A Gentle Introduction to Stochastic Differential Equations with Brownian Noise

Tommi Sottinen
School of Technology and Innovations
University of Vaasa
P.O. Box 700
FIN-65101 Vaasa
FINLAND
tommi.sottinen@uwasa.fi

June 8, 2023

Preface

These are notes for a 7.5 hour crash course on stochastic differential equations. There are 5 sections and each section is designed to be 90 min lecture session. In addition to the 7.5 hours of lectures the course has 2 exercise sessions, 90 min each.

Each section has 4 exercises. Completing 10 exercises is enough for a passing grade.

There are many excellent textbooks on stochastic differential equations with different level of mathematical sophistication. In writing these notes the author has used mainly Karatzas and Shreve [3], Øksendal [4], Revuz and Yor [5], and Schilling and Partzsch [6]. The author claims no originality. Indeed, most of the material here is copy/pasted from the above mentioned textbooks. Also, it should be noted that the proofs of these notes are sketchy at best. For more rigorous proofs the reader are referred to the textbooks mentioned above. We have, however, tried to give a flavor of proof in all cases except for Theorem 2.1 characterizing the space of Itô integrands, Theorem 2.3, the Doob maximal inequality, and Proposition 5.1 stating that Itô diffusions are Markovian. Rigorous proofs for all of these results can be found for example in Revuz and Yor [5].

In these notes we have simplified our story a little bit by assuming that all our processes and random variables are square-integrable.

The suggested measure-theoretical probability background for students for this course is given in Williams [8].

Tommi Sottinen

Contents

1 Brownian Motion

1.1 Brownian Motion as Martingale

The Brownian motion is arguably the most central stochastic process there is. It belongs to the intersection of many mathematical models: it is Gaussian, it is a Lévy process, it is a martingale. An impressive amount of formulas are known for the Brownian motion, see Borodin and Salminen [1].

A stochastic process $X = (X_t(\omega), t \geq 0, \omega \in \Omega)$ is a collection of random variables indexed by time t. The intrinsic filtration of a stochastic process X is $\mathbb{F}^X = (\mathscr{F}_t^X)_{t\geq 0}$ is given by $\mathscr{F}_t^X = \sigma(X_u; u \leq t)$. Intuitively this means that $\mathscr{F}_t^X$ is the information given by observing the process X over the time-interval $[0,t]$.

Below we give a qualitative definition of Brownian motion as a continuous Lévy process.

Definition 1.1 (Lévy Process)**.** A stochastic process L is a Lévy process if $L_0 = 0$, it has stationary and independent increments, and it has right-continuous paths with left limits. Let $s < t$. The stationarity of the increments means that the law of $L_t - L_s$ depends only on $t-s$ and not on t or s. The independence of the increments mean that $L_t - L_s$ if independent of the information (sigma-algebra) $\sigma(L_u; u \leq s)$ generated by the random variables L_s, $u \leq s$.

Definition 1.2 (Brownian motion)**.** A 1-dimensional stochastic process W is a Brownian motion if it is a centered continuous Lévy process with $\mathbb{E}[W_1^2] = 1$. A d-dimensional stochastic process is a Brownian motion if its components are independent 1-dimensional Brownian motions.

Remark 1.1. We will later show that the Brownian motion is Gaussian.

Let us recall the notions of conditional expectation and martingale. We define the conditional expectation only for square-integrable random variables.

Definition 1.3 (Conditional Expectation)**.** Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $X \in L^2(\Omega, \mathscr{F}, \mathbb{P})$ be a square-integrable random variable. Let $\mathscr{G}$ be a sub-sigma-algebra of $\mathscr{F}$. Then $L^2(\Omega, \mathscr{G}, \mathbb{P})$ is a subspace of $L^2(\Omega, \mathscr{F}, \mathbb{P})$ and the conditional expectation of $X \in L^2(\Omega, \mathscr{F}, \mathbb{P})$ is its orthogonal projection $\mathbb{E}[X|\mathscr{G}]$ to the subspace $L^2(\Omega, \mathscr{G}, \mathbb{P})$. In other words $\mathbb{E}[X|\mathscr{G}]$ is the $\mathscr{G}$-measurable random variable satisfying

$$\mathbb{E}[X|\mathscr{G}] \quad = \quad \operatorname*{argmin}_{Y \in L^2(\Omega,\mathscr{G},\mathbb{P})} \mathbb{E}[(X-Y)^2].$$

Remark 1.2 (Kolmogorov Definition of Conditional Expectation). $Y = \mathbb{E}[X|\mathscr{G}]$ if Y is $\mathscr{G}$-measurable and for all $A \in \mathscr{G}$ we have

$$\int_A X(\omega)\, \mathbb{P}[\mathrm{d}\omega] = \int_A Y(\omega)\, \mathbb{P}[\mathrm{d}\omega].$$

Below is a list of properties the conditional expectation satisfies

Proposition 1.1 (Properties of Conditional Expectation)**.** *Let* $X, X^{(m)}, Y, Z \in L^2(\Omega, \mathscr{F}, \mathbb{P})$ *and let* $\mathscr{G}$ *be a sub-sigma-algebra of* $\mathscr{F}$*. Let* $a, b \in \mathbb{R}$*.*

(i) $\mathbb{E}[\mathbb{E}[X|\mathscr{G}]] = \mathbb{E}[X]$.

(ii) *If* X *is* $\mathscr{G}$*-measurable, then* $\mathbb{E}[X|\mathscr{G}] = X$.

(iii) $\mathbb{E}[aX + bY|\mathscr{G}] = a\mathbb{E}[X|\mathscr{G}] + b\mathbb{E}[Y|\mathscr{G}]$.

(iv) *If* $X \geq 0$*, then* $\mathbb{E}[X|\mathscr{G}] \geq 0$.

(v) *If* $0 \leq X^{(m)} \uparrow X$*, then* $\mathbb{E}[X^{(m)}|\mathscr{G}] \uparrow \mathbb{E}[X|\mathscr{G}]$.

(vi) *If* $X^{(m)} \geq 0$*, then* $\mathbb{E}[\liminf X^{(m)}|\mathscr{G}] \leq \liminf \mathbb{E}[X^{(m)}|\mathscr{G}]$.

(vii) *If* $|X^{(m)}| \leq Y$ *and* $X^{(m)} \to X$*, then* $\mathbb{E}[X^{(m)}|\mathscr{G}] \to \mathbb{E}[X|\mathscr{G}]$.

(viii) *If* $c \colon \mathbb{R} \to \mathbb{R}$ *is convex, then* $\mathbb{E}[c(X)|\mathscr{G}] \geq c(\mathbb{E}[X|\mathscr{G}])$.

(ix) *If* $\mathscr{H}$ *is a sub-sigma-algebra of* $\mathscr{G}$*, then* $\mathbb{E}[\mathbb{E}[X|\mathscr{G}]|\mathscr{H}] = \mathbb{E}[X|\mathscr{H}]$.

(x) *If* Z *is* $\mathscr{G}$*-measurable, then* $\mathbb{E}[XZ|\mathscr{G}] = Z\mathbb{E}[X|\mathscr{G}]$.

(xi) *If* $\mathscr{H}$ *is independent of* $\sigma(\sigma(X), \mathscr{G})$ *then* $\mathbb{E}[X|\mathscr{G}, \mathscr{H}] = \mathbb{E}[X|\mathscr{G}]$*. In particular, if* X *is independent of* $\mathscr{G}$ *then* $\mathbb{E}[X|\mathscr{G}] = \mathbb{E}[X]$.

Proof. See Williams [8], Section 9.7. □

Remark 1.3. For the purposes of this course, it is enough to know the following: Let X be $\mathscr{G}$-measurable and let Y be independent of $\mathscr{G}$. Let $f \colon \mathbb{R}^2 \to \mathbb{R}$ be such that $f(X, Y)$ is square-integrable. Then

$$\mathbb{E}\left[f(X,Y)\middle|\mathscr{G}\right] \;=\; \mathbb{E}\left[f(x,Y)\right]_{x=X}.$$

Intuitively, a stochastic process $M = (M_t)_{t\geq 0}$ is a martingale for the filtration $\mathbb{F} = (\mathscr{F}_t)_{t\geq 0}$ if the best prediction of the "future" M_t given the "today" information $\mathscr{F}_s$ is the today-value M_s. Formal definition is as follows:

Definition 1.4 (Martingale)**.** Let M be a square-integrable stochastic process and let $\mathbb{F}$ be a filtration. If M_t is $\mathscr{F}_t$-measurable and for all $s < t$ and it holds that $M_s = \mathbb{E}[M_t|\mathscr{F}_s]$, we say that M is an $\mathbb{F}$-martingale. If $\mathbb{F} = \mathbb{F}^M$, we say simply that M is a martingale.

