Stochastic Analysis

King's College London
Version 1.4
October 2015

Markus Riedle

Preface

These notes are for a course on *Stochastic Analysis* at King's College London. Given the limited time and diverse background of the audience we will only consider stochastic integration with respect to Brownian motion. However in particular for application to Financial Mathematics, this is sufficient to study a wide range of models and to understand the major tools such as Girsanov's Theorem and Feynman-Kac formula.

The notes are intended to serve as an intermediate step to more advanced books such as the monographs by Karatzas and Shreve [9], Protter [17] or Revuz and Yor [18] among many others. The monographs by Klebaner [10], Kuo [12], Mikosch [13] and Oksendal [16] might be considered to be at a comparable level as these notes. One can also find some excellent online sources for these courses, such as the lecture notes [14], [21] and [23].

The mathematics for stochastic analysis can be quite technical, in particular if one is confronted with this part of probability theory the first time. In order to avoid that students get lost in these technical details some parts of the notes are written in small print. Moreover, due to the limited time, only a vanishing part of the theory can be presented here, and sometimes I cannot resist to say a little more, which is then also written in small print. In any case, text which is written in small print is not examinable.

Each chapter finishes with some exercises where the reader can apply the studied theory and results (and which is essential to understand the mathematics). The exercises are classified (very subjective) into the following categories:

- very basic;
- requires a bit of thinking/understanding, marked with $(*)$.
- slightly difficult, marked with $(**)$.
- rather straightforward but the topic and tools are not in the focus, marked with $(\natural)$.
- slightly difficult and not in the focus of these notes, marked with $(\sharp)$.

The symbols appear at the end of subquestions they are refereing to. All questions are examinable except those marked with $(\natural)$ or $(\sharp)$. Nevertheless, also these questions marked with $(\natural)$ or $(\sharp)$ might help to understand the content better, and thus they also might help to pass the exam.

The first three chapters are presented in a different order in classes in order to make it hopefully more interesting. Here we follow roughly the following order:

Definition 3.0.1 (Ex. 3.4.1)
Section 1.1 (Ex. 1.3.1 - 1.3.5, 1.3.8 - 1.3.10)
Proposition 3.2.1 (Ex. 3.4.1 - 3.4.6)
Proposition 3.2.2
Section 2.1 (Ex. 2.5.1 - 2.5.3, Ex. 2.5.6- 2.5.7)
Section 2.2
Theorem 3.2.3
Corollary 3.2.4
Section 1.2 (Ex. 1.3.6, 1.3.7)
Section 2.3 (Ex. 2.5.4, 2.5.5)
Section 3.3 (Ex. 3.4.7, 3.4.8

$\vdots$ (much later)
Section 2.4 (Ex. 2.5.8)

The numbers in bracket indicate the exercises you can tackle after this part of the course.

These notes benefited from the comments of a view students (and each comment is still welcome!). In particular, I want to mention here and thank very much Rasmus Søndergaard Pedersen (LGS) and Tomas Restrepo-Saenz (King's College).

Contents

1
The Tool: Stochastic Processes

Note: In the lecture classes the first three chapters of these notes are presented in a different order than in this printed version. The order can be found in the preface.

Let $(\Omega, \mathscr{A}, P)$ be a probability space. The integer $d \in \mathbb{N}$ is fixed and denotes the dimension of the underlying space $\mathbb{R}^d$. The Borel σ-algebra is denoted by $\mathfrak{B}(\mathbb{R}^d)$.

1.1. Some definitions

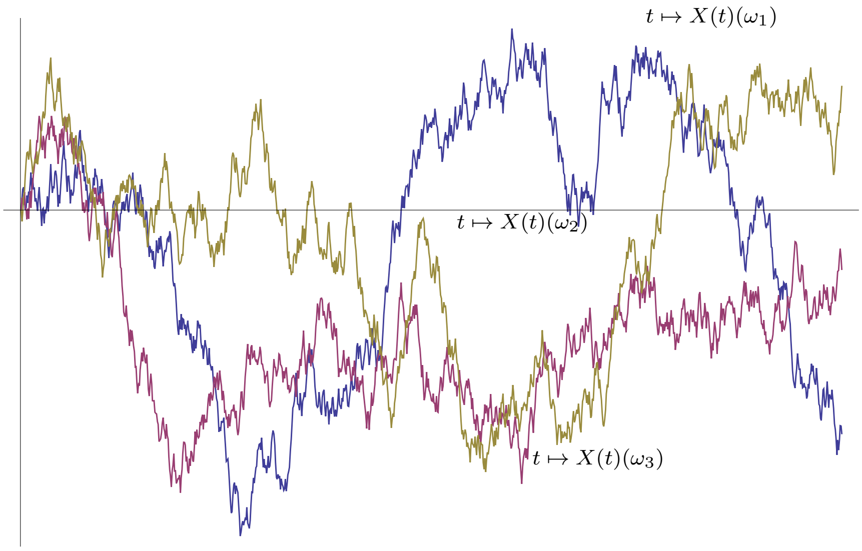
Recall that a random variable or random vector is a measurable mapping $X : \Omega \to \mathbb{R}^d$ with respect to $\mathscr{A}$ and $\mathfrak{B}(\mathbb{R}^d)$. In the case $d \geqslant 2$ the random variable X is also called random vector.

