Introduction to the Theory of Stochastic Differential Equations and Stochastic Partial Differential Equations

Tadahisa Funaki

November 26 - December 10, 2010

Contents

1 Some Basic Concepts in Probability Theory **2**
1.1 Basic Concepts 2
1.2 Several Different Notions of Convergence of Random Variables 3
1.3 Independence 4
1.4 Stochastic Processes 6

2 Brownian Motion **6**
2.1 Definition and Construction 6
2.2 Properties of the 1-dimensional Brownian Motion 8

3 Stochastic Integrals and Itô's Formula **9**
3.1 Stochastic Integrals 9
3.2 Itô's Formula 12

4 Stochastic Differential Equations **15**
4.1 What is a Stochastic Differential Equation? 15
4.2 Applications to PDE and the Feynman-Kac Formula 17
4.3 Vanishing Viscosity Limit under Non-uniqueness for the Limiting ODE 19

5 Stochastic Partial Differential Equations **20**
5.1 Definition of a Suitable Noise 21
5.2 Concepts of Solutions 22
5.3 Regularity of Solutions 22
5.4 The KPZ Equation 23

After briefly summarizing the basic concepts and facts in probability theory such as probability spaces, random variables, their convergences, independence, the Central Limit Theorem and Gaussian distributions, the Brownian motion will be introduced. The stochastic integrals are essential to discuss stochastic (partial) differential equations and Itô's formula plays a central role in the calculus related to them.

After these preparations, the SDEs and SPDEs will be discussed, the latter one with an explicit example, the KPZ equation.

1 Some Basic Concepts in Probability Theory

Modern probability theory is based on the abstract version of measure theory. In particular, σ-additivity plays an important role. The additivity of probability, that is if events A and B never occur simultaneously, then $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$ is clear. This additivity must be extended to σ-additivity, for example, to formulate the Strong Law of Large Numbers as I will explain later.

The Strong Law of Large Numbers was formulated by Borel in 1909. This is, of course, related to the foundation of Lebesgue's integral. Then Fréchet extended the measure theory to an abstract setting motivated by the probability theory around 1915. Finally, Kolmogorov settled the base of modern probability theory in his book in 1933.

1.1 Basic Concepts

(i) $(\Omega, \mathscr{F}, \mathbb{P})$ is called a probability space, if Ω is a certain set, $\mathscr{F}$ a σ-field on Ω and $\mathbb{P}$ a measure on $(\Omega, \mathscr{F})$ satisfying $\mathbb{P}[\Omega] = 1$. $\mathbb{P}$ is called a probability measure.
Remark. To define the (d-dimensional) Brownian Motion, we first consider all possible sample paths, so that we take $\Omega = C([0, \infty), \mathbb{R}^d)$ and define $\mathscr{F}, \mathbb{P}$ properly as we will see. But in fact, the choice of Ω is quite flexible. One can even take the Lebesgue space $([0,1], \mathscr{B}([0,1]), dx)$ to realize a Brownian Motion.

(ii) $\omega \in \Omega$ is called a sample, $A \in \mathscr{F}$ is an event.

(iii) Let $(S, \mathscr{S})$ be a measurable space. $X : \Omega \to S$ is called an (S-valued) random variable if it is $\mathscr{F} - \mathscr{S}$-measurable. In case of S being a topological space, e. g. $S = \mathbb{R}^d$, we usually take $\mathscr{S} = \mathscr{B}(S)$, the Borel field of S.

(iv) For $A \in \mathscr{F}$, A a.s. means $\mathbb{P}[A] = 1$. For two random variables $X, Y : \Omega \to S$ we write $X = Y$ a.s. for $\mathbb{P}[X = Y] = 1$.
Remark. We will usually omit writing ω, since Ω is considered to be taken just for convenience to describe random phenomena in mathematical terminology. Therefore

$$\{X = Y\} := \{\omega \in \Omega : X(\omega) = Y(\omega)\} \in \mathscr{F}$$

or

$$\{X \in A\} := \{\omega \in \Omega : X(\omega) \in A\} \in \mathscr{F}.$$

(v) For an $\mathbb{R}$-valued random variable X such that $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$, i. e.

$$\int_\Omega |X(\omega)| d\mathbb{P}[\omega] < \infty,$$

we write

$$\mathbb{E}[X] := \int_\Omega X(\omega) d\mathbb{P}[\omega]$$

and call it the expectation (or mean) of X.

(vi) For an S-valued random variable X, the image measure of $\mathbb{P}$ under X

$$\mathbb{P}_X[A] := \mathbb{P}[X \in A]$$

is called the distribution (or probability law) of X.

1.2 Several Different Notions of Convergence of Random Variables

Let X and $(X_n)_{n\in\mathbb{N}}$ be $\mathbb{R}$-valued random variables defined on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

(i) $X_n \xrightarrow{n\to\infty} X$ $\mathbb{P}$-a.s. $\quad :\Longleftrightarrow \quad \mathbb{P}\big[\lim_{n\to\infty} X_n = X\big] = 1$, where

$$\{\lim_{n\to\infty} X_n = X\} = \{\omega \in \Omega : \exists \lim_{n\to\infty} X_n(\omega) \text{ and } \lim_{n\to\infty} X_n(\omega) = X(\omega)\}.$$

(ii) $X_n \xrightarrow{n\to\infty} X$ in probability $\quad :\Longleftrightarrow \quad \forall \varepsilon > 0 : \lim_{n\to\infty} \mathbb{P}\big[|X_n - X| > \varepsilon\big] = 0.$

(iii) $X_n \xrightarrow{n\to\infty} X$ in L^p, $p \geq 1 \quad :\Longleftrightarrow \quad \mathbb{E}\big[|X_n - X|^p\big] = \int_\Omega |X_n - X|^p d\mathbb{P}[\omega] \xrightarrow{n\to\infty} 0.$

(iv) $X_n \xrightarrow{n\to\infty} X$ in law $\quad :\Leftrightarrow \quad \mathbb{P}_{X_n} \to \mathbb{P}_X$ weakly (weak*), i. e.

$$\forall \varphi \in C_b(\mathbb{R}) : \lim_{n\to\infty} \mathbb{E}\left[\varphi(X_n)\right] = \mathbb{E}\left[\varphi(X)\right].$$

The following implications hold: (1) ⇒ (2), (3) ⇒ (2), (2) ⇒ (4), as well as (2) ⇒ (1) and (3) ⇒ (1) along some subsequence. Also, we have to mention, that in (4) only the distribution matters, therefore it is not necessary for X_n, X to be defined on the same probability space. This leads us directly to the

Theorem 1.1 (Skorohod's representation theorem). *Assume (4) holds. Then we can construct a proper probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ and random variables $\tilde{X}, (\tilde{X}_n)_{n\in\mathbb{N}}$ on this space in such a way that*

$$\tilde{\mathbb{P}}_{\tilde{X}_n} = \mathbb{P}_{X_n}, \quad \tilde{\mathbb{P}}_{\tilde{X}} = \mathbb{P}_X \quad \textit{and} \quad \tilde{X}_n \xrightarrow{n\to\infty} \tilde{X} \ \tilde{\mathbb{P}}\textit{-a. s.}$$

Remark.

$$\{\lim_{n\to\infty} X_n = X\} = \bigcap_{k\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} \bigcap_{m>n} \{|X_m - X| < \frac{1}{k}\} \in \mathscr{F}$$

Therefore the σ-additivity is necessary to define the notion of $\mathbb{P}$-a. s. convergence.