Proposition 1.2. *The Brownian motion is a martingale.*

Proof. Let $\mathbb{F} = (\mathscr{F}_t)_{t\geq 0}$ be the intrinsic filtration of the Brownian motion. Let $t > s$. Write

$$W_t = W_s + (W_t - W_s).$$

Then W_s is $\mathscr{F}_s$-measurable and $W_t - W_s$ is independent of $\mathscr{F}_s$. Consequently,

$$\begin{aligned}\mathbb{E}\left[W_t \middle| \mathscr{F}_s\right] &= \mathbb{E}\left[W_s + (W_t - W_s) \middle| \mathscr{F}_s\right] \\ &= \mathbb{E}\left[W_s \middle| \mathscr{F}_s\right] + \mathbb{E}\left[W_t - W_s \middle| \mathscr{F}_s\right] \\ &= W_s + \mathbb{E}\left[W_t - W_s\right] \\ &= W_s\end{aligned}$$

showing that W is a martingale. □

Exercise 1.1. Suppose that we know that the 1-dimensional Brownian motion is Gaussian. Show that the (1-dimensional) law of $W_t - W_s$ conditioned on the intrinsic sigma-algebra $\mathscr{F}^W_s = \sigma(W_u; u \leq s)$ is Gaussian with mean 0 and variance $t - s$.

Remark 1.4. The following formula is useful later: Let M be a martingale. Let $s < t$. Then

$$\mathbb{E}\left[(M_t - M_s)^2 \middle| \mathscr{F}_s\right] = \mathbb{E}\left[M_t^2 - M_s^2 \middle| \mathscr{F}_s\right]. \tag{1.1}$$

To see that (1.1) holds, we first note that

$$\begin{aligned}\mathbb{E}\left[(M_t - M_s)^2 \middle| \mathscr{F}_s\right] &= \mathbb{E}\left[M_t^2 - 2M_sM_t + M_s^2 \middle| \mathscr{F}_s\right] \\ &= \mathbb{E}\left[M_t^2 \middle| \mathscr{F}_s\right] - \mathbb{E}\left[2M_tM_s \middle| \mathscr{F}_s\right] + \mathbb{E}\left[M_s^2 \middle| \mathscr{F}_s\right] \\ &= \mathbb{E}\left[M_t^2 \middle| \mathscr{F}_s\right] - 2M_s\mathbb{E}\left[M_t \middle| \mathscr{F}_s\right] + M_s^2 \\ &= \mathbb{E}\left[M_t^2 \middle| \mathscr{F}_s\right] - 2M_s^2 + M_s^2 \\ &= \mathbb{E}\left[M_t^2 \middle| \mathscr{F}_s\right] - M_s^2 \\ &= \mathbb{E}\left[M_t^2 \middle| \mathscr{F}_s\right] - \mathbb{E}\left[M_s^2 \middle| \mathscr{F}_s\right] \\ &= \mathbb{E}\left[M_t^2 - M_s^2 \middle| \mathscr{F}_s\right]\end{aligned}$$

showing the claim.

Exercise 1.2. Let W be a 1-dimensional Brownian motion. Show that the process $t \mapsto W_t^2 - t$ is a martingale.

1.2 Quadratic Variation of Brownian Motion

Informally, the paths of Brownian paths are so rough that

$$(\mathrm{d}W_t)^2 = \mathrm{d}t.$$

This is what Theorem 1.1 below says rigorously.

Theorem 1.1 (Quadratic Variation)**.** *Let W be a 1-dimensional Brownian motion. Let $\Pi = \{0 = t_0 < t_1 \cdots < t_m = t\}$ be a partition of $[0,t]$. Let $|\Pi| = \max_k(t_k - t_{k-1})$. Let $\Delta W_{t_k} = W_{t_k} - W_{t_{k-1}}$. Then*

$$\langle W \rangle_t \;=\; \lim_{|\Pi|\to 0} \sum_{t_k\in\Pi} (\Delta W_{t_k})^2 \;=\; t.$$

Proof. Denote $\Delta t_k = t_k - t_{k-1}$ and

$$Y_k \;=\; \left(\frac{\Delta W_{t_k}}{\sqrt{\Delta t_k}}\right)^2.$$

Now, because of stationarity and independence of the increments of the Brownian motion we have $\mathbb{E}[(\Delta W_{t_k})^2] = \Delta t_k$. Consequently, $\mathbb{E}[Y_k] = 1$. Now, note that the Y_k's are independent and identically distributed. Then, the claim follows from the law of the large numbers. Indeed, we have

$$\sum_{t_k\in\Pi} (\Delta W_{t_k})^2 \;=\; \sum_{t_k\in\Pi} \Delta t_k Y_k \;\sim\; \frac{t}{m}\sum_{k=1}^{m} Y_k.$$

This proves the claim. □

Exercise 1.3 (Nowhere Differentiability of Brownian Motion)**.** Show, by using Theorem 1.1, that the paths of Brownian motion are nowhere differentiable.

Exercise 1.4 (Quadratic Covariation)**.** Let $\Pi = \{0 = t_0 < t_1 \cdots < t_m = t\}$ be a partition of $[0,t]$. Let $|\Pi| = \max_k(t_k - t_{k-1})$. Let X and Y be continuous processes. Assume that the quadratic covariation

$$\langle X, Y \rangle_t \;=\; \lim_{|\Pi|\to 0} \sum_{t_k\in\Pi} (\Delta X_{t_k})(\Delta Y_{t_k})$$

exists. Show that the following polarization formula holds:

$$\langle X, Y \rangle \;=\; \frac{1}{2}\left(\langle X + Y \rangle - \langle X \rangle - \langle Y \rangle\right).$$

Remark 1.5. It should be noted that so far we have not assumed that the Brownian motion is Gaussian. Indeed, we shall prove this later in Section 3 in Proposition 3.1.

2 Itô Integration

2.1 Construction of Itô Integral

In what follows, $\mathbb{F} = (\mathscr{F}_t)_{t\geq 0}$ is the intrinsic filtration of the Brownian motion. In other words $\mathscr{F}_t = \sigma(W_u; u \leq t)$.

The indicator function $\mathbf{1}_A$ is defined as

$$\mathbf{1}_A(s) \quad = \quad \begin{cases} 1, & \text{if } s \in A, \\ 0, & \text{if } s \notin A. \end{cases}$$

Definition 2.1. The class of $\mathbb{F}$-predictable elementary stochastic process $\mathscr{E}$ contains processes of the form

$$H_s \quad = \quad \sum_{k=1}^{m} h_k \mathbf{1}_{(t_{k-1},t_k]}(s),$$

where h_k is $\mathscr{F}_{t_{k-1}}$-measurable and bounded.

Remark 2.1. Sometimes we write $H \in \mathscr{E}$ as

$$H_s \quad = \quad \sum_{k=1}^{m} H_{t_{k-1}} \mathbf{1}_{(t_{k-1},t_k]}(s).$$

If $H \in \mathscr{E}$ we can define the stochastic integral in the natural way as

$$\int_0^\infty H_s \, \mathrm{d}W_s \quad = \quad \sum_{t_k} H_{t_{k-1}} \Delta W_{t_k} \quad = \quad \sum_{k=1}^{m} h_k \Delta W_{t_k}, \tag{2.1}$$

where $\Delta W_{t_k} = W_{t_k} - W_{t_{k-1}}$.

Exercise 2.1. Calculate the expectations of the following sums

$$\sum_{k=1}^{m} W_{\frac{k-1}{m}} \left(W_{\frac{k}{m}} - W_{\frac{k-1}{m}} \right),$$

$$\sum_{k=1}^{m} W_{\frac{k-\frac{1}{2}}{m}} \left(W_{\frac{k}{m}} - W_{\frac{k-1}{m}} \right),$$

$$\sum_{k=1}^{m} W_{\frac{k}{m}} \left(W_{\frac{k}{m}} - W_{\frac{k-1}{m}} \right).$$

Conclude that Riemann–Stieltjes integration with respect to Brownian motion is practically impossible.

The following, elementary version of Itô isometry is the key ingredient in extending the simple Itô-intergral (2.1) beyond $\mathscr{E}$. We leave the following details of the proof as an exercise.