Definition 1.1.1. *Let I be a subset of $[0, \infty)$. A* stochastic process with values in $\mathbb{R}^d$ *is a family $(X(t) : t \in I)$ of random variables $X(t) : \Omega \to \mathbb{R}^d$ for each $t \in I$.*

In this course we most often consider stochastic processes in *continuous time*, that is $I = [0, T]$ for a constant $T > 0$ or $I = [0, \infty)$. If $I \subseteq \mathbb{N}_0$ then we say that $(X(t) : t \in I)$ is a stochastic process in *discrete time*. Both notations $X(t)$ and X_t are used in the literature but in this course the latter is typically used for stochastic processes in discrete time.

There are at least two different perspectives on stochastic processes:

- for each fixed $t \in I$ the object $X(t) : \Omega \to \mathbb{R}^d$ is a random variable and the stochastic process $(X(t) : t \in I)$ might be considered as an ordered family of random variables;
- for each fixed $\omega \in \Omega$ the collection $\{X(t)(\omega) : t \geqslant 0\}$ is a function
$$t \mapsto X(t)(\omega).$$
This mapping is called a *path* or a *trajectory* of X.



three paths of a stochastic process

If P-almost all (short: P-a.a.) paths of a stochastic process have a certain property, then we describe the stochastic process by this property, e.g. a continuous stochastic process $(X(t) : t \geqslant 0)$ means that for P-a.a. $\omega \in \Omega$ its trajectories $t \mapsto X(t)(\omega)$ are continuous. Here for P-a.a. $\omega \in \Omega$ means that there exists a set $\Omega_0 \in \mathscr{A}$ with $P(\Omega_0) = 1$ such that the property holds for all $\omega \in \Omega_0$, e.g. the trajectories $t \mapsto X(t)(\omega)$ are continuous for all $\omega \in \Omega_0$.

Example 1.1.2.

(a) Let $X_1, X_2, \dots$ be independent, identically distributed random variables with

$$P(X_1 = 1) = p \qquad P(X_1 = -1) = 1 - p,$$

for some fixed value $p \in (0,1)$. Define for each $t \geqslant 0$

$$R(t) := \begin{cases} 0, & \text{if } t \in [0,1), \\ X_1 + \dots + X_{[t]}, & \text{if } t \geqslant 1, \end{cases}$$

where $[t]$ denotes the largest integer smaller than t. It follows that $(R(t) : t \geqslant 0)$ is a stochastic process, the so-called *random walk*. Often this stochastic process is considered in discrete time by $(R_k : k \in \mathbb{N}_0)$ with $R_k := X_1 + \dots + X_k$ for $k \in \mathbb{N}$ and $R_0 = 0$.

(b) Let $X_1, X_2, \dots$ be independent, identically distributed random variables with exponential distribution with parameter $\lambda > 0$, that is

$$P(X_1 \leqslant x) = \begin{cases} 1 - e^{-\lambda x}, & \text{if } x \geqslant 0, \\ 0, & \text{else} \end{cases}.$$

Define $S_0 := 0$ and $S_n := X_1 + \cdots + X_n$ for all $n \in \mathbb{N}$. It follows that

$$N(t) := \begin{cases} 0, & \text{if } t = 0, \\ \max\{k \in \{0,1,2,\dots\} : S_k \leqslant t\}, & \text{if } t > 0, \end{cases}$$

defines a stochastic process $(N(t) : t \geqslant 0)$, the *Poisson process with intensity* $\lambda > 0$.

Poisson processes often model the occurrences of a sequence of discrete events, if the time intervals between successive events are exponentially distributed.

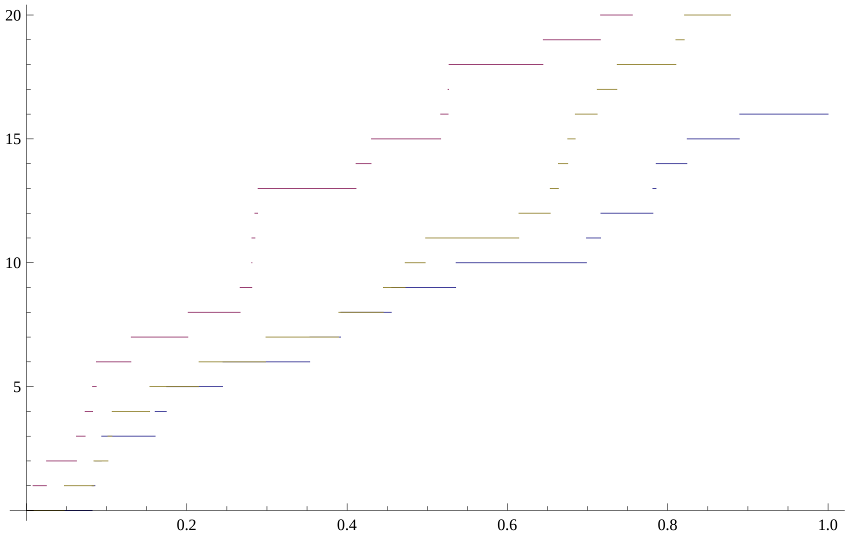
(c) A stochastic process $(X(t) : t \geqslant 0)$ is called *Gaussian* if for every $0 \leqslant t_1 \leqslant \ldots \leqslant t_n$ and $n \in \mathbb{N}$ the random vector