Definition 1.2. Let $(A_n)_{n\in\mathbb{N}} \subset \mathscr{F}$. We say "$A_n$ occurs infinitely often" if

$$\mathbb{P}\left[\limsup_{n\to\infty} A_n\right] = \mathbb{P}\left[\cap_{n\in\mathbb{N}} \cup_{m>n} A_m\right] = 1.$$

Proposition 1.3 (Borel-Cantelli). *Let $(A_n)_{n\in\mathbb{N}} \subset \mathscr{F}$. If $\sum_{n\in\mathbb{N}} \mathbb{P}\left[A_n\right] < \infty$ then*

$$\mathbb{P}\left[\limsup_{n\to\infty} A_n\right] = 0.$$

1.3 Independence

The concept of independence is one of the main components of probability theory. Related to this, we will consider different σ-fields on Ω at the same time, which is typical in probability theory.

Definition 1.4. We introduce the following definitions.

(i) Let $A, B \in \mathscr{F}$. Then A and B are independent iff $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]$.

(ii) Let $\{A_i\}_{1\le i\le n} \subset \mathscr{F}$ be a finite family of sets. Then $\{A_i\}_{1\le i\le n}$ are independent, iff for all $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ holds

$$\mathbb{P}\left[A_{i_1} \cap \dots A_{i_k}\right] = \mathbb{P}\left[A_{i_1}\right] \cdot \dots \cdot \mathbb{P}\left[A_{i_k}\right].$$

(iii) Let $\{A_\lambda\}_{\lambda\in\Lambda} \subset \mathscr{F}$ be a family of sets, with Λ possibly being uncountable. Then we say $\{A_\lambda\}_{\lambda\in\Lambda}$ are independent, iff for all finite sequences $\lambda_1, \dots, \lambda_n \in \Lambda$ the family of sets $\{A_{\lambda_k}\}_{1\le k\le n}$ is independent.

(iv) Let $\{\mathscr{F}_\lambda\}_{\lambda\in\Lambda}$ be a family of σ-fields on Ω, with Λ possibly being uncountable. Then we say $\{\mathscr{F}_\lambda\}_{\lambda\in\Lambda}$ are independent, iff for all finite sequences $\lambda_1, \dots, \lambda_n \in \Lambda$ and all $A_{\lambda_i} \in \mathscr{F}_{\lambda_i}$ the family of sets $\{A_{\lambda_k}\}_{1\le k\le n}$ is independent.

(v) Let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a family of random variables on a common probability space $(\Omega, \mathscr{F}, \mathbb{P})$, with Λ possibly being uncountable. Then we say $\{X_\lambda\}_{\lambda\in\Lambda}$ are independent, iff the generated σ-fields $\{\sigma\{X_\lambda\}\}_{\lambda\in\Lambda}$ are independent. Hereby

$$\sigma\{X\} := \{\{X \in A\} : A \in \mathscr{S}\}.$$

Proposition 1.5. *Let X, Y be two independent $\mathbb{R}$-valued random variables.*

(i) Let $f, g \in \mathscr{B}(\mathbb{R})$. Then $f(X)$ and $g(Y)$ are independent.

(ii) Let $X, Y \in L^1(\Omega, \mathscr{F}, \mathbb{P})$. Then $XY \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ and $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$.

Theorem 1.6 (Strong Law of Large Numbers/ Kolmogorov's second law)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{R}$-valued independent random variables defined on the same probability space with identical distributions (short iid). If $X_1 \in L^1(\Omega, \mathscr{F}, \mathbb{P})$, then the Strong Law of Large Numbers holds, i. e.*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{n\to\infty} m = \mathbb{E}\left[X_1\right] \quad \mathbb{P}\text{-a. s.}.$$

The next theorem explains why Gaussian random variables appear quite naturally in physics and statistics.

Theorem 1.7 (Central Limit Theorem (CLT))**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{R}^d$-valued iid random variables. We assume $\mathbb{E}\left[X_1\right] = 0$ and that all covariances exist and are finite, therefore we can define the symmetric $d \times d$-matrix*

$$V := \operatorname{Cov} X_1 = (\mathbb{E}\left[X_1^{(i)} X_1^{(j)}\right])_{1\le i,j\le d}.$$

Then we have a Central Limit Theorem, i e.

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k \xrightarrow{n\to\infty} X \sim \mathscr{N}(0, V) \quad \textit{weakly}.$$

Here $\mathcal{N}(0, V)$ is the Gaussian distribution on $\mathbb{R}^d$ with mean 0 and covariance V

$$\mu_{0,V}(dx) = \frac{1}{(2\pi)^{\frac{d}{2}}(\det V)^{\frac{1}{2}}} e^{-\frac{1}{2}xV^{-1}x} dx.$$

Remark. If $\det V = 0$, then $\mu_{0,V}$ is degenerate, e. g. in $d = 1$ that means $\mathcal{N}(0,0) = \delta_0$.

For the proof of the CLT the characteristic function of a random variable X is used. It is the Fourier transform of the distribution, namely

$$\psi_X(u) := \mathbb{E}\left[e^{iuX}\right], u \in \mathbb{R}^d.$$

For $X \sim \mathcal{N}(0, V)$ we have $\psi_X(u) := e^{-\frac{1}{2}uVu}$.

1.4 Stochastic Processes

Definition 1.8. A collection of $\mathbb{R}^d$-valued random variables $X = (X_t)_{t\geq 0}$ indexed by $t \in [0, \infty)$ is called a (continuous time) stochastic process.

X is called continuous, if $\mathbb{P}\left[\{X_.(\omega) \text{ is continuous in } t\}\right] = 1$.

2 Brownian Motion

The random motion of a particle in $\mathbb{R}^d$, which refreshes the memory of its past at every time and has continuous trajectories, is called the Brownian Motion.

2.1 Definition and Construction

Definition 2.1. An $\mathbb{R}$-valued stochastic process $B = (B_t)_{t\geq 0}$ defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is called a Brownian Motion, iff

(i) $B_0 = 0$ $\mathbb{P}$-a. s.,

(ii) B is continuous in time and

(iii) for all $0 = t_0 < t_1 < \cdots < t_n$ the increments $(B_{t_i} - B_{t_{i_1}})_{1\leq i\leq n}$ are independent and $B_t - B_s \sim \mathcal{N}(0, t-s)$ for $0 \leq s \leq t$.

The property (iii) is equivalent to

$$\mathbb{P}\Big[B_{t_i} - B_{t_{i-1}} \in A_i, 1 \le i \le n\Big] = \int_{A_1} dx_1 \cdots \int_{A_n} dx_n \prod_{i=1}^{n} p(t_i - t_{i-1}, x_i) \tag{2.1}$$

for $A_1, \dots, A_n \in \mathscr{B}(R)$, where

$$p(t,x) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}, \quad t > 0, x \in \mathbb{R}$$

is the heat kernel (Gaussian density) on $\mathbb{R}$. By a change of variables $x_i = y_i - y_{i-1}$ with $y_0 = 0$, the equation (2.1) is equivalent to

$$\mathbb{P}\Big[B_{t_i} \in C_i, 1 \le i \le n\Big] = \int_{C_1} dy_1 \cdots \int_{C_n} dy_n \prod_{i=1}^{n} p(t_i - t_{i-1}, y_i - y_{i-1}) \tag{2.2}$$

for $C_1, \dots, C_n \in \mathscr{B}(R)$. This is sometimes written as

$$\mathbb{P}\Big[B_{t_i} \in dy_i, 1 \le i \le n\Big] = \prod_{i=1}^{n} p(t_i - t_{i-1}, y_i - y_{i-1}) dy_i, \quad y_0 = 0.$$

Theorem 2.2. *There exists a Brownian Motion on a certain probability space.*

There are several proofs for this theorem, involving

(i) Kolmogorov's extension theorem together with (2.2),

(ii) a Fourier series expansion,

(iii) Invariance principles and approximation by a random walk and

(iv) existence of Gaussian systems.