Exercise 2.2. Let $H \in \mathscr{E}$. Show that

$$\sum_{t_k, t_\ell :\, t_k < t_\ell} \sum \mathbb{E}\left[H_{t_{k-1}} H_{t_{\ell-1}} \Delta W_{t_k} \Delta W_{t_\ell}\right] \;=\; 0$$

and

$$\mathbb{E}\left[(\Delta W_{t_k})^2 \middle| \mathscr{F}_{t_{k-1}}\right] \;=\; \Delta t_k.$$

Lemma 2.1 (Simple Itô Isometry)**.** *Let $H \in \mathscr{E}$. Then*

$$\mathbb{E}\left[\left(\int_0^\infty H_s \,\mathrm{d}W_s\right)^2\right] \;=\; \int_0^\infty \mathbb{E}[H_s^2]\,\mathrm{d}s$$

Proof. Now,

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^\infty H_s \,\mathrm{d}W_s\right)^2\right] &= \mathbb{E}\left[\left(\sum_{t_k} H_{t_{k-1}} \Delta W_{t_k}\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{t_k} H_{t_{k-1}} \Delta W_{t_k}\right)\left(\sum_{t_\ell} H_{t_{\ell-1}} \Delta W_{t_\ell}\right)\right] \\
&= \sum_{t_k} \sum_{t_\ell} \mathbb{E}\left[H_{t_{k-1}} H_{t_{\ell-1}} \Delta W_{t_k} \Delta W_{t_\ell}\right] \\
&= 2 \sum_{t_k, t_\ell :\, t_k < t_\ell} \sum \mathbb{E}\left[H_{t_{k-1}} H_{t_{\ell-1}} \Delta W_{t_k} \Delta W_{t_\ell}\right] + \sum_{t_k} \mathbb{E}\left[H_{t_{k-1}}^2 (\Delta W_{t_k})^2\right] \\
&= \sum_{t_k} \mathbb{E}\left[H_{t_{k-1}}^2 (\Delta W_{t_k})^2\right] \\
&= \sum_{t_k} \mathbb{E}\left[\mathbb{E}\left[H_{t_{k-1}}^2 (\Delta W_{t_k})^2 \middle| \mathscr{F}_{t_{k-1}}\right]\right] \\
&= \sum_{t_k} \mathbb{E}\left[H_{t_{k-1}}^2 \mathbb{E}\left[(\Delta W_{t_k})^2 \middle| \mathscr{F}_{t_{k-1}}\right]\right] \\
&= \sum_{t_k} \mathbb{E}\left[H_{t_{k-1}}^2\right] \Delta t_k \\
&= \int_0^\infty \mathbb{E}[H_s^2]\,\mathrm{d}s,
\end{aligned}$$

which proves the claim. □

Now we extend the Itô integral from $\mathscr{E}$ to a space we call $\mathscr{L}^2$. The idea is the following. Let L^2 be the space of (centered) square-integrable random variables endowed with the norm $\|X\|^2 = \mathbb{E}[X^2]$. Let us endow $\mathscr{E}$ with the norm

$$\|H\|^2 \;=\; \int_0^\infty \mathbb{E}[H_s^2]\,\mathrm{d}s. \tag{2.2}$$

Then the Itô isometry of Lemma 2.1 states that the mapping

$$H \quad \mapsto \quad \int_0^\infty H_s \, \mathrm{d}W_s$$

is an isometry from $\mathscr{E}$ to L^2. Now, it follows immediately that this mapping extends to the closure $\bar{\mathscr{E}}$ under the norm (2.2), and the Itô-isometry of Lemma 2.1 holds for this extended Itô integral. In other words we have the following

Definition 2.2 (Itô Integral)**.** Let $H \in \mathscr{L}^2 = \bar{\mathscr{E}}$. Let $H^{(m)} \in \mathscr{E}$ approximate H. Then

$$\int_0^\infty H_s \, \mathrm{d}W_s \quad = \quad \lim_{m\to\infty} \int_0^\infty H_s^{(m)} \, \mathrm{d}W_s.$$

The following is immediate from the definition.

Theorem 2.1 (Itô Isometry)**.** *Let* $H \in \mathscr{L}^2$*. Then*

$$\mathbb{E}\left[\left(\int_0^\infty H_s \, \mathrm{d}W_s\right)^2\right] \quad = \quad \int_0^\infty \mathbb{E}\left[H_s^2\right] \mathrm{d}s.$$

We end this subsection by characterizing, the space $\mathscr{L}^2 = \bar{\mathscr{E}}$.

Proposition 2.1 (The Space $\mathscr{L}^2$)**.** *A stochastic process* $H \in \mathscr{L}^2$ *if and only if*

(i) $(t, \omega) \mapsto H_t(\omega)$ *is jointly measurable.*

(ii) H_t *is* $\mathscr{F}_t$*-measurable.*

(iii) *For all* $T > 0$*,*

$$\mathbb{E}\left[\int_0^T H_s^2 \, \mathrm{d}s\right] \quad < \quad \infty.$$

Proof. See Øksendal [4], Section 3.1. □

2.2 Properties of Itô Integral

The Itô integral $\int_0^\infty H_s \, \mathrm{d}W_s$ is linear in terms of the integrator H, this is obvious. It is also linear in terms of the integration limits. Indeed, let us define

$$\int_a^b H_s \, \mathrm{d}W_s \quad = \quad \int_0^\infty H_s \mathbf{1}_{(a,b]}(s) \, \mathrm{d}W_s,$$

then we have the following.

Proposition 2.2. *Let $H \in \mathscr{L}^2$. Then for all $t < T$ we have*

$$\int_0^T H_s \, \mathrm{d}W_s = \int_0^t H_s \, \mathrm{d}W_s + \int_t^T H_s \, \mathrm{d}W_s$$

Exercise 2.3. Prove Proposition 2.2.

Theorem 2.2 (Martingale Property)**.** *Let $H \in \mathscr{L}^2$. Then the process*

$$\int_0^\cdot H_s \, \mathrm{d}W_s$$

is a continuous square-integrable martingale

The main part of the proof of Theorem 2.2 is left as an exercise.

Exercise 2.4. Let $H \in \mathscr{E}$. Show that the process

$$\int_0^\cdot H_s \, \mathrm{d}W_s$$

is a square-integrable martingale continuous Martingale.

Proof of Theorem 2.2. Let $H^{(m)} \in \mathscr{E}$ approximate H in $\mathscr{L}^2$. Then $X^{(m)} = \int_0^\cdot H_s^{(m)} \, \mathrm{d}W_s$ is a square-integrable martingale by Exercise 2.4. The martingale property follows by the dominated convergence theorem for conditional expectations (see Proposition 1.1). For the continuity we need the Doob maximal inequality (see Theorem 2.3 later). The continuity then follows from the following estimates:

$$\begin{aligned}
&\mathbb{P}\left[\sup_{t \le T} \left| \int_0^t (H_s^n - H_s) \, \mathrm{d}W_s \right| > \varepsilon \right] \\
&= \frac{1}{\varepsilon^2} \mathbb{E}\left[\sup_{t \le T} \left| \int_0^t (H_s^n - H_s) \, \mathrm{d}W_s \right|^2 \right] \\
&\le \frac{4}{\varepsilon^2} \mathbb{E}\left[\left| \int_0^T (H_s^n - H_s) \, \mathrm{d}W_s \right|^2 \right] \\
&= \frac{4}{\varepsilon^2} \mathbb{E}\left[\int_0^T |H_s^n - H_s|^2 \, \mathrm{d}W_s \right] \\
&\to 0.
\end{aligned}$$

This means that $\int_0^\cdot H_s^n \, \mathrm{d}W_s$ converges to $\int_0^\cdot H_s \, \mathrm{d}W_s$ uniformly in probability. By passing to a subsequence (n') we obtain that

$$\sup_{t \le T} \left| \int_0^t H_s^{n'} \, \mathrm{d}W_s - \int_0^t H_s \, \mathrm{d}W_s \right| \to 0$$

almost surely. The continuity now follows from the fact that uniform limit of continuous functions is continuous. $\square$

We end this subsection by a version of the so-called Doob maximal inequality that we need later in the course in Section 4 in connection to existence and uniqueness of the solutions of stochastic differential equations. We consider n-dimensional processes. We denote the Euclidean norm $|b|^2 = \sum_i b_i^2$ for vectors and the Frobenius norm $|\sigma|^2 = \sum_i \sum_k \sigma_{ik}^2$ for matrices. Let $W = (W^1, \ldots, W^d)$ be a d-dimensional Brownian motion and let H be $\mathbb{R}^{n\times d}$-valued $\mathscr{L}^2$-process meaning that $H^{ik} \in \mathscr{L}^2$ for all i and k. By

$$X_t = \int_0^t H_s \,\mathrm{d}W_s$$

we mean the n-dimensional stochastic process $X = (X^1, \ldots, X^n)$ whose i^{th} component is given by

$$X_t^i \;=\; \sum_{k=1}^d \int_0^t H_s^{ik} \,\mathrm{d}W_s^k,$$

so that formally

$$X_t \;=\; \int_0^t H_s \,\mathrm{d}W_s \;=\; \int_0^t \begin{bmatrix} H_s^{11} & H_s^{12} & \cdots & H_s^{1d} \\ H_s^{22} & H_s^{22} & \cdots & H_s^{1d} \\ \vdots & \vdots & & \vdots \\ H_t^{n1} & H_t^{n2} & \cdots & H_s^{nd} \end{bmatrix} \begin{bmatrix} \mathrm{d}W_s^1 \\ \mathrm{d}W_s^2 \\ \vdots \\ \mathrm{d}W_s^d \end{bmatrix}.$$

