \sum_k \|\mu_k \varphi\|_H^2 \leqslant -\delta \|\varphi\|_V^2 + C_0 \|\varphi\|_H^2.$$

("coercivity condition" $\Leftrightarrow$ *superellipticity)*

(A2). $\|A\varphi\|_{V'} \leqslant C_A \|\varphi\|_V$.

Then there is existence and uniqueness for the above equations.

That means there is a $u \in L^2(\Omega: C([0,T]); H) \cap L^2(\Omega: C([0,T]); V)$, moreover

$$E \sup_{t \leqslant T} \|u(t)\|_H^2 + E \int_0^T \|u(t)\|_V^2 \mathrm{d}t \leqslant C E\left(\|u_0\|_H^2 + \int_0^T \|f\|_{V'}^2 \mathrm{d}t + \sum_k \int_0^T \|g_k\|_H^2 \mathrm{d}t \right)$$

Interpretation: If $H = L^2$, $V = H^1$, $u(t)$ is cont. in L^2 and has one derivative in x which is square-integrable. (We might have also used $H = H^1$ and $V = H^2$, in which case u is cont. in H^1 and has two derivatives which are square-integrable.)

Now consider the following fact leading up to the *energy equality:* Suppose we have a function $u(t) \in L^2(0,T)$ and a generalized derivative $u'(t) \in L^2(0,T) \Rightarrow u(t)$ is continuous on $[0,T]$ and

$$\begin{aligned} u(t) &= \int_0^T u'(s)\mathrm{d}s, \\ |u(t)|^2 &= 2\int_0^t u(s)u'(s)\mathrm{d}s. \end{aligned}$$

Proof: Homework.

In the infinite-dimensional setting, we have a very analogous statement:

Suppose $u(t) \in L^2([0,T];V)$ and $u'(t) \in L_2([0,T];V')$. Then $u(t) \in C([0,T];H)$ and

$$\|u(t)\|_H^2 = 2\int_0^t [u'(s), u(s)]\mathrm{d}s.$$

[Lectures 14-15 not typed, notes available from Prof. Rozovsky]

[April 10, 2007, Lototsky, Lecture 16]

9 SPDE with space-time white noise

$$\mathrm{d}u = u_{xx}\mathrm{d}t + g(u)\mathrm{d}W(t,x)$$

on $0 < x < \pi$ with

$$\begin{aligned} u|_{t=0} &= u_0, \\ u|_{x=0} = u|_{x=\pi} &= 0, \\ u_t|_{x=0} = u_t|_{x=\pi} &= 0. \end{aligned}$$

Two different ways of writing this equation are

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

or

$$\mathrm{d}u = u_{xx}\mathrm{d}t + \sum_{k=1}^{\infty} g(u)h_k\mathrm{d}W_k(t).$$

Theorem 9.1. (Walsh, Lecture Notes in Mathematics 1180, 1984)
If $u_0 \in C^\infty$, then $u \in C_t^{0,1/4-\varepsilon} \cap C_x^{0,1/2-\varepsilon}$.

Three kinds of space-time white noise:

- Brownian sheet — $W(t,x) = \mu([0,t] \times [0,x])$
- Cylindrical/Brownian motion — family of Gaussian random variables $B_t = B_t(h)$, $h \in H$ a Hilbert space, $E[B_t(h)] = 0$, $E[B_t(h)B_s(g)] = (h,g)_H\ (t \wedge s)$
- Space-time white noise $\mathrm{d}W(t,x) = \frac{\partial^2 W}{\partial t \partial x} = \sum_{k=1}^\infty h_k(x)\mathrm{d}W_k(t)$, where $\{h_k\}$ is assumed a Basis of the Hilbert space we're in — if $\{h_k, k \geqslant 1\}$ is a complete orthonormal system, then $\{B_t(h_k), k \geqslant 1\}$-independet standard Brownian motion.

Connection between the three: If $H = L^2(\mathbb{R})$ or $H = L^2(0,\pi)$, then

$$B_t(h) = \int \frac{\partial W}{\partial x} h(x)\mathrm{d}x,$$

and

$$B_t(x) = B_t(\chi_{[0,X]}) = \sum_{k=1}^\infty \int_0^x (h_k(y)\mathrm{dyW}_k(t)) = W(t,x)$$

9.1 A closer look

Consider $g(u) \equiv 1$.

$$\mathrm{d}u = u_{xx}\mathrm{d}t + \sum_{k=1}^{\infty} h_k(x)\mathrm{d}W_k(t),$$

where we assume that

$$h_k(x) = \sqrt{\frac{2}{\pi}}\sin(k\,x).$$

Observe that, strictly, the series on the RHS diverges in L^2. Now consider the setting of a Sobolev space

$$H^\gamma = H^\gamma((0,\pi)),$$

with

$$\|f\|_\gamma^2 = \sum_{k=1}^{\infty} k^{2\gamma} f_k^2, \quad f_k = \int_0^\pi f(x)h_k(x)\mathrm{d}x$$

for $\gamma \in \mathbb{R}$. Now consider

$$M(t,x) = \sum_{k=1}^{\infty} h_k(x)W_k(t) \in H^\gamma,$$

i.e.

$$E\|M\|_\gamma^2 = t\sum_{\gamma=1}^{\infty} k^{2\gamma} < \infty$$

if $\gamma < -1/2$.

$$u(t) = u_0 + \int_0^t A\,u\mathrm{d}s + M(t),$$

where

$$A = \frac{\partial^2}{\partial x^2}: H^{\gamma+1} \to H^{\gamma-1}.$$

Then

$$\exists! u \in L^2(\Omega; L^2(0,T); H^{\gamma+1}) \cap L^2(\Omega; C(0,T); H^\gamma)$$

for all $\gamma < -1/2$, so u is almost in $H^{1/2}$ for almost all t.