In particular, one can construct the Brownian Motion on the space $W := C([0,\infty), \mathbb{R})$, resp. $W_0 := \{w \in W : w(0) = 0\}$. The distribution of B on W_0 is called the Wiener measure.

In the same way, one can talk of a d-dimensional Brownian Motion, which is an $\mathbb{R}^d$-valued stochastic process $B = ((B_t^{(1)})_{t\ge 0}, \dots (B_t^{(d)})_{t\ge 0})$ where $B^{(i)}$ are independent 1-dimensional Brownian Motions. In this case (2.2) holds with

$$p(t,x) = \frac{1}{(2\pi t)^{\frac{d}{2}}} e^{-\frac{|x|^2}{2t}}, \quad t \ge 0, x \in \mathbb{R}^d.$$

2.2 Properties of the 1-dimensional Brownian Motion

Corollary 2.3 (Moments and Covariance). *For the* 1*-dimensional Brownian Motion holds*

$$\mathbb{E}\left[B_t^{2n-1}\right] = 0, n \in \mathbb{N}, \quad \mathbb{E}\left[B_t^2\right] = t, \quad \mathbb{E}\left[B_t^4\right] = 3t^2, \dots$$

and

$$\mathbb{E}\left[B_t B_s\right] = t \wedge s (= \min\{t, s\}).$$

Proposition 2.4 (Scaling Invariance). *Let B be a* 1*-dimensional Brownian Motion and $\alpha > 0$. Then the scaled process $(\alpha B_{t\alpha^{-2}})_{t\geq 0}$ is again a Brownian Motion.*

Next we want to focus on path properties of the Brownian Motion. One can show, that the total variation of B on every interval $[T_1, T_2]$ is $\mathbb{P}$-a. s. infinite for all $0 \leq T_1 < T_2$. Therefore we consider the quadratic variation on the interval $[0, 1]$. Let $t_k = \frac{k}{n}, 1 \leq k \leq n$ be a partition of $[0, 1]$ and X_n be the sum of the squared increments

$$X_n := \sum_{k=1}^{n} (B_{t_k} - B_{t_{k-1}})^2.$$

Then we compute the L^2-distance

$$\mathbb{E}\left[(X_n - 1)^2\right] = \mathbb{E}\left[(\sum_{k=1}^{n} (B_{t_k} - B_{t_{k-1}})^2 - \frac{1}{n})^2\right] = \sum_{k=1}^{n} \mathbb{E}\left[((B_{t_k} - B_{t_{k-1}})^2 - \frac{1}{n})^2\right] = \frac{2}{n} \to 0,$$

where we used the independence of the increments and the moment calculations in Corollary 2.3.

From the fact, that the quadratic variation is finite we can easily deduce that the total variation has to be infinite. We can indeed estimate

$$\sum_{k=1}^{n} (B_{t_k} - B_{t_{k-1}})^2 \leq V_0^1(B) \cdot \sup_k |B_{t_k} - B_{t_{k-1}}|$$

and because the last term tends to 0, every function with finite total variation $V_0^1(B)$ has quadratic variation 0. Therefore the Brownian Motion cannot have finite total variation.

The computation on the quadratic variation suggests that B_t is $\frac{1}{2}$-Hölder continuous in t $\mathbb{P}$-a. s. In fact, though this is not really true, B_t is $(\frac{1}{2} - \delta)$-Hölder continuous in t $\mathbb{P}$-a. s. for all $\delta \in (0, 1/2)$.

Let $\mathscr{F}_t^B$ be the σ-field generated by $\{B_s; s \leq t\}$ and null sets of $\mathbb{P}$. Then, $\left(\mathscr{F}_t^B\right)_{t\geq 0}$ is a reference family (i.e. $\mathscr{F}_t^B$ are sub σ-fields of $\mathscr{F}$ such that $\mathscr{F}_s^B \subset \mathscr{F}_t^B$ for all $0 \leq s < t$), B is $\left(\mathscr{F}_t^B\right)_{t\geq 0}$-adapted (i.e. B_t is $\mathscr{F}_t^B$-measurable for all $t \geq 0$) and the increment $B_t - B_s$ is independent of $\mathscr{F}_s^B$ for all $0 \leq s \leq t$.

3 Stochastic Integrals and Itô's Formula

3.1 Stochastic Integrals

We now develop the calculus based on the Brownian Motion, usually called stochastic calculus or Itô's calculus. Since the derivative $\dot{B}_t = \frac{d}{dt}B_t$ does not exist, we want to give a meaning to the integral

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

for a certain class of functions $f(s,\omega)$. Since B also does not have finite variation, such integrals cannot be defined as usual Stieltjes integrals.

Example 3.1. Before defining the stochastic integrals in general, let us consider the integral

$$\int_0^1 B_s dB_s$$

defined as followed. We take again the partition $t_k = \frac{k}{n}, 1 \leq k \leq n$ of $[0,1]$ and look at the Riemann sums

$$S_n := \sum_{k=1}^{n} B_{s_k}(B_{t_k} - B_{t_{k-1}}), \quad s_k \in [t_{k-1}.t_k].$$

For the usual Stieltjes integrals, $\lim_{n\to\infty} S_n$ exists and does not depend on the choice of s_k. However, for the stochastic integral this is not the case. If we set $\underline{S}_n$ as the sum with $s_k = t_{k-1}$ (evaluation at the left edges) and $\overline{S}_n$ as the sum with $s_k = t_k$ (evaluation at the right edges), we get

$$\overline{S}_n - \underline{S}_n = \sum_{k=1}^{n}(B_{t_k} - B_{t_{k-1}})^2 \to 1 \text{ in } L^2.$$

Thus, the limit of the stochastic integral is supposed to depend on the choice of s_k. This is due to the fact, that the Brownian Motion has only finite quadratic variation. We will see, that the evaluation at the left edge $s_k = t_{k-1}$ is convenient for our purposes.

In the following fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a reference family of σ-fields $\left(\mathscr{F}_t\right)_{t\geq 0}$, with $\mathscr{F}_s \subset \mathscr{F}_t \subset \mathscr{F}$ for all $0 \leq s < t$. Furthermore let B be an $\left(\mathscr{F}_t\right)_{t\geq 0}$-Brownian Motion, i. e.

B is a Brownian Motion and $(\mathscr{F}_t)_{t\geq 0}$-adapted (the random variable B_t is $\mathscr{F}_t$-measurable for each t) and for all $0 \leq s \leq t$ the increment $B_t - B_s$ is independent of $\mathscr{F}_s$.