Theorem 2.3 (Doob Maximal Inequality)**.** *Let $X = (X^1, X^2, \ldots, X^n)$ be a continuous martingale. Then*

$$\mathbb{E}\left[\sup_{t\le T} |X_t|^2\right] \;\le\; 4\mathbb{E}\left[|X_T|^2\right]$$

The proof of Theorem 2.3 is very technical. So we omit it. (The interested reader should consult Revuz and Yor [5].) Let us just note that if

$$X_t = \int_0^t H_s \,\mathrm{d}W_s$$

then by using the Itô isometry, we can rewrite the Doob maximal inequality as

$$\mathbb{E}\left[\sup_{t\le T}\left|\int_0^t H_s \,\mathrm{d}W_s\right|^2\right] \;\le\; 4\int_0^T \mathbb{E}\left[|H_s|^2\right] \mathrm{d}s \tag{2.3}$$

3 Itô Formula

The main tool, indeed the only tool, we have for Itô integrals is the so-called Itô formula. This is a change-of-variables formula for Itô integrals. We begin by stating and proving the formula for the simple case of 1-dimensional Brownian motion is Subsection 3.1 and in Subsection 3.2 we state and prove the Itô formula for n-dimensional Itô diffusions.

3.1 Homogeneous 1-Dimensional Itô Formula

The change-of-variables formula for classical calculus states that

$$\mathrm{d}f(X_t) = \frac{\mathrm{d}f}{\mathrm{d}x}(X_t)\,\mathrm{d}X_t \tag{3.1}$$

This formula is true if X is differentiable. Indeed, in this case the formula can be rewritten in a more familiar form as

$$\frac{\mathrm{d}f(X_t)}{\mathrm{d}t} = \frac{\mathrm{d}f(X_t)}{\mathrm{d}x}\frac{\mathrm{d}X_t}{\mathrm{d}t}.$$

Setting $x(t) = X_t$ we get an even more familiar form

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\mathrm{d}f}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}.$$

If $X = W$ is the Brownian motion, then the formula (3.1) is no longer true, and $\frac{\mathrm{d}W_t}{\mathrm{d}t}$ does not make classical sense. Informally, the reason is that as $\mathrm{d}t$ tends to zero $(\mathrm{d}W_t)^2$ tends to $\mathrm{d}t$ and not to zero. This implies that we have the formula

$$\begin{aligned}\mathrm{d}f(W_t) &= \frac{\mathrm{d}f}{\mathrm{d}x}(W_t)\,\mathrm{d}W_t + \frac{1}{2}\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(W_t)\,(\mathrm{d}W_t)^2 \\ &= \frac{\mathrm{d}f}{\mathrm{d}x}(W_t)\,\mathrm{d}W_t + \frac{1}{2}\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(W_t)\,\mathrm{d}t \qquad (3.2)\end{aligned}$$

Of course, the differential equation (3.2) has to be understood as an integral equation. The precise statement of formula (3.2) is given below in Theorem 3.1. For that we recall as an exercise a version of Taylor formula

Exercise 3.1 (2nd Order Taylor formula)**.** Let f be smooth enough. Show that

$$f(x) - f(a) = \frac{\mathrm{d}f}{\mathrm{d}x}(a)(x-a) + \frac{1}{2}\frac{\mathrm{d}^2f}{\mathrm{d}x^2}(a)(x-a)^2 + R(x,a),$$

where $R(a,x) \le \varepsilon(|a-x|)(x-y)^2$ and $\varepsilon(y) \to 0$ as $y \to 0$.

Theorem 3.1 (Itô Formula for 1-Dimensional Brownian motion)**.** *Let $f\colon \mathbb{R} \to \mathbb{R}$ be smooth enough. Then*

$$f(W_t) = f(W_0) + \int_0^t \frac{\mathrm{d}f}{\mathrm{d}x}(W_s)\,\mathrm{d}W_s + \frac{1}{2}\int_0^t \frac{\mathrm{d}^2f}{\mathrm{d}x^2}(W_s)\,\mathrm{d}s. \tag{3.3}$$

Proof. Let $\Pi = \{0 = t_0 < t_1 < \cdots < t_m = t\}$ be a partition of $[0,t]$. Let $\Delta W_{t_k} = W_{t_k} - W_{t_{k-1}}$ and $\Delta f(W_{t_k}) = f(W_{t_k}) - f(W_{t_{k-1}})$. Then, by Taylor formula,

$$\Delta f(W_{t_k}) \;\simeq\; \frac{\mathrm{d}f}{\mathrm{d}x}(W_{t_{k-1}})\Delta W_{t_k} + \frac{1}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(W_{t_{k-1}})(\Delta W_{t_k})^2 + \varepsilon(\Delta W_{t_{k-1}})(\Delta W_{t_k})^2$$

Summing over the partition Π this yields

$$\begin{aligned} f(W_t) - f(W_0) &\simeq \sum_{t_k\in\Pi} \frac{\mathrm{d}f}{\mathrm{d}x}(W_{t_{k-1}})\Delta W_{t_k} + \frac{1}{2}\sum_{t_k\in\Pi}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(W_{t_{k-1}})(\Delta W_{t_k})^2 \\ &\quad + \sum_{t_k\in\Pi}\varepsilon(\Delta W_{t_{k-1}})(\Delta W_{t_k})^2 \\ &= I_1(\Pi) + I_2(\Pi) + I_3(\Pi). \end{aligned}$$

Now, by the definition of Itô integral

$$I_1(\Pi) \;\to\; \int_0^t \frac{\mathrm{d}f}{\mathrm{d}x}(W_s)\,\mathrm{d}W_s$$

as $|\Pi| \to 0$.

Let us then consider the sum $I_2(\Pi)$. By the quadratic variation of Brownian motion (Theorem 1.1) we have informally

$$(\Delta W_{t_k})^2 \;\to\; (\mathrm{d}W_t)^2 \;=\; \mathrm{d}t.$$

Consequently,

$$\begin{aligned} I_2(\Pi) &= \frac{1}{2}\sum_{t_k\in\Pi}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(W_{t_{k-1}})(\Delta W_{t_k})^2 \\ &\to \frac{1}{2}\int_0^t \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(W_s)\,(\mathrm{d}W_s)^2 \\ &= \frac{1}{2}\int_0^t \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(W_s)\,\mathrm{d}s \end{aligned}$$

as $|\Pi| \to 0$.

It remains to show that $I_3(\Pi) \to 0$ as $|\Pi| \to 0$. Since W has (uniformly) continuous paths we have, by using the quadratic variation Theorem 1.1 that

$$\begin{aligned} I_3(\Pi) &= \sum_{t_k\in\Pi}\varepsilon(\Delta W_{t_{k-1}})(\Delta W_{t_k})^2 \\ &\le \sup_{t_k\in\Pi}\varepsilon(\Delta W_{t_{k-1}})\sum_{t_k\in\Pi}(\Delta W_{t_k})^2 \\ &\to 0. \end{aligned}$$

This finishes the proof. □

Remark 3.1 (Itô Formula and Quadratic Variation). It should be noted that the process W in the Itô formula (3.3) need not be Brownian motion. Indeed, only things that were needed in the proof where (i) the process W is continuous and (ii) $(\mathrm{d}W_t)^2 = \mathrm{d}t$. We refer to Föllmer [2] (see Sondermann [7] of English translation) on further discussion on how to construct stochastic calculus for quadratic variation processes.

Exercise 3.2. Calculate by using Itô formula

$$\int_0^t W_s \,\mathrm{d}W_s.$$

We end this subsection by showing that a continuous Lévy process (i.e. the Brownian motion) is Gaussian.

Proposition 3.1. *The Brownian motion is Gaussian.*

Proof. We only give the proof in 1-dimensional case. The multidimensional case follows simply by considering the independent components separately.