$$Z := \big(X(t_1), \dots, X(t_n)\big) : \Omega \to \mathbb{R}^n$$

is normally distributed in $\mathbb{R}^n$. In this case, the distribution of the random vector Z is characterised by

$$\Big(E[X(t_1)], \dots, E[X(t_n)]\Big) \quad \text{and} \quad \Big(\operatorname{Cov}(X(t_i), X(t_j))\Big)_{i,j=1}^{n}.$$

In contrast to the other two examples it is not clear if Gaussian stochastic processes exist!



three paths of a Poisson process

Since the paths of a stochastic process are random it is not obvious which processes are considered to be the same. In fact, there are two different notions for the equivalence of two stochastic processes.

Definition 1.1.3. *Let* $X = (X(t) : t \in I)$ *and* $Y = (Y(t) : t \in I)$ *be two stochastic processes.*

(a) *The stochastic processes X and Y are called a* modification of each other *if*

$$P\big(X(t) = Y(t)\big) = 1 \qquad \textit{for all } t \in I.$$

(b) *The stochastic processes X and Y are called* indistinguishable

$$P\big(X(t) = Y(t) \textit{ for all } t \in I\big) = 1.$$

In order that the definition of indistinguishable for stochastic processes in continuous time makes sense it must be true that

$$\{X(t) = Y(t) \quad \text{for all } t \geqslant 0\} \in \mathscr{A}.$$

Since for a general probability space $(\Omega, \mathscr{A}, P)$ this need not to be true this requirement is implicitly part of the definition.

It follows directly form the definition that if two stochastic processes X and Y are indistinguishable then they are also modification of each other. For stochastic processes in continuous time with continuous paths also the converse direction is true (see Exercise 1.3.1). In general this is not true as the following example shows.

Example 1.1.4. Let $\Omega = [0, \infty)$, $\mathscr{A} = \mathfrak{B}([0, \infty))$ and P be a probability measure on $\mathscr{A}$ which has a density. Define two stochastic processes $(X(t) : t \geqslant 0)$ and $(Y(t) : t \geqslant 0)$ by

$$X(t)(\omega) = \begin{cases} 1, & \text{if } t = \omega, \\ 0, & \text{otherwise,} \end{cases} \qquad Y(t)(\omega) = 0 \qquad \text{for all } t \geqslant 0 \text{ and all } \omega \in \Omega.$$

Then X and Y are modification of each other but X and Y are not indistinguishable.

In typical applications such as financial mathematics or physics, a stochastic process models the evolution of a particle or a share price in time. If the stochastic process is observed at a fixed time t in such models, then only its values at time t and prior to time t are known. Thus, at a time t only some specific information is available and the amount of information increases as the time increases. Mathematically, the amount of information available at different times is modelled by a filtration:

Definition 1.1.5. *A family $\{\mathscr{F}_t\}_{t \in I}$ of σ-algebras $\mathscr{F}_t \subseteq \mathscr{A}$ for all $t \in I$ with*

$$\mathscr{F}_s \subseteq \mathscr{F}_t \qquad \textit{for all } 0 \leqslant s \leqslant t$$

is called a filtration*.*

Definition 1.1.6. *A stochastic process $X := (X(t) : t \in I)$ is called* adapted with respect to a filtration *$\{\mathscr{F}_t\}_{t \in I}$ if $X(t)$ is $\mathscr{F}_t$-measurable for every $t \in I$.*

Then there is no ambiguity we sometimes say that a stochastic process is adapted without mentioning the underlying filtration explicitly.

You might think of a filtration as the description of the information available at different times. The σ-algebra $\mathscr{F}_t$ of a filtration $\{\mathscr{F}_t\}_{t \in I}$ represents the information which is available at time t. One can think that the random outcome $\omega \in \Omega$ is already specified but we are only told at time t for all sets in the σ-algebra $\mathscr{F}_t$ whether this ω is in the set or not. The more sets there are in $\mathscr{F}_t$, the more information we obtain of an $\mathscr{F}_t$-measurable random variable. If $(X(t) : t \geqslant 0)$ is an adapted stochastic process this means that the random function $t \mapsto X(t)(\omega)$ is already specified on the interval $[0, \infty)$ (by fixing $\omega \in \Omega$) but we know at time s only the values of the function on the interval $[0, s]$ but not on (s, ∞).

Example 1.1.7. Let $X := (X(t) : t \geqslant 0)$ be an $\{\mathscr{F}_t\}_{t\geqslant 0}$-adapted stochastic process in the following examples.