Fix $T > 0$ to define the class of possible integrands we will look at.

$$\mathscr{L}_T^2 = \mathscr{L}_T^2(\mathscr{F}_t)_{t\geq 0} = \{f \in L^2([0,T] \times \Omega) : f \text{ is } (\mathscr{F}_t)_{t\geq 0}\text{-adapted}\}$$

Lemma 3.2 (Approximation of $f \in \mathscr{L}_T^2$)**.** *Consider a step process, i. e. a process in*

$$\mathscr{S}_T = \{f : f(t,\omega) = \sum_{k=1}^n f_k(\omega)\chi_{[t_{k-1},t_k)}(t), t \in [0,T], 0 = t_0 < t_1 < \cdots < t_n = T,$$
$$n \in \mathbb{N} \text{ and } f_k \text{ is } \mathscr{F}_{t_{k-1}}\text{-measurable and bounded in } \omega\}.$$

Then for every $f \in \mathscr{L}_T^2$ there exists a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathscr{S}_T$ such that

$$\|f - f_n\|_{L^2([0,T]\times\Omega)} \xrightarrow{n\to\infty} 0.$$

Idea of Proof First consider a cutoff by $f_m(t,\omega) := (f(t,\omega) \wedge m) \vee -m$. Then it clearly holds $\|f - f_m\|_{L^2([0,T]\times\Omega)} \to 0$. So we can assume f is bounded.

By smearing the function in t via $f^\varepsilon(t,\omega) := \varepsilon^{-1}\int_{(t-\varepsilon)\vee 0}^t f(s,\omega)ds$ we can assume, that f is continuous in time. This approximation maintains the adaptedness to $(\mathscr{F}_t)_{t\geq 0}$.

Finally we approximate the continuous in time function f by

$$f_n(t,\omega) = \sum_{k=1}^n f_(t_{k-1},\omega)\chi_{[t_{k-1},t_k)}(t)$$

with $t_k = \frac{k}{n}T$. By taking the evaluation at the left edge we maintain the adaptedness.

□

Definition 3.3. For every function $f \in \mathscr{S}_T$ we define the stochastic Integral as

$$I_t(f) = \int_0^t f(s)dB_s := \sum_{k=1}^n f_k(B_{t\wedge t_k} - B_{t\wedge t_{k-1}}), t \in [0,T].$$

Lemma 3.4. *For $f \in \mathscr{S}_T$ holds*

(i) $\mathbb{E}\left[I_t(f)\right] = 0$ *and*

(ii) $\mathbb{E}\left[I_t(f)^2\right] = \mathbb{E}\left[\int_0^t f(s)^2 ds\right] = \|f\|^2_{L^2([0,t]\times\Omega)}$.

Proof For the first claim observe, that since f_k is $\mathscr{F}_{t_{k-1}}$-measurable and therefore independent of the increment $(B_{t\wedge t_k} - B_{t\wedge t_{k-1}})$ we can use Proposition 1.5 and Corollary 2.3 and get the desired result.

For the second claim we use the same properties, which yield first that

$$\mathbb{E}\left[I_t(f)^2\right] = \mathbb{E}\left[\sum_{k=1}^{n} f_k^2 (B_{t\wedge t_k} - B_{t\wedge t_{k-1}})^2\right],$$

since all off-diagonal terms have expectation 0. By exploiting again the independence, we see that

$$= \sum_{k=1}^{n} \mathbb{E}\left[f_k^2\right] \mathbb{E}\left[(B_{t\wedge t_k} - B_{t\wedge t_{k-1}})^2\right] = \mathbb{E}\left[\sum_{k=1}^{n} f_k^2((t\wedge t_k) - (t\wedge t_{k-1}))\right].$$

□

Now we can construct the stochastic integral for general integrand $f \in \mathscr{L}_T^2$. We take a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathscr{S}_T$ with $f_n \to f$ in $L^2([0,T]\times\Omega)$ and consider for $0 < t < T$ the sequence of stochastic integrals $(I_t(f_n))_{n\in\mathbb{N}}$. By Lemma 3.4 this is a Cauchy sequence in $L^2([0,T]\times\Omega)$ and we call its limit $I_t(f) = \int_0^t f(s)dB_s$, the stochastic integral of f with respect to B.

One can extend the results in Lemma 3.4 to all $f \in \mathscr{L}_T^2$ and obtain the so called Itô isometry:

$$\mathbb{E}\left[I_t(f)^2\right] = \mathbb{E}\left[\int_0^t f(s)^2 ds\right] \text{ and } \mathbb{E}\left[I_t(f)I_t(g)\right] = \mathbb{E}\left[\int_0^t f(s)g(s)ds\right].$$

As a matter of fact, based on the theory of martingales one can show a uniform convergence in time

$$\mathbb{E}\left[\sup_{0<t<T} (I_t(f_m) - I_t(f_n))^2\right] \to 0$$

as $m,n\to\infty$. This shows that the stochastic integral $I_t(f)$ is a continuous stochastic process.

Since $T > 0$ is arbitrary, we can define the stochastic integral for all $f \in \mathscr{L}^2$

$$\mathscr{L}^2 := \{f : f|_{[0,T]} \in \mathscr{L}_T^2 \text{ for all } T > 0\}.$$

In order to deal with d-dimensional processes in the next subsection, we need the next lemma.

Lemma 3.5. *Let $B = (B^{(1)}, \dots, B^{(d)})$ be a d-dimensional Brownian Motion. Then*

$$\mathbb{E}\left[\int_0^t f(s)dB_s^{(i)} \int_0^t g(s)dB_s^{(j)}\right] = \delta_{ij}\mathbb{E}\left[\int_0^t f(s)g(s)ds\right].$$

By using the independence of the different components of the d-dimensional Brownian Motion the result follows easily from the same computations as above.

3.2 Itô's Formula

As a motivation for the following we first consider $f, g \in C^1(\mathbb{R})$. Then

$$\frac{d}{dt}f(g(t)) = f'(g(t))\frac{dg}{dt}$$

and of course in the integrated form

$$f(g(t)) - f(g(0)) = \int_0^t f'(g(s))dg(s).$$

We want to introduce a similar result for the stochastic integral, i. e. $g(t) = B_t$. As we will see later, a correction term of second order appears

$$f(B_t) - f(B_0) = \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds. \tag{3.3}$$

Example 3.6. We come back to the example we discussed before:

$$\int_0^1 B_s dB_s = \lim_{n\to\infty} \underline{S}_n = \lim_{n\to\infty} \sum_{k=1}^n B_{t_{k-1}}(B_{t_k} - B_{t_{k-1}})$$

with $t_k = \frac{k}{n}$. Using the identity $2B_{t_{k-1}} = (B_{t_k} + B_{t_{k-1}}) - (B_{t_k} - B_{t_{k-1}})$ yields

$$= \lim_{n\to\infty} \frac{1}{2}\left[\sum_{k=1}^n (B_{t_k}^2 - B_{t_{k-1}}^2) - \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2\right] = \frac{1}{2}(B_1^2 - B_0^2) - \frac{1}{2}.$$

Thus we have shown, that for $f(x)=x^2$, therefore $f'(x)=2x$ and $f''(x)=2$

$$B_1^2 = f(B_1) = 2\int_0^t B_s dB_s + 1 = f(B_0) + \int_0^1 f'(B_s)dB_s + \frac{1}{2}\int_0^1 f''(B_s)ds.$$