Let $f(x) = \mathrm{e}^{\mathrm{i}\theta x}$. Then, by the 1-dimensional Itô formula we have

$$\begin{aligned} \mathrm{d}f(W_t) &= \frac{\mathrm{d}f}{\mathrm{d}x}(W_t)\,\mathrm{d}W_t + \frac{1}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(W_t)\,\mathrm{d}t \\ &= \mathrm{i}\theta f(W_t)\mathrm{d}W_t - \frac{1}{2}\theta^2 f(W_t)\,\mathrm{d}t. \end{aligned} \tag{3.4}$$

Now we note that

$$\phi_t(\theta) \;=\; \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\theta W_t}\right] \;=\; \mathbb{E}\left[f(W_t)\right].$$

Consequently, by taking the expectation in (3.4), we obtain the integral equation

$$\phi_t(\theta) \;=\; 1 - \frac{1}{2}\theta^2 \int_0^t \phi_s(\theta)\,\mathrm{d}s.$$

The solution of this integral equation is the Gaussian characteristic function

$$\phi_t(\theta) \;=\; \mathrm{e}^{-\frac{1}{2}\theta^2 t},$$

which shows the claim. □

3.2 Itô Formula for Itô Diffusions

Let $W = (W^1, \ldots, W^d)$ be a d-dimensional Brownian motion. Let $b\colon [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma\colon [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$.

We consider the stochastic differential equation that is defined componentwise as

$$\mathrm{d}X_t^i \;=\; b_i(t, X_t)\,\mathrm{d}t + \sum_{k=1}^d \sigma_{ik}(t, X_t)\,\mathrm{d}W_t^k, \qquad X_0^i = \xi_i, \tag{3.5}$$

for each $i = 1, \ldots, n$. Of course, to be precise, the stochastic differential equation (3.5) should be understood as the componentwise integral equation

$$X_t^i = \xi_i + \int_0^t b_i(s, X_s)\, \mathrm{d}s + \sum_{k=1}^d \int_0^t \sigma_{ik}(s, X_s)\, \mathrm{d}W_s^k,$$

for each $i = 1, \ldots, n$. In the above the initial value ξ is $\mathscr{F}_0$-measurable and the filtration $\mathbb{F} = (\mathscr{F}_t)_{t \geq 0}$ is such that the Brownian motion W is an $\mathbb{F}$-martingale. In practice, this typically means that the sigma-algebra $\mathscr{F}_0$ and the Brownian motion W are independent.

Sometimes we write (3.5) shortly as

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

or

$$X_t = \xi + \int_0^t b(s, X_s)\, \mathrm{d}s + \int_0^t \sigma(s, X_s)\, \mathrm{d}W_s$$

Solutions X of the stochastic differential equations (3.5) are called Itô diffusions. In this subsection we simply assume that the solution X of (3.5) exists. We will prove the existence and uniqueness later in Section 4.

The proof of multidimensional Itô formula for Itô diffusions is, like in the simple Brownian motion case, based on Taylor formula. Therefore we give a version of the Taylor formula as an exercise.

Exercise 3.3 (Multidimensional 2^{nd} Order Taylor Formula)**.** Let f be smooth enough. Show the following multidimensional second order Taylor formula:

$$\begin{aligned} & f(x) - f(a) \\ &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x_i - a_i)(x_j - a_j) + R(x, a), \end{aligned} \tag{3.6}$$

where $R(x, a) \leq \varepsilon(x - a)|x - a|^2$ and $\varepsilon(y) \to 0$ as $y \to 0$, and $|\cdot|$ is the Euclidean norm $|y|^2 = \sum_{i=1}^n y_i^2$.

Below is the Itô formula for Itô diffusions in differential form.

Theorem 3.2 (Itô Formula for Itô Diffusions)**.** *Let X be the Itô diffusion (3.5). Let $f\colon [0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ be smooth enough. Then*

$$\mathrm{d}f(t, X_t) = \frac{\partial f}{\partial t}(t, X_t)\, \mathrm{d}t + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t, X_t)\, \mathrm{d}X_t^i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(t, X_t)\, \mathrm{d}X_t^i \mathrm{d}X_t^j, \tag{3.7}$$

where $\mathrm{d}X_t^i \mathrm{d}X_t^j$ is calculated according to the quadratic variation rules $\mathrm{d}W_t^i \mathrm{d}W_t^j = \delta_{ij} \mathrm{d}t$ and $\mathrm{d}W_t^i \mathrm{d}t = \mathrm{d}t \mathrm{d}W_t^i = (\mathrm{d}t)^2 = 0$. Here δ_{ij} is the Kronecker delta: $\delta_{ii} = 1$ and $\delta_{ij} = 0$ if $i \neq j$.

Proof. The formula (3.7) follows from the Taylor formula (3.6) an the proof of the one-dimensional simple Itô formula. The calculation rules $\mathrm{d}W_t^i \mathrm{d}W_t^j = \delta_{ij}\mathrm{d}t$, $\mathrm{d}W_t^i \mathrm{d}t = \mathrm{d}t\mathrm{d}W_t^i = 0$ follow from quadratic variation calculus. We omit the details. □

Remark 3.2. As in the 1-dimensional Brownian case, the multidimensional Itô diffusion Itô formula (3.7) remains true, if we only assume that (i) X is continuous (ii) X is quadratic covariation process meaning that

$$\langle X^i, X^j \rangle_t \;=\; \lim_{|\Pi|\to 0} \sum_{t_k\in\Pi} \Delta X^i_{t_k} \Delta X^j_{t_k}$$

exists. Indeed, then

$$\mathrm{d}X_t^i \mathrm{d}X_t^j \;=\; \mathrm{d}\langle X^i, X^j\rangle_t.$$

Remark 3.3. The Itô formula (3.7) can be written in many ways in terms of the underlying Brownian motion. We have, for example,

$$\begin{aligned}
\mathrm{d}f(t,X_t) &= \frac{\partial f}{\partial t}(t,X_t)\,\mathrm{d}t + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t,X_t)\mathrm{d}X_t^i \\
&\quad + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(t,X_t)(\sigma\sigma^\top)_{ij}(t,X_t)\,\mathrm{d}t \\
&= \frac{\partial f}{\partial t}(t,X_t)\,\mathrm{d}t + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t,X_t) b_i(t,X_t)\,\mathrm{d}t + \sum_{k=1}^d\sum_{i=1}^n \frac{\partial f}{\partial x_i}(t,X_t)\sigma_{ik}(t,X_t)\mathrm{d}W_t^k \\
&\quad + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(t,X_t)(\sigma\sigma^\top)_{ij}(t,X_t)\,\mathrm{d}t.
\end{aligned}$$

Indeed, we have

$$\mathrm{d}X_t^i \mathrm{d}X_t^j \;=\; (\sigma\sigma^\top)_{ij}(t,X_t)\,\mathrm{d}t,$$

since according to the rules $\mathrm{d}W_t^i \mathrm{d}W_t^j = \delta_{ij}\mathrm{d}t$, $\mathrm{d}W_t^i \mathrm{d}t = \mathrm{d}t\mathrm{d}W_t^i = (\mathrm{d}t)^2 = 0$,

$$\begin{aligned}
\mathrm{d}X_t^i \mathrm{d}X_t^j &= \left[b_i(t,X_t)\,\mathrm{d}t + \sum_{k=1}^d \sigma_{ik}(t,X_t)\mathrm{d}W_t^k\right]\left[b_j(t,X_t)\,\mathrm{d}t + \sum_{\ell=1}^d \sigma_{j\ell}(t,X_t)\mathrm{d}W_t^\ell\right] \\
&= \sum_{k=1}^d\sum_{\ell=1}^d \sigma_{ik}(t,X_t)\sigma_{j\ell}(t,X_t)\,\mathrm{d}W_t^k \mathrm{d}W_t^\ell \\
&= \sum_{k=1}^d \sigma_{ik}(t,X_t)\sigma_{jk}(t,X_t)\,\mathrm{d}t \\
&= (\sigma\sigma^\top)_{ij}(t,X_t)\,\mathrm{d}t.
\end{aligned}$$

Exercise 3.4 (Integration-by-Parts)**.** Let X and Y be 1-dimensional Itô diffusions. Show the following integration-by-parts formula:

$$\mathrm{d}(X_t Y_t) \;=\; X_t\,\mathrm{d}Y_t + Y_t\,\mathrm{d}X_t + \mathrm{d}X_t\,\mathrm{d}Y_t.$$

4 Stochastic Differential Equations

In this section we consider stochastic differential equations driven by a d-dimensional Brownian motion. We begin with examples in the 1-dimensional case and end with proving an existence and uniqueness theorem in the multivariate case.

4.1 Examples

In general stochastic differential equations cannot be solved analytically. This is not surprising. Indeed, ordinary differential equations cannot usually be solved analytically. However, in this section we give a few examples (and exercises) where analytical solutions can be found. In this subsection we only consider the 1-dimensional case.