(a) The set $A = \{\omega \in \Omega : X(s)(\omega) \geqslant 29.4 \text{ for all } s \leqslant 3.14\}$ is an element in $\mathscr{F}_{3.14}$. The set A is even in each $\mathscr{F}_s$ for $s \geqslant 3.14$.

(b) If X has continuous trajectories then the stochastic process $Y := (Y(t) : t \geqslant 0)$ defined by
$$Y(t) := \sup_{s\in[0,t]} |X(s)|$$
is adapted to the same filtration as X is.[1]

(c) The stochastic process $Y := (Y(t) : t \geqslant 0)$ defined by
$$Y(t) := \sup_{s\in[0,t+1]} |X(s)|$$
is not adapted to the same filtration as X is (if we do not consider a very pathological situation).

Most often one assumes the filtration which is generated by the process X itself, that is at a time t the σ-algebra $\mathscr{F}_t$ contains all information which is decoded by X restricted to the interval $[0, t]$. This is described more formally by
$$\begin{aligned} \mathscr{F}_t^X : &= \sigma\Big(X(s) : s \in [0,t]\Big) \\ &:= \sigma\Big(\big(X(s)\big)^{-1}([a,b]) : s \in [0,t], -\infty < a \leqslant b < \infty\Big), \end{aligned}$$
which means "the smallest σ-algebra" containing all the preimages[2]
$$\big(X(s)\big)^{-1}([a,b]) := \{\omega \in \Omega : X(s)(\omega) \in [a,b]\}$$
for all $s \in [0, t]$ and all $a, b \in \mathbb{R}$.

Example 1.1.8. Let $\Omega = \{1, 2, 3\}$, $\mathscr{A} = \mathscr{P}(\Omega)$ and $P(\{\omega\}) = \frac{1}{3}$ for each $\omega \in \Omega$. Define a stochastic process $(X(t) : t \geqslant 0)$ by $X(t)(\omega) = \max\{t - \omega, 0\}$. Then the filtration generated by the stochastic process X computes as
$$\mathscr{F}_t^X = \begin{cases} \{\emptyset, \Omega\}, & \text{if } t \in [0,1], \\ \{\emptyset, \Omega, \{1\}, \{2,3\}\}, & \text{if } t \in (1,2], \\ \mathscr{P}(\Omega), & \text{if } t > 2. \end{cases}$$

[1] Why do I require that X has continuous paths?

[2] note, I really mean preimage and not the inverse of a function

We end this subsection with some more technical definitions on filtration and measurability of stochastic processes. These notions will not be in the center of our attention but they are necessary to present the results in these notes correctly.

Let $(X(t) : t \geqslant 0)$ be a stochastic process and $\{\mathscr{F}_t^X\}_{t\geqslant 0}$ the generated filtration. For some technical reasons, see for example Exercise 1.1.10, we have to enlarge the filtration by the set of the so-called *null-sets*:

$$\mathscr{N} := \{N \subseteq \Omega : \exists B \in \mathscr{A} \text{ such that } N \subseteq B,\, P(B) = 0\}.$$

Define the *augmented*[3] *filtration* $\{\mathfrak{F}_t^X\}_{t\geqslant 0}$ *generated by* X by

$$\mathfrak{F}_t^X := \sigma(\mathscr{F}_t^X \cup \mathscr{N}) \qquad \text{for all } t \geqslant 0.$$

Thus, each σ-algebra $\mathfrak{F}_t^X$ contains all null-sets $N \in \mathscr{N}$.

Apart from including the set $\mathscr{N}$ in the filtration, often the so-called right-continuity of a filtration is required. For example, this plays later an important role in part (b) of Proposition 1.2.7.

Definition 1.1.9. *A filtration* $\{\mathscr{F}_t\}_{t\geqslant 0}$ *satisfies the* usual conditions *if*

(a) $\mathscr{N} \subseteq \mathscr{F}_0$. *(completion)*

(b) $\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s$ *for all* $t \geqslant 0$. *(right-continuity)*

Example 1.1.10. In Example 1.1.4 the stochastic processes X and Y are modifications of each other. But although the stochastic process Y is adapted to the generated filtration $\{\mathscr{F}_t^Y\}_{t\geqslant 0}$ the modification X is not adapted to the same filtration.

This problem can not occur if the filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ is required to be complete. Then, if X and Y are modification of each other it follows for every $B \in \mathfrak{B}(\mathbb{R})$ and $t \geqslant 0$

$$\{Y(t) \in B\} = \{X(t) \in B\} \backslash \{Y(t) \notin B, X(t) \in B\} \cup \{Y(t) \in B, X(t) \notin B\}.$$

Completeness of the filtration guarantees for all $t \geqslant 0$ that

$$\{Y(t) \notin B, X(t) \in B\},\, \{Y(t) \in B, X(t) \notin B\} \in \mathscr{F}_t,$$

since both sets are subsets of the set $\{X(t) \neq Y(t)\}$.