We see, that the correction term actually appears with our choice of the stochastic integral. If we choose a different evaluation point, say $s_k = \frac{1}{2}(t_{k-1}+t_k)$ the middle point of the division times, and define the stochastic integral by

$$\int_0^1 B_s \circ dB_s := \lim_{n\to\infty} S_n := \lim_{n\to\infty} \sum_{k=1}^n B_{s_k}(B_{t_k} - B_{t_{k-1}})$$

this correction term will not appear. In detail we calculate S_n

$$S_n = \sum_{k=1}^n B_{s_k}(B_{t_k} - B_{s_k} + B_{s_k} - B_{t_{k-1}}) = \sum_{k=1}^n B_{s_k}(B_{t_k} - B_{s_k}) + \sum_{k=1}^n B_{s_k}(B_{s_k} - B_{t_{k-1}})$$

and with the same trick as above

$$\begin{aligned} &= \frac{1}{2}\left[\sum_{k=1}^n (B_{t_k}^2 - B_{s_k}^2) + \sum_{k=1}^n (B_{s_k}^2 - B_{t_{k-1}}^2) - \sum_{k=1}^n (B_{t_k} - B_{s_k})^2 + \sum_{k=1}^n (B_{s_k} - B_{t_{k-1}})^2\right] \\ &\quad = \frac{1}{2}\left[(B_1^2 - B_0^2 - \sum_{k=1}^n (B_{t_k} - B_{s_k})^2 + \sum_{k=1}^n (B_{s_k} - B_{t_{k-1}})^2\right] \to \frac{1}{2}B_1^2. \end{aligned}$$

Thus we see that $B_1^2 = \int_0^1 B_s \circ dB_s$ without the correction term.

We now want to state the Itô's formula, a general form of (3.3). Therefore we introduce the $\mathbb{R}^d$-valued stochastic process $X=(X^1,\dots,X^d)$ given by

$$X_t^i = X_0^i + \sum_{k=1}^N \int_0^t a_k^i(s,\omega)dB_s^k + \int_0^t b^i(s,\omega)ds, \quad 1\le i\le d, \tag{3.4}$$

where X_0^i are $\mathscr{F}_0$-measurable random variables, B an N-dimensional $(\mathscr{F}_t)_{t\ge 0}$-Brownian Motion, $a_k^i \in \mathscr{L}^2$ and b^i bounded and $(\mathscr{F}_t)_{t\ge 0}$-adapted.

Theorem 3.7 (Itô's formula)**.** *Let $f \in C_b^2(\mathbb{R}^d)$ and X given by* (3.4)*. Then*

$$
\begin{aligned}
f(X_t) = f(X_0) + \sum_{i=1}^{d}\sum_{k=1}^{N}\int_0^t \frac{\partial f}{\partial x^i}(X_s)a_k^i(s)dB_s^k + \sum_{i=1}^{d}\int_0^t \frac{\partial f}{\partial x^i}(X_s)b^i(s)ds \\
+ \frac{1}{2}\sum_{i,j=1}^{d}\sum_{k=1}^{N}\int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)a_k^i(s)a_k^j(s)ds, \quad t \geq 0, \quad \mathbb{P}\text{-a. s.}.
\end{aligned} \tag{3.5}
$$

There are several different proofs for this theorem, one can start considering $a_k^i \in \mathscr{S}_T$ and taking the limit afterwards. We omit a detailed proof, instead we give an efficient interpretation of this formula.

Remark. (i) Usually equations (3.4) and (3.5) are written in terms of stochastic differentials

$$dX_t^i = a_k^i(t)dB_t^k + b^i(t)dt,$$

using Einstein's convention. Note that this expression has only a formal meaning and mathematically we always need to go back to the integrated forms. Then, under the usual calculus for differentiable functions, we have

$$df(X_t) = \frac{\partial f}{\partial x^i}(X_t)dX_t^i$$

but in the stochastic calculus, we have

$$df(X_t) = \frac{\partial f}{\partial x^i}(X_t)dX_t^i + \frac{1}{2}\frac{\partial^2 f}{\partial x^i \partial x^j}(X_t)dX_t^i dX_t^j,$$

i. e. we take account of the second order term in the Taylor expansion. To compute $dX_t^i dX_t^j$, we apply the relations:

$$dB_t^i dB_t^j = \delta_{ij}dt \quad \text{and} \quad dB_t^i dt = (dt)^2 = 0.$$

The first relation is related to Lemma 3.5 and we may think of $dB_t \sim \sqrt{t}$, since B has almost $\frac{1}{2}$-Hölder continuous paths. This relation gives us

$$dX_t^i dX_t^j = \sum_{k=1}^{N} a_k^i(t)a_k^j(t)dt$$

and thus, we obtain Itô's formula.

(ii) As the last example suggests, if we use another definition for the stochastic integral, namely the Stratonovich stochastic integral

$$\int_0^t f(s) \circ dB_s$$

the correction term in Itô's formula does not appear. However, the class of possible integrands is more restrictive than the one for the Itô integral.

4 Stochastic Differential Equations

4.1 What is a Stochastic Differential Equation?

As a first example we consider an ordinary differential equation on $\mathbb{R}^d$ given by a vector field $b : \mathbb{R}^d \to \mathbb{R}^d$ and

$$\frac{dX_t}{dt} = b(X_t). \tag{4.6}$$

We now want to get into the situation, that each time t is independently perturbed by a noise, so that we formally consider

$$\frac{dX_t}{dt} = b(X_t) + \dot{B}_t,$$

where $\dot{B}_t = \frac{dB_t}{dt}$ with B being a d-dimensional Brownian Motion is seen as a noise independent in time (as a limit of independent increments), though $\dot{B}_t$ does not exist. In addition to that, the noise may depend on the space variable by changing the equation to

$$\frac{dX_t}{dt} = b(X_t) + \sigma(X_t)\dot{B}_t,$$

where $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times N}$ and B is an N-dimensional $\left(\mathscr{F}_t\right)_{t\geq 0}$-Brownian Motion.

More precisely this equation is written in terms of stochastic differentials

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \tag{4.7}$$

i. e.

$$dX_t^i = b^i(X_t)dt + \sigma_k^i(X_t)dB_t^k, \quad 1 \leq i \leq d.$$

Definition 4.1. We call an $\mathbb{R}^d$-valued continuous $\left(\mathscr{F}_t\right)_{t\geq 0}$-adapted stochastic process $X = (X_t)_{t\geq 0}$ defined on $(\Omega, \mathscr{F}, \mathbb{P})$ a solution of the stochastic differential equation (4.7) starting in $x \in \mathbb{R}^d$, if it satisfies the stochastic integral equation

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s, \quad t \geq 0, \quad \mathbb{P}\text{-a. s..} \tag{4.8}$$

Remark. We always consider the situation $\sigma^i_k(X_\cdot) \in \mathscr{L}^2(\mathscr{F}_t)_{t\geq 0}$ and $b(X_\cdot) \in L^1_{\mathrm{loc}}([0,\infty))$ $\mathbb{P}$-a. s., so that the integrals are well defined.