We assume a Fourier point of view, so that

$$u(t,x) = \sum_{k=1}^{\infty} u_k(t)h_k(x)$$

and

$$\mathrm{d}u_k = -k^2 u_k + \mathrm{d}W_k(t).$$

Then

$$u_k(t) = \int_0^t e^{-k^2(t-s)}\mathrm{d}W_k(s).$$

Next, note

> *Kolmogorov's criterion*: If
>
> $$E|X(x) - X(y)|^p < C|x-y|^{d+q}$$
>
> for $x \in \mathbb{R}^d$, then $X \in C^{0,q/p-\varepsilon}$ for all $\varepsilon > 0$.

Now, consider try to prove its assumption:

$$\begin{aligned} E|u(t,x) - u(t,y)|^p &= E\left|\sum_{k=1}^{\infty} u_k(t)(h_k(x) - h_k(y))\right|^p \\ &\overset{\text{BDG}}{\leqslant} C\left(\sum_{k=1}^{\infty} \frac{1}{2k^2}(1 - e^{-2k^2 t})|h_k(x) - h_k(y)|^2\right)^{p/2} \\ &\overset{(*)}{\leqslant} C|x-y|^{(1/2-\varepsilon)p}. \end{aligned}$$

where we've used the BDG (Burkholder/Davis/Gundy) Inequality, i.e.

$$E[M_T^p] \leqslant C\,E\langle M\rangle_T^{p/2},$$

where M is assumed a martingale, which we can achieve by fixing time t to T in the expression for u_k above. Next, note

$$E[u_k^2(t)] \;=\; \int_0^t e^{-2k^2(t-s)}\mathrm{d}s = \frac{1}{2k^2}(1-e^{-2k^2t}),$$

also quadration variation if we fix time as hinted above.

Once we get to $(*)$ above, realize that we *want*

$$\sum k^{2\delta-2} < \infty,$$

and usethe fact that

$$|h_k(x) - h_k(y)| \sim |\sin(k\,x) - \sin(k\,y)| \leqslant C(K|x-y|)^\delta$$

for $2\delta - 2 < -1$, i.e. $\delta < 1/2$, i.e. $\delta = 1/2 - \varepsilon$.

So altogether, we obtain $E|u(t,x) - u(t,y)|^p \leqslant C|x-y|^{(1/2-\varepsilon)p}$. Thus

$$u \in C_x^{1/2-\varepsilon-\frac{1}{p}-\varepsilon} = C_x^{1/2-\varepsilon}.$$

9.2 Mild solutions

Our u above is "a solution" to our SPDE, but not in the variational sense defined so far. So we need a more general idea of what a solution is, to subsume both concepts. If you have a general PDE

$$\dot{u} = A(t)U,$$

then $u(t) = \Phi_{t,0}u_0$. Then

$$\dot{u} = A(t)u + f(t)$$

gives us

$$u(t) = \Phi_{t,0}u_0 + \int_0^t \Phi_{t,s}f(s)\mathrm{d}s.$$

For example, if we have

$$\frac{\partial u}{\partial t} = u_{xx},$$

then

$$\Phi_{t,0}\colon f \mapsto \int_0^t G(t,x,y)f(y)\mathrm{d}y,$$

where *Greeen's function* is given by

$$G(t,x,y) = \sum_{k=1}^\infty e^{-k^2t}h_k(x)h_k(y)$$

if

$$\mathrm{d}u = u_{xx}\mathrm{d}t + \sum_k h_k(x)\mathrm{d}W_k, \quad u_0 = 0.$$

Then

$$u(t,x) = \sum_{k=1}^\infty \int_0^t \int_0^\pi G(t-s,x,y)h_k(y)\mathrm{d}y\mathrm{d}W_k(s).$$

Now for

$$\mathrm{d}u = u_{xx}\mathrm{d}t + \sum g(u)h_k\mathrm{d}W_k,$$

we write

$$u(t,x) \;=\; \int_0^\pi G(t,x,y)u_0(y)\mathrm{d}y + \sum_{k=1}^\infty \int_0^t \int_0^\pi G(t-s,x,y)g(u(y))h_k(y)\mathrm{d}y\mathrm{d}W_k(s).$$

Then you define a *mild solution* to be a solution to the above integral equation.

Now try

$$\begin{aligned} E|u(t,x_1)-u(t,x_2)|^p &\sim E\bigg|\sum_k \iint G(t-s,x_1,y)-G(t-s,x_2,y)h_k(y)g(u(s,y))\mathrm{d}y\mathrm{d}W_k(s)\bigg|^p \\ &\leqslant E\bigg(\sum_k \int_0^t \bigg|\int_0^\pi (G(t-s,x_1,y)-G(t-s,x_2,y))h_k(y)g\mathrm{d}y\bigg|^2\mathrm{d}s\bigg)^{p/2} \\ &= E\bigg(\int_0^t \int_0^\pi |G(t-s,x_1,y)-G(t-s,x_2,y)|^2 g^2(u(x,y))\mathrm{d}y\mathrm{d}s\bigg)^{p/2}. \end{aligned}$$

Then came Krylov (1996) and turned this “hard analysis” into clever “soft analysis” or so.