Example 4.1 (Geometric Brownian motion)**.** Let W be a 1-dimensional Brownian motion and let μ and σ be constants. Consider the stochastic differential equation

$$\mathrm{d}S_t \quad = \quad \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}W_t. \tag{4.1}$$

We want to solve this by using Itô formula. An educated guess is to consider the function

$$f(t,x) = \exp\left\{ \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma x \right\}.$$

Then

$$\begin{aligned}
\frac{\partial f}{\partial t}(t,x) &= \left(\mu - \frac{\sigma^2}{2} \right) f(t,x), \\
\frac{\partial f}{\partial x}(t,x) &= \sigma f(t,x), \\
\frac{\partial^2 f}{\partial x^2}(t,x) &= \sigma^2 f(t,x).
\end{aligned}$$

Then, by Itô formula, we have that

$$\begin{aligned}
\mathrm{d}f(t,W_t) &= \frac{\partial f}{\partial t}(t,W_t)\,\mathrm{d}t + \frac{\partial f}{\partial x}(t,W_t)\,\mathrm{d}W_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,W_t)\,(\mathrm{d}W_t)^2 \\
&= \left(\mu - \frac{\sigma^2}{2} \right) f(t,W_t)\,\mathrm{d}t + \sigma f(t,W_t)\,\mathrm{d}W_t + \frac{\sigma^2}{2} f(t,W_t)\,\mathrm{d}t \\
&= \mu f(t,W_t)\,\mathrm{d}t + \sigma f(t,W_t)\,\mathrm{d}W_t.
\end{aligned}$$

So, we see that the process

$$S_t \quad = \quad f(t,W_t) \quad = \quad S_0 \exp\left\{ \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right\}$$

solves the stochastic differential equation (4.1).

Exercise 4.1. Solve the stochastic differential equation

$$\mathrm{d}X_t \quad = \quad X_t\,\mathrm{d}t + \mathrm{d}W_t.$$

Example 4.2 (Ornstein–Uhlenbeck Process)**.** Let $\theta > 0$ a constant. Let $\sigma \in \mathbb{R}$ be a constant. The Ornstein–Uhlenbeck process U is the solution of the Langevin equation

$$\mathrm{d}U_t \quad = \quad -\theta U_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \qquad U_0 = \xi. \tag{4.2}$$

To solve the Langevin equation (4.2) one can mimic he solution of linear ordinary differential equation. Indeed, let $\dot{W}_t = \frac{\mathrm{d}}{\mathrm{d}t}W_t$ be the (nonexistent) time derivative of the Brownian motion W. Then, at least informally, we can rewrite (4.2) as the linear equation

$$\frac{\mathrm{d}}{\mathrm{d}t}U_t \quad = \quad -\theta U_t + \sigma\dot{W}_t.$$

This is a linear first order non-homogeneous equation and classical theory of ordinary differential equations suggests the solution

$$U_t \quad = \quad \xi\mathrm{e}^{-\theta t} + \mathrm{e}^{-\theta t}\int_0^t \mathrm{e}^{\theta s}\sigma\dot{W}_s\,\mathrm{d}s.$$

Since, informally $\dot{W}_t\,\mathrm{d}t = \mathrm{d}W_t$, this suggests that the solution of (4.2) is

$$U_t \quad = \quad \xi\mathrm{e}^{-\theta t} + \sigma\int_0^t \mathrm{e}^{-\theta(t-s)}\,\mathrm{d}W_s. \tag{4.3}$$

This is indeed the solution.

Another way to find the solution is to use the Itô formula (or rather integration by parts formula) with the function

$$f(t,x) = \mathrm{e}^{\theta t}x.$$

Indeed,

$$\begin{aligned}
\mathrm{d}f(t,U_t) \quad &= \quad \mathrm{d}\left(\mathrm{e}^{\theta t}U_t\right) \\
&= \quad \mathrm{e}^{\theta t}\mathrm{d}U_t + \theta U_t\mathrm{e}^{\theta t}\,\mathrm{d}t \\
&= \quad \mathrm{e}^{\theta t}\left(-\theta U_t\,\mathrm{d}t + \sigma\mathrm{d}W_t + \theta U_t\,\mathrm{d}t\right) \\
&= \quad \mathrm{e}^{\theta t}\sigma\mathrm{d}W_t.
\end{aligned}$$

Integrating this differential gives us

$$\mathrm{e}^{\theta t}U_t \quad = \quad \xi + \int_0^t \mathrm{e}^{\theta s}\,\mathrm{d}W_s,$$

i.e, we obtain the solution (4.3).

Exercise 4.2 (Ornstein–Uhlenbeck with Drift)**.** Consider the Ornstein–Uhlenbeck process with constant drift μ:

$$\mathrm{d}V_t \quad = \quad \theta(\mu - V_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t. \tag{4.4}$$

Solve the stochastic differential equation (4.4).

4.2 Existence and Uniqueness

Let $W = (W^1, \dots, W^d)$ be a d-dimensional Brownian motion. Let $b\colon [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma\colon [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$. In this section we provide existence and uniqueness result for the stochastic differential equation

$$\mathrm{d}X_t \quad = \quad b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t \tag{4.5}$$

In this section $|\cdot|$ will denote the Euclidean norm $|b|^2 = \sum_j b_j^2$ for vectors and the Frobenius norm $|\sigma|^2 = \sum_j \sum_k \sigma_{jk}^2$ for matrices.

Remark 4.1. We note that if $\sigma \equiv 0$ then the stochastic differential equation (4.5) is just a deterministic differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} X_t^j \quad = \quad b_j(t, X_t), \qquad X_0^j = \xi_j,$$

$j = 1, \dots n$. The theory of ordinal differential equations suggests we need growth and Lipschitz conditions for b. Then the solution of the differential equation can be then constructed by using Picard iterations.

Assumption 4.1 (Growth and Lipschitz)**.** We assume the following growth condition for the coefficients b and σ

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \quad \leq \quad M_T\,(1 + |x|)^2 \tag{4.6}$$

for all $x \in \mathbb{R}^n$ and $t \in [0,T]$.

We also assume the following Lipschitz condition for the coefficients b and σ:

$$|b(t,x) - b(t,y)|^2 + |\sigma(t,x) - \sigma(t,y)|^2 \quad \leq \quad L_T |x - y|^2 \tag{4.7}$$

for all $x, y \in \mathbb{R}^n$ and and $t \in [0,T]$.

We begin by proving a stability result for the stochastic differential equations (4.5) that implies the uniqueness of the results.

Theorem 4.1 (Stability)**.** *Let X and $\tilde{X}$ be two solutions of (4.5) with $X_0 = \xi \in L^2(\mathbb{P})$ and $\tilde{X}_0 = \tilde{\xi} \in L^2(\mathbb{P})$. Assume that the Lipschitz condition (4.7) hold. Then*

$$\mathbb{E}\left[\sup_{t\leq T} |X_t - \tilde{X}_t|^2\right] \quad \leq \quad C_T \mathbb{E}\left[|\xi - \tilde{\xi}|^2\right].$$

Before going to the proof of Theorem 4.1 we give an auxiliary result called Grönwall inequality needed in the proof as an exercise.

Exercise 4.3 (Grönwall Lemma)**.** Suppose $u, a, b\colon [0,\infty) \to [0,\infty)$ satisfy

$$u(T) \quad \leq \quad a(T) + \int_0^T b(t) u(t)\,\mathrm{d}t.$$

Show that

$$u(T) \leq a(T) + \int_0^T a(t) b(t) \mathrm{e}^{\int_t^T b(s)\mathrm{d}s} \mathrm{d}t$$

Proof of Theorem 4.1. Let us first note that

$$X_t - \tilde{X}_t = \xi - \tilde{\xi} + \int_0^t \left[b(s, X_s) - b(s, \tilde{X}_s)\right] \mathrm{d}s + \int_0^t \left[\sigma(s, X_s) - b(s, \tilde{X}_s)\right] \mathrm{d}W_s$$

Since

$$(a + b + c)^2 \leq 3a^2 + 3b^2 + 3c^2,$$

we obtain

$$\begin{aligned} \mathbb{E}\left[\sup_{t\leq T} |X_t - \tilde{X}_t|^2\right] \leq\ & 3\mathbb{E}\left[|\xi - \tilde{\xi}|^2\right] \\ & +3\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t \left[b(s, X_s) - b(s, \tilde{X}_s)\right] \mathrm{d}s\right|^2\right] \\ & +3\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t \left[\sigma(s, X_s) - \sigma(s, \tilde{X}_s)\right] \mathrm{d}W_s\right|^2\right] \end{aligned}$$