It is well known that if the coefficients of an ordinary differential equation are Lipschitz continuous, then it has a unique solution. This can be extended to the stochastic differential equations. We introduce the notation

$$\|\sigma\|^2 = \sum_{i,k} (\sigma^i_k)^2 \quad \text{and} \quad |b|^2 = \sum_i (b^i)^2.$$

Theorem 4.2. *Assume that for the coefficients of* (4.7) *exists a constant $K > 0$ such that*

$$\|\sigma(x) - \sigma(y)\| + |b(x) - b(y)| \leq K|x - y| \quad \textit{for all } x, y \in \mathbb{R}^d.$$

Then for every $x \in \mathbb{R}^d$ there exists a solution X of (4.7) *starting in x and the following pathwise uniqueness holds: If X, X' are two solutions starting in x, then*

$$\mathbb{P}\Big[X_t = X'_t, \textit{ for all } t \geq 0\Big] = 1.$$

Remark. The Lipschitz continuity implies the linear growth condition on σ and b:

$$\|\sigma(x)\| + |b(x)| \leq K'(|x| + 1).$$

Proof As in the deterministic setting, the usual method of Picard's successive approximations can be applied. Set $X^{(1)}_t = x$ and

$$X^{(n)}_t = x + \int_0^t b(X^{(n-1)}_s)ds + \int_0^t \sigma(X^{(n-1)}_s)dB_s \quad n \geq 2.$$

Then one can show

$$\mathbb{E}\left[\sup_{0\leq s\leq t} |X^{(n)}_s - X^{(n-1)}_s|^2\right] \leq C \int_0^t \mathbb{E}\Big[|X^{(n-1)}_s - X^{(n-2)}_s|^2\Big]\, ds,$$

which can be used to conclude existence via a fixed point theorem. The details are omitted. □

4.2 Applications to PDE and the Feynman-Kac Formula

In this section, we want to point out the relation between stochastic differential equations and parabolic resp. elliptic partial differential equations. So in the following let X be the unique solution to (4.7). For $f \in C_b^2(\mathbb{R}^d)$ define the differential operator

$$Lf(x) := \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}(x)\frac{\partial^2 f}{\partial x^i \partial x^j}(x) + \sum_{i-1}^{d} b^i(x)\frac{\partial f}{\partial x^i},$$

where $a = (a^{ij})_{1\le i,j\le d} = \sigma\sigma^T$.

Proposition 4.3. *If the distribution of X_t has a density $u(t,x)$ with respect to the Lebesgue measure dx in $\mathbb{R}^d$, i. e.*

$$\mathbb{P}\left[X_t \in dx\right] = u(t,x)dx,$$

then $u(t,x)$ satisfies Kolmogorov's forward equation

$$\begin{aligned}\frac{\partial u}{\partial t} &= L^* u(t,x)\\ &= \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial x^i \partial x^j}(a^{ij}(x)u(t,x)) - \sum_{i=1}^{d}\frac{\partial}{\partial x^i}(b^i(x)u(t,x)), \quad t \ge 0, x \in \mathbb{R}^d \end{aligned} \tag{4.9}$$

in a weak sense.

Proof The origin of the Kolmogorov's forward equation can be seen in the Itô's formula. Since for a test function $f \in C_c^2(\mathbb{R}^d)$ it holds

$$f(X_t) = f(x) + \int_0^t Lf(X_s)ds + \text{stochastic integral.}$$

By taking the expectation the stochastic integral vanishes and we get

$$\mathbb{E}\left[f(X_t)\right] = f(x) + \int_0^t \mathbb{E}\left[Lf(X_s)\right] ds.$$

Since $\mathbb{E}\left[f(X_t)\right] = \int_{\mathbb{R}^d} u(t,x)f(x)dx$ holds, we derive the weak form of Kolmogorov's forward equation

$$\frac{d}{dt}\langle u(t), f\rangle = \langle u(t), Lf\rangle.$$

□

As a next step we want to introduce a similar result, without the assumption that X admits a density w. r. t. the dx. Fix functions $\varphi, V, g \in C_b(\mathbb{R}^d)$ and consider the parabolic equation

$$\begin{aligned} \frac{\partial}{\partial t}u(t) &= Lu(t) + Vu(t) + g, \quad t > 0, \text{ on } \mathbb{R}^d \\ u(0) &= \varphi \end{aligned} \tag{4.10}$$

Proposition 4.4 (Feynman-Kac formula). *Assume that the parabolic equation* (4.10) *has a smooth solution* $u \in C^{1,2}((0,\infty)\times\mathbb{R}^d)\cap C([0,\infty)\times\mathbb{R}^d)$ *growing at most polynomially in* x*, i. e.*

$$\forall T > 0\ \exists C_T > 0, p \geq 1 : |u(t,x)| \leq C_T(1+|x|^p) \text{ for all } t \in [0,T], x \in \mathbb{R}^d.$$

Then we have the representation

$$u(t,x) = \mathbb{E}\left[\varphi(X(t,x))e^{\int_0^t V(X(s,x))ds} + \int_0^t g(X(s,x))e^{\int_0^s V(X(r,x))dr}ds\right],$$

where $X(t,x)$ *is the solution to* (4.7) *at time* t *starting in* x.

Proof Fix $T > 0$ and set

$$M_t := u(T-t, X(t,x))e^{\int_0^t V(X(s,x))ds} + \int_0^t g(X(s,x))e^{\int_0^s V(X(r,x))dr}ds.$$

We can apply Itô's formula because of the polynomial growth of u and get

$$\begin{aligned} dM_t = {} & \frac{\partial u}{\partial x^i}(T-t,X(t,x))\sigma_k^i(X(t,x))e^{\int_0^t V(X(s,x))ds}dB_t^k + \Big(-\frac{\partial u}{\partial t}(T-t,X(t,x)) \\ & + Lu(T-t,X(t,x)) + V(X(t,x))u(T-t,X(t,x)) + g(X(t,x))\Big)e^{\int_0^t V(X(s,x))ds}dt. \end{aligned}$$

Since u satisfies the equation (4.10) the term in parentheses vanishes, thus $\mathbb{E}[M_T] = M_0 = u(T,x)$, what was to prove.

□

After this short paragraph about parabolic equations we want to turn our attention to the Dirichlet problem for an elliptic equation. Therefore let $D \subset \mathbb{R}^d$ be a bounded, open set with smooth boundary ∂D. $\varphi \in C(\partial D)$ is the given Dirichlet boundary data and we define the stopping times

$$\tau(x) := \inf\{t \geq 0 : X(t,x) \in \partial D\},$$

where $X(t,x)$ is the solution to (4.7) at time t starting in x. You can interpret this as the first hitting time for the boundary ∂D.