Definition 1.1.11. *The stochastic process* $X := (X(t) : t \geqslant 0)$ *is called* measurable *if the mapping*

$$\mathbb{R}_+ \times \Omega \to \mathbb{R}^d, \qquad (t, \omega) \mapsto X(t)(\omega)$$

is measurable with respect to $\mathfrak{B}(\mathbb{R}_+) \otimes \mathscr{A}$ *and* $\mathfrak{B}(\mathbb{R}^d)$.

[3] Oxford dictionary: augmented: adjective 1) having been made greater in size or value

One of the reasons to require the measurability is that we often want to define the time integral of a stochastic process and we want to take the expectation of the resulting new random variable. More specifically, if the stochastic process $(X(t) : t \geqslant 0)$ has trajectories $s \mapsto f_\omega(s) := X(s)(\omega)$ which are integrable functions $f_\omega : \mathbb{R}_+ \to \mathbb{R}$ for all $\omega \in \Omega$ then

$$Y(t) := \int_0^t X(s)\, ds$$

defines a random variable $Y(t) : \Omega \to \mathbb{R}$ for each $t \geqslant 0$. If in addition the stochastic process X is measurable then Fubini's theorem can be applied and it follows that

$$E\left[\int_0^t |X(s)|\, ds\right] = \int_0^t E[|X(s)|]\, ds.$$

1.2. Stopping Times

Definition 1.2.1. *A random variable* $\tau : \Omega \to [0, \infty]$ *is called a* stopping time of the filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ *if*

$$\{\tau \leqslant t\} \in \mathscr{F}_t \qquad \text{for all } t \geqslant 0.$$

Example 1.2.2.

(a) Every constant $\tau(\omega) := c$ for all $\omega \in \Omega$ for a constant $c \geqslant 0$ is a stopping time.

(b) An important example is the hitting time of a constant $a \in \mathbb{R}$ for a stochastic process $(X(t) : t \geqslant 0)$:

$$\tau_a : \Omega \to [0, \infty], \qquad \tau_a(\omega) := \inf\{t \geqslant 0 : X(t)(\omega) = a\}.$$

This mapping is not always a stopping time but in Proposition 1.2.7 we give conditions under which this is true.

(c) Let $(X(t) : t \geqslant 0)$ be an adapted stochastic process with continuous paths and let $A \subseteq \mathbb{R}$ be a set. A typical example of a random variable which is not a stopping time is

$$Y : \Omega \to [0, \infty], \quad Y(\omega) := \sup\{t \geqslant 0 : X(t)(\omega) \in A\}.$$

This random variable describes the last time that the stochastic process X visits the set A.

Example 1.2.3. In the setting of Example 1.1.8 define

$$\tau : \Omega \to [0, \infty), \qquad \tau(\omega) := \inf\{t \geqslant 0 : X(t)(\omega) > 0\}.$$

Then τ is not a stopping time as $\{\tau \leqslant 1\} = \{1\}$ but $\{1\} \notin \mathscr{F}_1$.

But τ is a so-called optional time. A mapping $\tau : \Omega \to [0, \infty)$ is called an *optional time* if $\{\tau < t\} \in \mathscr{F}_t$ for all $t \geqslant 0$.

Financial Mathematics 1. American options do not have a fixed exercise time, the holder can exercise an American option at any time agreed at conclusion of the contract. By means of stopping times the value of an American option can be defined and in some cases calculated explicitly. Let the stochastic process $(S(t) : t \in [0,T])$ model a share price. Then the American call option with strike price K can be described by the stochastic process $(A(t) : t \in [0,T])$ for $A(t) := \max\{S(t) - K, 0\}$ and its (non-discounted) value at time t is given by

$$V(t) = \sup_{\tau \in \Upsilon} E[A(\tau)|\mathscr{F}_t^S],$$

where $\Upsilon := \{\tau : \Omega \to [0,T] : \tau \text{ is stopping time with respect to } \{\mathscr{F}_t^S\}_{t\in[0,T]}\}$.

Another application of stopping times in Financial Mathematics are default times in credit risk models. Here, typically

$$\tau := \inf\{t \geqslant 0 : V(t) \leqslant 0\} \quad \text{or} \quad \tau := \inf\{t \geqslant 0 : V(t) \leqslant c\},$$

where the stochastic process $(V(t) : t \geqslant 0)$ models the firm value and c is a safety barrier.

Theorem 1.2.4. *If τ and φ are stopping times then*

(a) *$\tau + \varphi$ is a stopping time;*

(b) *$\tau \wedge \varphi := \min\{\tau, \varphi\}$ is a stopping time;*

(c) *$\tau \vee \varphi := \max\{\tau, \varphi\}$ is a stopping time;*

Proposition 1.2.5. *Let $(X(t) : t \geqslant 0)$ be a stochastic process with continuous paths and adapted to a filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ and let τ be a stopping time with respect to the same filtration.*

(a) *The mapping*

$$X(\tau) : \Omega \to \mathbb{R}, \qquad X(\tau)(\omega) := \begin{cases} X(\tau(\omega))(\omega), & \text{if } \tau(\omega) < \infty, \\ 0, & \text{else} \end{cases}$$

defines a random variable.