Let us consider the second term above. The supremum is eliminated by noticing that

$$\sup_{t\leq T}\left|\int_0^t \left[b(s, X_s) - b(s, \tilde{X}_s)\right] \mathrm{d}s\right| \leq \int_0^T \left|b(s, X_s) - b(s, \tilde{X}_s)\right| \mathrm{d}s$$

Then we can use the Cauchy–Schwarz (or Jensen) inequality to obtain

$$\left(\int_0^T \left|b(s, X_s) - b(s, \tilde{X}_s)\right| \mathrm{d}s\right)^2 \leq T\int_0^T \left|b(s, X_s) - b(s, \tilde{X}_s)\right|^2 \mathrm{d}s$$

Finally, by using the Lipschitz condition (4.7) and taking expectations, we obtain

$$\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t \left[b(s, X_s) - b(s, \tilde{X}_s)\right] \mathrm{d}s\right|^2\right] \leq L_T^2 T\int_0^T \mathbb{E}\left[|X_s - \tilde{X}_s|^2\right] \mathrm{d}s$$

Let us consider the third term. Here we need the maximal inequality for stochastic integrals (Theorem 2.3) and the Lipschitz condition (4.7). We obtain

$$\begin{aligned} \mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t \left[\sigma(s, X_s) - \sigma(s, \tilde{X}_s)\right] \mathrm{d}W_s\right|^2\right] & \leq 4\int_0^T \mathbb{E}\left[\left|\sigma(s, X_s) - \sigma(s, \tilde{X}_s)\right|^2\right] \mathrm{d}s \\ & \leq 4L_T^2 \int_0^T \mathbb{E}\left[\left|X_s - \tilde{X}_s\right|^2\right| ds \end{aligned}$$

Plugging in the estimates for the second and third term we obtain

$$\begin{aligned} \mathbb{E}\left[\sup_{t\leq T} |X_t - \tilde{X}_t|^2\right] & \leq 3\mathbb{E}\left[|\xi - \tilde{\xi}|^2\right] + 3L_T^2(T + 4)\int_0^T \mathbb{E}\left[|X_s - \tilde{X}_s|^2\right] \mathrm{d}s \\ & \leq 3\mathbb{E}\left[|\xi - \tilde{\xi}|^2\right] + 3L_T^2(T + 4)\int_0^T \mathbb{E}\left[\sup_{r\leq s} |X_r - \tilde{X}_r|^2\right] \mathrm{d}s \end{aligned}$$

Now, we denote

$$
\begin{aligned}
u(T) &= \mathbb{E}\left[\sup_{t\le T}|X_t-\tilde{X}_t|^2\right], \\
a(t) &= 3\mathbb{E}\left[|\xi-\tilde{\xi}|^2\right], \\
b(t) &= 3L^2(T+4)
\end{aligned}
$$

and apply the Grönwall lemma of Exercise 4.3. The claim follows from this. □

Corollary 4.1 (Uniqueness)**.** *Suppose* (4.7) *holds. Then the solution of the stochastic differential equation* (4.5) *is unique.*

Let us then show the existence of the solution to (4.5) by using Picard iterations. Again, we leave some technical details as an exercise.

Exercise 4.4. Let $X^{(m)}$ be the Picard iteration defined as

$$
\begin{aligned}
X_t^{(0)} &= \xi, \\
X_t^{(m+1)} &= \xi + \int_0^t b(s, X_s^{(m)})\,\mathrm{d}s + \int_0^t \sigma(s, X_s^{(m)})\,\mathrm{d}W_s.
\end{aligned}
$$

(i) Assume the growth condition (4.6). Show, as in the proof of Theorem 4.1, that

$$
\mathbb{E}\left[\sup_{t\le T}|X_t^{(m+1)}-\xi|^2\right] \le C_T\mathbb{E}\left[\sup_{t\le T}\left(1+X_t^{(m)}\right)^2\right].
$$

(ii) Assume the Lipschitz condition (4.7). Show, as in the proof of Theorem 4.1, that

$$
\mathbb{E}\left[\sup_{t\le T}|X_t^{(m+1)}-X_t^{(m)}|^2\right] \le C_T\int_0^T \mathbb{E}\left[\sup_{u\le s}|X_u^{(m)}-X_u^{(m-1)}|^2\right]\mathrm{d}s.
$$

Theorem 4.2 (Existence)**.** *Assume the growth condition* (4.6) *and the Lipschitz conditions* (4.7) *hold. Then the stochastic differential equation* (4.5) *admits a solution* X*.*

Proof. We set the Picard iterations as

$$
\begin{aligned}
X_t^{(0)} &= \xi, \\
X_t^{(m+1)} &= \xi + \int_0^t b(s, X_s^{(m)})\,\mathrm{d}s + \int_0^t \sigma(s, X_s^{(m)})\,\mathrm{d}W_s.
\end{aligned}
$$

By Exercise 4.4 we have

$$
\mathbb{E}\left[\sup_{t\le T}|X_t^{(m+1)}-\xi|^2\right] \le C_T\mathbb{E}\left[\sup_{t\le T}\left(1+X_t^{(m)}\right)^2\right]. \tag{4.8}
$$

This shows that the iterations $X^{(m)}$ are well-defined.

Let us then consider the convergence of the iterations. By Exercise 4.4 we have

$$\mathbb{E}\left[\sup_{t\leq T}|X_t^{(m+1)} - X_t^{(m)}|^2\right] \;\leq\; C_T\int_0^T \mathbb{E}\left[\sup_{u\leq s}|X_u^{(m)} - X_u^{(m-1)}|^2\right]\mathrm{d}s. \tag{4.9}$$

Set

$$\alpha_m(T) \;=\; \mathbb{E}\left[\sup_{t\leq T}|X_t^{(m+1)} - X_t^{(m)}|^2\right].$$

Then, the inequality (4.9) takes the form

$$\alpha_m(T) \;\leq\; C_T\int_0^T \alpha_{m-1}(s)\mathrm{d}s.$$

Iterating this with m and noticing that $C_s \leq C_T$ we see that

$$\alpha_m(T) \;\leq\; c_T^m\frac{T^m}{m!}\alpha_0(T).$$

By the estimate (4.8) we have

$$\alpha_0(T) \;=\; C_T\mathbb{E}\left[(1+|\xi|)^2\right] < \infty.$$

Consequently, $\alpha_m(T)$, $m \in \mathbb{N}$, converges exponentially fast. This means that the uniform L^2 limit

$$X = \lim_{m\to\infty} X^{(m)}$$

exists in the sense that

$$\lim_{m\to\infty}\mathbb{E}\left[\sup_{t\leq T}|X_t - X_t^{(m)}|^2\right] \;=\; 0.$$

Finally, by using the dominated convergence theorem, we see that the limit X satisfies the stochastic differential equation (4.5). □

5 Itô Diffusions and Partial Differential Equations

5.1 Dynkin Formula

In this subsection we consider the homogneus n dimensional Itô diffusion as a Markov process and consider its generator and its connection to partial differential equations. A homogeneous Itô diffusion is the solution X of the following time-independent stochastic differential equation

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \qquad X_0 = x. \tag{5.1}$$

As before (5.1) should be understood as the componentwise integral equations

$$X_t^i = x_i + \int_0^t b_i(X_s)\,\mathrm{d}s + \sum_{k=1}^d \int_0^t \sigma_{ik}(X_s)\,\mathrm{d}W_s^k,$$

$i = 1, \ldots, n$. Here we have written x_i instead of ξ_i to emphasize the fact that the initial point $X_0 = x$ of the Itô diffusion is deterministic.

Recall that a process is Markovian intuitively if its past is independent of the future given the present. A rigorous definition is the following.

Definition 5.1 (Markov Process). A process $X = (X_t)_{t\geq 0}$ is a time-homogeneous strong Markov process if for all bounded Borel functions $f\colon \mathbb{R}^n \to \mathbb{R}$ and stopping times τ

$$\mathbb{E}\left[f(X_{\tau+h})\middle|\mathscr{F}_\tau^X\right] = \mathbb{E}^{X_\tau}\left[f(X_h)\right]$$

Here we have written $\mathbb{E}^{X_\tau} = \mathbb{E}[\,\cdot\,|X_\tau]$.

Proposition 5.1. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be bounded Borel function. Let τ be stopping time. Let X be given by* (5.1)*. Then*

$$\mathbb{E}^x\left[f(X_{\tau+h})\middle|\mathscr{F}_\tau^X\right] = \mathbb{E}^{X_\tau}\left[f(X_h)\right],$$

i.e., X is a time-homogeneous strong Markov process.