Proposition 4.5. *Assume* $\mathbb{P}[\tau(x) < \infty] = 1$ *for all* $x \in D$ *and the elliptic equation*

$$\begin{aligned} Lu + Vu + g = 0, \quad x \in D \\ u = \varphi, \quad x \in \partial D \end{aligned} \tag{4.11}$$

has a smooth solution u*. Then*

$$u(x) = \mathbb{E}\left[\varphi(X(\tau, x)) e^{\int_0^\tau V(X(s,x))ds} + \int_0^\tau g(X(s,x)) e^{\int_0^s V(X(r,x))dr} ds \right], \quad x \in \overline{D},$$

where $\tau = \tau(x)$*. In particular the solution to* (4.11) *is unique.*

Proof We use a similar argument as above, namely we define

$$M_t := u(X(t,x)) e^{\int_0^t V(X(s,x))ds} + \int_0^t g(X(s,x)) e^{\int_0^s V(X(r,x))dr} ds, \quad 0 \le t < \tau.$$

Then we can show that, computing dM_t, only the stochastic integral remains, which shows M_t is a martingale. By Doob's optimal sampling theorem for martingales

$$\mathbb{E}\left[M_{\tau(x)} \right] = M_0 = u(x).$$

□

4.3 Vanishing Viscosity Limit under Non-uniqueness for the Limiting ODE

Even though the uniqueness of solutions for the ODE (4.6) does not hold, the uniqueness (in law sense) holds for the SDE (4.7) if the diffusion coefficient $\sigma(x)$ is non-degenerate. Let us consider the SDE on $\mathbb{R}$:

$$dX_t^\epsilon = b(X_t^\epsilon) dt + \epsilon dB_t, \tag{4.12}$$

where $b(x) = \mathrm{sgn} x \cdot \sqrt{|x|}$ and $\epsilon > 0$. It is well-known that the ODE (with $\epsilon = 0$) admits several solutions: $X_t = (t - t_0)^2/4$ and $-(t - t_0)^2/4$ with arbitrary $t_0 \ge 0$.

Proposition 4.6. *The solution* $X_t^\epsilon, t \in [0, T]$ *of the SDE* (4.12) *starting at* 0 *(i.e.* $X_0^\epsilon = 0$*) weakly converges to*

$$X_t = \begin{cases} t^2/4, & \text{with probability } 1/2, \\ -t^2/4, & \text{with probability } 1/2, \end{cases}$$

as $\epsilon \downarrow 0$.

The proof goes as follows. First note that the corresponding large deviation rate function is given by

$$I(x) = \frac{1}{2}\int_0^T (\dot{x}_t - b(x_t))^2 dt,$$

for a path $x = (x_t)_{t\in[0,T]}$, see Funaki ('05, p.121). Thus the limits of X^ϵ as $\epsilon \downarrow 0$ must concentrate on the set of minimizers of I, that is, the paths $X_t = \pm(t-t_0)^2/4$ with some $t_0 \in [0, T]$. However, one can compute the expectation $E[\tau^\epsilon]$ of the hitting time τ^ϵ of X^ϵ to $\{\pm 1\}$ (this is well-known in probability theory, but see, for example, Karatzas-Shreve ('91, p.344)) and this shows that t_0 should be 0 in the limit. The symmetry proves the proposition. (I thank Tokuzo Shiga for helpful discussions.) This proposition tells that the random noise forces the particle starting at the origin to leave there immediately. The vanishing viscosity limit picks up the most natural solutions from several candidates.

Remark. If b is Lipschitz continuous and $b(0) = 0$, then X_t^ϵ converges to $X_t \equiv 0$ as $\epsilon \downarrow 0$, since it is the unique minimizer of I.

5 Stochastic Partial Differential Equations

In the last section we saw that "SDE = ODE + random noise". This noise in time was modeled by a Brownian motion. Now we want to look at partial differential equations, so this should translate to "SPDE = PDE + random space-time noise". Examples for these kind of equations are the stochastic Navier-Stokes equations, stochastic reaction-diffusion equations and equations in control theory, filtering and finance.

Let us consider for $u = u(t, x)$, $t \geq 0$, $x \in \mathbb{R}^d$ the equation

$$\frac{\partial u}{\partial t} = Au + B(u) + C\dot{W}_t(x), \tag{5.13}$$

where

(i) $A = \sum_{|\alpha|\leq 2m} a_\alpha(x) D^\alpha$ with $a_\alpha \in C_b^\infty(\mathbb{R}^d)$, $m \in \mathbb{N}$ and $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}_+^d$. The derivative is the usual $D^\alpha = \left(\frac{\partial}{\partial x^1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x^d}\right)^{\alpha_d}$. The coefficients should satisfy some uniform parabolicity condition

$$\inf_{x,\sigma\in\mathbb{R}^d, |\sigma|=1} (-1)^{m+1} \sum_{|\alpha|=2m} a_\alpha(x)\sigma^\alpha > 0,$$

where $\sigma^\alpha = \sigma_1^{\alpha_1} \cdots \sigma_d^{\alpha_d}$ for $\sigma = (\sigma_1, \ldots, \sigma_d)$.

(ii) $B(u)$ is a nonlinear term.

(iii) $C = \sum_{|\alpha|\le l} c_\alpha(x)D^\alpha$ with $c_\alpha \in C_b^\infty(\mathbb{R}^d)$, $l \in \mathbb{Z}$. l may be negative, then it is seen as an integral operator.

(iv) $\dot{W}_t(x)$ is a space-time White Noise, i. e. a Gaussian noise with mean 0 and covariance structure
$$\mathbb{E}\left[\dot{W}_t(x)\dot{W}_s(y)\right] = \delta(t-s)\delta(x-y). \tag{5.14}$$
This notation is only used by physicists, we define the noise in a correct way later on.

Remark. For $\dot{B}_t = \frac{dB_t}{dt}$ we formally have
$$\mathbb{E}\left[\dot{B}_t\dot{B}_s\right] = \delta(t-s),$$
since for $\varphi, \psi \in C_c^\infty((0,\infty))$ we formally have (since the derivative does not exist)
$$\begin{aligned}\int_0^\infty dt \int_0^\infty ds \varphi(t)\psi(s)\mathbb{E}\left[\dot{B}_t\dot{B}_s\right] &= \mathbb{E}\left[\langle\varphi,\dot{B}\rangle\langle\psi,\dot{B}\rangle\right] \\ = \mathbb{E}\left[\langle\dot{\varphi},B\rangle\langle\dot{\psi},B\rangle\right] &= \int_0^\infty dt \int_0^\infty ds \dot{\varphi}(t)\dot{\psi}(s)\mathbb{E}\left[B_tB_s\right] \\ = \int_0^\infty \dot{\varphi}(t)\int_0^\infty \dot{\psi}(s)(t\wedge s)ds dt &= \int_0^\infty \varphi(t)\psi(t)dt.\end{aligned}$$

5.1 Definition of a Suitable Noise

In this section we want to extend the definition of the d-dimensional Brownian Motion to an infinite dimensional setting. Let H be a separable Hilbert space, for simplicity $H = L^2(\mathbb{R}^d, dx)$ and $(e_i)_{i\in\mathbb{N}}$ a complete orthonormal system of H. We now take a sequence of independent 1-dimensional Brownian Motions $(B_t^i)_{i\in\mathbb{N}}$ and want to define
$$W_t(x) := \sum_{i\in\mathbb{N}} B_t^i e_i(x), \quad x \in \mathbb{R}^d.$$
Unfortunately $W_t \notin H$, but we can extend H to a larger space $\tilde{H}$ of generalized functions. In fact, if we take a sequence $(\lambda_i > 0)_{i\in\mathbb{N}}$ such that $\sum_{i\in\mathbb{N}}\lambda_i < \infty$ we can define a weaker norm than $\|\cdot\|_H$ by
$$\|f\|_{\tilde{H}}^2 := \sum_{i\in\mathbb{N}} \lambda_i \langle f, e_i\rangle_H^2, \quad f \in H.$$
Then if $\tilde{H}$ is the completion of H under $\|\cdot\|_{\tilde{H}}$, we actually get $W_t \in \tilde{H}$ $\mathbb{P}$-a. s. Formally we have
$$\mathbb{E}\left[\|W_t\|_{\tilde{H}}^2\right] = \sum_{i\in\mathbb{N}}\lambda_i\mathbb{E}\left[(B_t^i)^2\right] = t\sum_{i\in\mathbb{N}}\lambda_i < \infty.$$

Nevertheless the scalar product with elements in H is always definable as $\langle W_t, f\rangle_H = W_t(\varphi) := \sum_{i\in\mathbb{N}} B_t^i \langle\varphi, e_i\rangle_H$, since

$$\begin{aligned}\mathbb{E}\left[\langle W_t,\varphi\rangle_H \langle W_s,\psi\rangle_H\right] &= \sum_{i,j\in\mathbb{N}} \mathbb{E}\left[B_t^i B_s^j\right] \langle\varphi, e_i\rangle_H \langle\psi, e_j\rangle_H \\ &= (t\wedge s)\sum_{i\in\mathbb{N}} \langle\varphi, e_i\rangle_H \langle\psi, e_j\rangle_H = (t\wedge s)\langle\varphi,\psi\rangle_H,\end{aligned}$$

which is finite. This computation suggests (5.14) for the covariance structure of $\dot{W}_t(x)$.