(b) *The random variables*

$$X(t \wedge \tau) : \Omega \to \mathbb{R}, \qquad X(t \wedge \tau)(\omega) = \begin{cases} X(t)(\omega), & \text{if } t \leqslant \tau(\omega), \\ X(\tau(\omega))(\omega), & \text{else}, \end{cases}$$

form a continuous stochastic process $(X(t \wedge \tau) : t \geqslant 0)$ adapted to $\{\mathscr{F}_t\}_{t\geqslant 0}$.

Proof. See [9, Prop. I.2.18]. □

The stochastic process $(X(t \wedge \tau) : t \geqslant 0)$ defined in Proposition 1.2.5 is called a *stopped process*.

Example 1.2.6. Let $(X(t) : t \geqslant 0)$ be a continuous stochastic process and define for a constant $a \in \mathbb{R}$

$$\tau_a : \Omega \to [0, \infty], \qquad \tau_a(\omega) := \inf\{t \geqslant 0 : X(t)(\omega) = a\}.$$

Then we have $X(\tau) = a$.

An important example of a stopping time is provided by the following result. Note that we employ a standard convention that the infimum of the empty set is infinity.

Proposition 1.2.7. *Let $(X(t) : t \geqslant 0)$ be an $\{\mathscr{F}_t\}_{t\geqslant 0}$-adapted stochastic process with continuous paths and define for $A \in \mathfrak{B}(\mathbb{R})$ the random time*

$$\tau_A : \Omega \to [0, \infty], \qquad \tau_A(\omega) := \inf\{t \geqslant 0 : \, X(t)(\omega) \in A\}.$$

(a) *If A is a closed set then τ_A is a stopping time of $\{\mathscr{F}_t\}_{t\geqslant 0}$.*

(b) *If A is an open set and the filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ satisfies the usual condition then τ_A is a stopping time of $\{\mathscr{F}_t\}_{t\geqslant 0}$.*

Proposition 1.2.7 clarifies part (b) in Example 1.2.2: if X is a continuous, stochastic process then the hitting time

$$\tau_a := \inf\{t \geqslant 0 : \, X(t) = a\}$$

of a constant $a \in \mathbb{R}$ is a stopping time of the filtration $\{\mathscr{F}_t^X\}_{t\geqslant 0}$. This is due to the simple fact that $\{a\}$ is a closed set.

Part (b) of Proposition 1.2.7 is true under much more generality, see Début Theorem.

Example 1.2.8. In Example 1.2.3, the random time τ is of the form τ_A for $A = (0, \infty)$ as defined in Proposition 1.2.7. However, τ is not a stopping time as showed in Example 1.2.3, and Proposition 1.2.7 cannot be applied since the filtration is not right-continuous.

1.3. Exercises

1. Let $X = (X(t) : t \in I)$ and $Y = (Y(t) : t \in I)$ be two stochastic processes.
   (a) Show that if X and Y are indistinguishable then they are modification of each other.
   (b) Assume that $I = [0, \infty)$ and that X and Y have continuous trajectories. Show that if X and Y are modifications then they are also indistinguishable.

2. Let T be a non-negative random variable with a density and define two stochastic processes X and Y by

$$X(t)(\omega) = 0, \qquad Y(t)(\omega) = \begin{cases} 0, & \text{if } T(\omega) \neq t, \\ 1, & \text{if } T(\omega) = t, \end{cases} \qquad \text{for all } \omega \in \Omega, t \geqslant 0.$$

   Show that X is a modification of Y but that X and Y are not indistinguishable.

3. Let $(Y(t) : t \geqslant 0)$ be a modification of $X(t) : t \geqslant 0)$. Show that then the finite-dimensional distributions coincide:

$$P(X(t_1) \in B_1, \dots, X(t_n) \in B_n) = P(Y(t_1) \in B_1, \dots, Y(t_n) \in B_n)$$

for all $t_1, \dots, t_n \geqslant 0$ and $B_1, \dots, B_n \in \mathscr{A}$. (∗)

4. (from [16]). Let $\Omega = \{1, 2, 3, 4, 5\}$.
   (a) Find the smallest σ-algebra $\mathscr{C}$ containing

   $$\mathscr{S} := \{\{1, 2, 3\}, \{3, 4, 5\}\}.$$

   (b) Is the random variable $X : \Omega \to \mathbb{R}$ defined by

   $$X(1) = X(2) = 0, \quad X(3) = 10, \quad X(4) = X(5) = 1$$

   measurable with respect to $\mathscr{C}$?
   (c) Find the σ-algebra $\mathscr{D}$ generated by $Y : \Omega \to \mathbb{R}$ and defined by

   $$Y(1) = 0, \quad Y(2) = Y(3) = Y(4) = Y(5) = 1.$$

5. Let $\Omega = [0, 1]$ and $\mathscr{A} = \mathfrak{B}([0, 1])$. Define a stochastic process $(X_n : n \in \mathbb{N})$ by

   $$X_n(\omega) := 2\omega \mathbb{1}_{[0,1-\frac{1}{n}]}(\omega).$$

   Show that the generated filtration $\{\mathscr{F}_n^X\}_{n\in\mathbb{N}}$ is given by (∗)

   $$\mathscr{F}_n^X = \Big\{ A \cup B : A \in \mathfrak{B}\left((0, 1 - \tfrac{1}{n}]\right),\, B \in \big\{\emptyset,\, \{0\} \cup (1 - \tfrac{1}{n}, 1]\big\}\Big\}$$