Proof. See Øksendal [4] Theorem 7.1.2 and Theorem 7.2.4 □

Since the Itô diffusion (5.1) is Markovian, its probability law can be described by the transition semigroup $(P_t)_{t\geq 0}$ given by

$$P_t f(x) = \mathbb{E}^x\left[f(X_t)\right],$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$ is smooth. The transition semigroup itself can be written by using its generator A formally as

$$P_t = \mathrm{e}^{tA},$$

where

$$Af(x) = \lim_{t\downarrow 0} \frac{P_t f(x) - f(x)}{t}.$$

Exercise 5.1. Let B be d-dimensional Brownian motion. Show that its generator is given by the Laplacian as

$$Af(x) \;=\; \frac{1}{2}\triangle f(x) \;=\; \frac{1}{2}\sum_{k=1}^{d}\frac{\partial^2 f}{\partial x_k^2}(x).$$

The generator of time-homogeneous Itô diffusion (5.1) is

$$Af(x) \;=\; \sum_{i=1}^{n} b_i(x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{d}(\sigma\sigma^\top)_{ik}(x)\frac{\partial^2 f}{\partial x_i \partial x_k}(x). \tag{5.2}$$

Indeed, the claim follows from the following Dynkin formula.

Theorem 5.1 (Dynkin Formula)**.** *Let X be the Itô diffusion given by (5.1). Let τ be stopping time such that $\mathbb{E}^x[\tau] < \infty$. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a bounded Borel function. Then*

$$\mathbb{E}^x\left[f(X_\tau)\right] \;=\; f(x) + \mathbb{E}^x\left[\int_0^\tau Af(X_s)\,\mathrm{d}s\right]. \tag{5.3}$$

Proof. By Itô formula (see Theorem 3.2 and Remark 3.3) we have

$$\begin{aligned}
\mathrm{d}f(X_t) &= \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(X_t)\,\mathrm{d}X_t^i + \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_j}(X_t)\,\mathrm{d}X_t^i\mathrm{d}X_t^j \\
&= \sum_{i=1}^{n} b_i(X_t)\frac{\partial f}{\partial x_i}(X_t)\,\mathrm{d}t + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\sigma\sigma^\top)_{ij}(X_t)\frac{\partial^2 f}{\partial x_j \partial x_j}(X_t)\,\mathrm{d}t \\
&\quad + \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(X_t)\sum_{k=1}^{d}\sigma_{ik}(X_t)\mathrm{d}W_t^k \\
&= Af(X_t)\mathrm{d}t + \text{ martingale.}
\end{aligned}$$

The claim follows by integrating and taking expectations. □

Exercise 5.2. Let U be the 1-dimensional Ornstein–Uhlenbeck process, i.e.,

$$\mathrm{d}U_t \;=\; -\theta U_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t.$$

Find its generator.

We end this section by giving a Monte Carlo type of method for solving Dirichlet boundary problem for partial differential equations.

Example 5.1. Suppose we have a domain $D \subset \mathbb{R}^n$ and given boundary data $f(x) = \varphi(x)$ on ∂D and we assume that $Af(x) = 0$ on D. Here A is the differential operator (5.2). Then Dynkin formula (5.3) gives us immediately a way to simulate the solution of the equation $Af(x) = 0$. Indeed, we have

$$f(x) \;=\; \mathbb{E}^x\left[\varphi(X_{\tau_D})\right],$$

where τ_D is the first time the Itô diffusion X exits the domain D.

5.2 Feynman–Kac Formula

In this subsection we consider time-dependent n-dimensional Itô diffusion

$$\mathrm{d}X_t = b(t,X_t)\,\mathrm{d}t + \sigma(t,X_t)\,\mathrm{d}W_t, \qquad X_0 = x \tag{5.4}$$

and its connection to deterministic partial differential equations. The generator of such time-inhomogeneous Itô diffusion is

$$Af(x) = \sum_{j=1}^{n} b_j(t,x)\frac{\partial f}{\partial x_j}(x) + \frac{1}{2}\sum_{k=1}^{d}\sum_{j=1}^{n}(\sigma\sigma^\top)_{jk}(t,x)\frac{\partial^2 f}{\partial x_j \partial x_k}(x).$$

Theorem 5.2 (Feynman–Kac Formula)**.** *Let $u = u(t,x)$, $V = V(t,x)$ and $f = f(t,x)$ be real-valued. Consider the backward partial differential equation*

$$\frac{\partial u}{\partial t} + Au - Vu + f = 0 \tag{5.5}$$

with boundary condition $u(T,x) = \phi(x)$. The solution of (5.5) can be written as

$$u(t,x) = \mathbb{E}^{t,x}\left[\int_t^T \mathrm{e}^{-\int_t^r V(u,X_u)\mathrm{d}u} f(r,X_r)\,\mathrm{d}r + \mathrm{e}^{-\int_t^T V(u,X_u)\mathrm{d}u}\phi(X_T)\right], \tag{5.6}$$

where $\mathbb{E}^{x,t} = \mathbb{E}[\,\cdot\,|X_t = x]$ and X is the Itô diffusion (5.4).

Proof. We only prove that *if* the solution to (5.5) exists, then it is necessarily of the form (5.6).

The key idea is to use integration-by-parts and Itô formula to the process ($s \geq t$)

$$Y_s = \mathrm{e}^{-\int_t^s V(u,X_u)\mathrm{d}u} u(s,X_s) + \int_t^s \mathrm{e}^{-\int_t^r V(u,X_u)\mathrm{d}u} f(r,X_r)\,\mathrm{d}r \tag{5.7}$$

After some simplifications that are left as an exercise, we obtain

$$\begin{aligned}\mathrm{d}Y_s &= \mathrm{e}^{-\int_t^s V(u,X_u)\mathrm{d}u}\left(-V(s,X_s)u(s,X_s) + f(s,X_s) + \frac{\partial u}{\partial s}(s,X_s) + Au(s,X_s)\right)\mathrm{d}s \\ &\quad + \mathrm{e}^{-\int_t^s V(u,X_u)\mathrm{d}u}\sigma(s,X_s)\frac{\partial u}{\partial x}(s,X_s)\,\mathrm{d}W_s. \end{aligned} \tag{5.8}$$

Since the terms in the parentheses sum up to zero, we obtain

$$\mathrm{d}Y_s = \mathrm{e}^{-\int_t^s V(u,X_u)\mathrm{d}u}\sigma(s,X_s)\frac{\partial u}{\partial x}(s,X_s)\,\mathrm{d}W_s$$

Now, by integrating we obtain

$$Y_T - Y_t = \int_t^T \mathrm{e}^{-\int_t^s V(u,X_u)\mathrm{d}u}\sigma(s,X_s)\frac{\partial u}{\partial x}(s,X_s)\,\mathrm{d}W_s.$$

Taking conditonal expectations we obtain, since the right hand side above is an Itô integral,

$$\mathbb{E}\left[Y_T|X_t = x\right] = \mathbb{E}\left[Y_t|X_t = x\right] = u(t,x),$$

which finishes the proof. □

Exercise 5.3. Finalize the proof of Theorem 5.2, i.e., let Y be given by (5.7). Show that (5.8) holds.

Exercise 5.4 (Black–Scholes Partial Differential Equation)**.** Let $V = V(t,x)$. Let r and σ be constants. Consider the so-called Black–Scholes partial differential equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2} + rx\frac{\partial V}{\partial x} - rV \;=\; 0,$$

with boundary condition $V(T,x) = (x-K)^+$. Show that the solution of this partial differential equation is

$$V(t,x) \;=\; N(d_1)x + N(d_2)K\mathrm{e}^{-r(T-t)},$$

where

$$N(z) \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} \mathrm{e}^{-\frac{1}{2}w^2}\,\mathrm{d}w$$

and

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{x}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right], \\ d_2 &= d_1 - \sigma\sqrt{T-t}. \end{aligned}$$

References

[1] A. N. Borodin and P. Salminen, *Handbook of Brownian motion—facts and formulae*, Probability and its Applications, Birkhäuser Verlag, Basel, second ed., 2002.

[2] H. Föllmer, *Calcul d'Itô sans probabilités*, in Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French), vol. 850 of Lecture Notes in Math., Springer, Berlin, 1981, pp. 143–150.

[3] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, vol. 113 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 1991.

[4] B. Øksendal, *Stochastic differential equations*, Universitext, Springer-Verlag, Berlin, sixth ed., 2003. An introduction with applications.

[5] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, vol. 293 of Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, third ed., 1999.

[6] R. L. Schilling and L. Partzsch, *Brownian motion*, De Gruyter, Berlin, 2012. An introduction to stochastic processes, With a chapter on simulation by Björn Böttcher.

[7] D. Sondermann, *Introduction to stochastic calculus for finance*, vol. 579 of Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, 2006. A new didactic approach.

[8] D. Williams, *Probability with martingales*, Cambridge Mathematical Textbooks, Cambridge University Press, Cambridge, 1991.