5.2 Concepts of Solutions

Definition 5.1. $u(t,x)$ is called a solution of (5.13) in the sense of generalized functions, if it satisfies

$$\langle u(t),\varphi\rangle_H = \langle u_0,\varphi\rangle_H + \int_0^t \{\langle u(s), A^*\varphi\rangle_H + \langle B(u(s)),\varphi\rangle_H\} ds + W_t(C^*\varphi)$$

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

Definition 5.2. u is called a mild solution of (5.13), if it satisfies

$$u(t) = T(t)u_0 + \int_0^t T(t-s)B(u(s))ds + \int_0^t T(t-s)C\,dW_s,$$

where $T(t)$ is the semigroup generated by the operator A (in a proper space).

In some typical cases the two notions of solutions are equivalent.

5.3 Regularity of Solutions

Since the noise $W_t(x)$ only lives in a bad space $\tilde{H}$ we need the regularizing properties of the operator A. We want to consider the linear equation, therefore we set $B(u) = 0$ and we get the following result on the regularity of solutions, apply Kolmogorov-Čentsov's theorem (see, for example, Kunita ('90)) noting Lemma 2.3 in Funaki ('91) and the Gaussian property of the solutions.

Proposition 5.3. *Let $u(t,x)$ be the solution to the linear equation corresponding to* (5.13). *Then*

$$u(t,x) \in \bigcap_{\delta>0} C^{\alpha-\delta,\beta-\delta}((0,\infty)\times\mathbb{R}^d),$$

with

$$\alpha = \frac{2m-2l-d}{4m} \quad \textit{and} \quad \beta = \frac{2m-2l-d}{2}.$$

If the noise is the space-time White Noise (i. e. $l=0$) and $A=\Delta$ (i. e. $m=1$) we have

$$u(t,x) \in \bigcap_{\delta>0} C^{\frac{1}{2}-\frac{d}{4}-\delta, 1-\frac{d}{2}-\delta}((0,\infty)\times\mathbb{R}^d).$$

Therefore the solution lives in the usual function spaces only, when $d=1$. If $d=2$, the solution is already a generalized function. Of course, this can be improved, if we take l to be negative, which results in a more regular noise in the space variable, a so-called colored noise.

If we consider the stochastic Navier-Stokes equations as an example for equation (5.13), the nonlinear term $u\cdot\nabla u$ appears, so that the noise has to be colored, or some special way to interpret the nonlinearity is required.

5.4 The KPZ Equation

Kardar-Parisi-Zhang ('86) introduced the following stochastic partial differential equation for a height function $h(t,x)$ of a randomly growing interface.

$$\frac{\partial h}{\partial t} = \frac{1}{2}\frac{\partial^2 h}{\partial x^2} - \frac{1}{2}\left(\frac{\partial h}{\partial x}\right)^2 + \dot{W}_t(x), \quad t>0, x\in\mathbb{R}, \tag{5.15}$$

where $\dot{W}_t(x)$ is the space-time White Noise. Without the nonlinear term we have seen that $h\in\bigcap_{\delta>0} C^{\frac{1}{4}-\delta,\frac{1}{2}-\delta}$ and for such h the space derivative $\frac{\partial h}{\partial x}$ is not definable. Therefore (5.15) does not have a solution in the usual sense.

Consider the slope of the interface $u(t,x) = \frac{\partial h}{\partial x}(t,x)$. For this function we obtain the stochastic viscous Burgers' equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} - \frac{1}{2}\frac{\partial}{\partial x}u^2 + \frac{\partial}{\partial x}\dot{W}_t(x).$$

Here the noise has even less regularity, but a formal application of the Hopf-Cole transformation

$$Z(t,x) := e^{\int_x^\infty u(t,y)dy} = e^{h(t,\infty)-h(t,x)}$$

or setting

$$Z(t,x) := e^{-h(t,x)} \tag{5.16}$$

leads to the linear SPDE

$$\frac{\partial Z}{\partial t} = \frac{1}{2}\frac{\partial^2 Z}{\partial x^2} - Z\dot{W}_t(x). \tag{5.17}$$

This stochastic partial differential equation has a unique continuous solution and we interpret (5.15) by (5.17) with $h(t,x) = -\log Z(t,x)$. Such h is called the Hopf-Cole solution of (5.15) and was introduced by Bertini and Giancomin in ('97).

Remark ($\frac{1}{3}$-power law)**.** Recently Balász, Quastel and Seppäläinen ('10) showed that if we choose $Z(0,x) = e^{-B(x)}$ with B being a two-sided Brownian Motion independent of the noise $\dot{W}_t(x)$ (i.e. we consider a stationary solution) then

$$ct^{\frac{2}{3}} \le \mathrm{Var}(h(t,0)) \le Ct^{\frac{2}{3}},$$

i. e. the fluctuations of $h(t,0)$ are of order $t^{\frac{1}{3}}$. This is a different behavior from the Central Limit Theorem. See also recent works by Sasamoto and Spohn ('10) who gave much precise behaviors.

Acknowledgement. The author thanks Professor Yoshihiro Shibata who gave him the opportunity to deliver one of the special lectures on Mathematical Fluid Dynamics at Waseda University. He also thanks Martin Sauer who did superb typing job of the notes.

References

[1] L. Bertini and G. Giacomin (1997): *Stochastic Burgers and KPZ equations from particle systems*, Commun. Math. Phys., **183**, 571–607.

[2] T. Funaki (1991): *Regularity properties for stochastic partial differential equations of parabolic type*, Osaka J. Math., **28**, 495–516.

[3] T. Funaki (2005): *Stochastic Differential Equations*, (in Japanese) Iwanami.

[4] G. Da Prato, J. Zabczyk (1992): *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, Cambridge University Press.

[5] G. Da Prato, J. Zabczyk (1996): *Ergodicity for Infinite Dimensional Systems*, London Mathematical Society Lecture Notes, Volume 229, Cambridge University Press.

[6] I. Karatzas, S.E. Shreve (1991): *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, Volume 113, Springer, 2nd edition.

[7] H. Kunita (1990): *Stochastic flows and stochastic differential equations*, Cambridge University Press.

[8] D. Revuz, M. Yor (2005): *Continuous Martingales and Brownian Motion*, Grundlehren der mathematischen Wissenschaften, Volume 293, Springer, 3rd edition.