6. Show part (c) of Theorem 1.2.4.

7. Let σ, τ be stopping times with respect to a filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$.
   (a) Show that

   $$\mathscr{F}_\tau := \{A \in \mathscr{A} : A \cap \{\tau \leqslant t\} \in \mathscr{F}_t \quad \text{for all } t \geqslant 0\}.$$

   is a σ-algebra;
   (b) Show that if $\sigma \leqslant \tau$ then the σ-algebras defined in (a) satisfy $\mathscr{F}_\sigma \subseteq \mathscr{F}_\tau$;

8. There are four different "standard" definitions for Poisson processes, which are equivalent. We show that the definition in Example 1.1.2 and the following are equivalent:
   Definition 1.3.1. *A stochastic process* $(N(t) : t \geqslant 0)$ *is called a* Poisson process with intensity $\lambda > 0$ *if*
   (i) $N(0) = 0$*;*

(ii) *for every* $0 \leqslant t_1 \leqslant \ldots \leqslant t_n$, $n \in \mathbb{N}$ *the random variables*

$$N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$$

*are independent (*independent increments*);*

(iii) *for all* $0 \leqslant s \leqslant t$ *and* $k \in \mathbb{N}$ *we have*

$$P(N(t) - N(s) = k) = \frac{(\lambda(t-s))^k}{k!} e^{-\lambda(t-s)}.$$

*(*stationary and Poisson distributed increments*)*

The proof of the implication from Example 1.1.2 to Definition 1.3.1 can be divided into smaller steps, where I follow Section 6 in [3]. Let N be a Poisson process as constructed in Example 1.1.2 based on the random variables $X_1, X_2, \ldots$.

(a) $P(N(t) = k) = \dfrac{(\lambda t)^k}{k!} e^{-\lambda t}$ for all $t \geqslant 0$, $k \in \mathbb{N}_0$. ($\sharp$)

(b) Define for some $t > 0$ the sum $R_n^t := Y_1^t + \cdots + Y_n^t$ where

$$Y_1^t := S_{N(t)+1} - t, \quad Y_n^t := X_{N(t)+n}, \qquad n = 2, 3, \ldots.$$

Show that the stochastic process $(Q^t(s) : s \geqslant 0)$ defined by

$$Q^t(s) := \begin{cases} 0, & \text{if } s = 0, \\ \max\{k \in \{0, 1, 2, \ldots\} : R_k^t \leqslant s\}, & \text{if } s > 0, \end{cases}$$

obeys P-a.s. the equality $Q^t(s) = N(t+s) - N(t)$ for all $s \geqslant 0$. ($\natural$)

(c) Show that for all $k \in \mathbb{N}$ and $s_i \geqslant 0$, $i = 0, \ldots k$, we have

$$P\big(Y_1^t > s_1, \ldots, Y_k^t > s_k | N(t)\big) = P(X_1 > s_1) \cdots P(X_k > s_k). \tag{1.3.1}$$

(Hint: start with $k = 1$.) ($\sharp$)

(d) The stochastic process $(Q^t(s) : s \geqslant)$ is a Poisson process (in the sense of Example 1.1.2) and it is independent of $N(t)$ with $Q^t(s) \stackrel{\mathscr{D}}{=} N(s)$ for all $s \geqslant 0$. ($\natural$)

(e) Conclude that every stochastic process of the form constructed in Example 1.1.2.b satisfies Definition 1.3.1. ($\natural$)

Do not forget to show the converse. ($\natural$)

9. Assume that the service of busses starts at 8pm and then they arrive according to a Poisson process of rate $\lambda = 4$ per hour. John starts to wait for a bus at 8pm.
   (a) What is the expected waiting time for the next bus?
   (b) At 8:30pm John is still waiting. What is now the expected waiting time?

10. Show that if N_1 and N_2 are independent Poisson processes with intensity $\lambda_1 > 0$ and $\lambda_2 > 0$, respectively, then $N_1 + N_2$ is also a Poisson process.

2 Martingales

We state the following definition for martingales for both in discrete time ($I \subseteq \mathbb{N}_0$) or in continuous time ($I \subseteq [0, \infty)$).

2.1. Definition

Definition 2.1.1. *Let* $\{\mathscr{F}_t\}_{t\in I}$ *be a filtration. An adapted stochastic process* $(M(t) : t \in I)$ *is called*

(a) a martingale with respect to $\{\mathscr{F}_t\}_{t\in I}$ *if*

  (i) $E[|M(t)|] < \infty$ *for all* $t \in I$*;*

  (ii) $E[M(t)|\mathscr{F}_s] = M(s)$ P*-a.s. for all* $0 \leqslant s \leqslant t$ *and* $s, t \in I$.

(b) a submartingale *if*

  (i) $E[|M(t)|] < \infty$ *for all* $t \in I$*;*

  (ii) $E[M(t)|\mathscr{F}_s] \geqslant M(s)$ P*-a.s. for all* $0 \leqslant s \leqslant t$ *and* $s, t \in I$.

(c) a supermartingale *if*

  (i) $E[|M(t)|] < \infty$ *for all* $t \in I$*;*

  (ii) $E[M(t)|\mathscr{F}_s] \leqslant M(s)$ P*-a.s. for all* $0 \leqslant s \leqslant t$ *and* $s, t \in I$.

In Definition 2.1.1, the condition in part (i) is the technical condition such that the conditional expectation considered in part (ii) is defined. A martingale $(M(t) : t \in I)$ can be considered as a stochastic process which describes a “fair” game in the following sense: the best approximation of the future value $M(t)$ given all information available today at time s, equals the value $M(s)$ observed today. In other words, the martingale M has no systematic up- or downwards movements.

Example 2.1.2.

(a) Let $X_1, X_2, \dots$ be independent, integrable random variables with $E[X_k] = 0$ for all $k \in \mathbb{N}$. Then

$$S_n := X_1 + \cdots + X_n$$

defines a martingale $(S_n : n \in \mathbb{N})$ in discrete time with respect to the filtration $\mathscr{F}_n^S := \sigma(S_1, \dots, S_n)$ for $n \in \mathbb{N}$.

(b) Let $(N(t) : t \geqslant 0)$ be a Poisson process with intensity λ. Then $(N(t) - \lambda t) : t \geqslant 0)$ is a martingale with respect to the filtration $\mathscr{F}_t^N := \sigma(N(s) : s \leqslant t)$, see Exercise 2.5.1.

(c) Let X be a random variable with $E[|X|] < \infty$ and let $\{\mathscr{F}_t\}_{t \in I}$ be a filtration. Then $Y(t) := E[X|\mathscr{F}_t]$ defines a martingale $(Y(t) : t \in [0, T])$ with respect to $\{\mathscr{F}_t\}_{t \in I}$, see Exercise 2.5.2.

Financial Mathematics 2. *Efficient market hypotheses* requires that asset prices in financial markets reflect all relevant information about an asset, which is called *informationally efficient*. There are different forms: the *weak* efficient market hypotheses postulates that asset prices cannot be predicted from historical information about prices. In particular, this means that one cannot beat the market in the long run by using strategies based on historical share prices. The *strong* efficient market hypothesis assumes that the share prices reflect all information, public and private, and one cannot beat the market in the long run even if one includes all private and public information.

If $(S(t) : t \geqslant 0)$ denotes the share prices then the weak efficient market hypothesis means that $(\exp(-\mu t)S(t) : t \geqslant 0)$ is a martingale with respect to the generated filtration $\{\mathscr{F}_t^S\}_{t \geqslant 0}$. Here, μ denotes the expected growth rate of the share price. The strong efficient market hypothesis requires that $(\exp(-\mu t)S(t) : t \geqslant 0)$ is a martingale with respect to the much larger filtration $\{\mathscr{G}_t\}_{t \geqslant 0}$, where $\mathscr{G}_t$ contains all private and public information available at time t.

2.2. Equalities and Inequalities

Martingales satisfy some important relations and probably the most important one is the so-called Doob's maximal inequality which we will introduce below. However, some other important relations follow easily directly from the martingale property.

Example 2.2.1. A martingale $(M(t) : t \in I)$ with respect to $\{\mathscr{F}_t\}_{t \in I}$ and with $E[|M(t)|^2] < \infty$ for all $t \in I$ satisfies:

(a) $E[(M(t) - M(s))^2 | \mathscr{F}_s] = E[M^2(t)|\mathscr{F}_s] - M^2(s)$ for all $t \geqslant s$ and $s, t \in I$.

(b) $E[(M(t) - M(s))^2] = E[M^2(t) - M^2(s)]$ for all $s, t \in I$.

(c) $E[M(s)(M(t) - M(s))] = 0$ for all $0 \leqslant s \leqslant t$ and $s, t \in I$ (orthogonality of increments).

In the following result we consider martingales in continuous times. An analogue result for martingales in discrete time is available but there the assumption of continuous martingale does not make much sense.

Theorem 2.2.2. *(Doob's maximal inequality)*
Let $(M(t) : t \in [0,T])$ be a continuous martingale (or a non-negative submartingale). Then we have

(a) *for $p \geqslant 1$ and $\lambda > 0$ that*

$$\lambda^p P\left(\sup_{t\in[0,T]} M(t) \geqslant \lambda\right) \leqslant E[|M(T)|^p].$$

(b) *for $p > 1$ that*

$$E\left[\sup_{t\in[0,T]} |M(t)|^p\right] \leqslant \left(\frac{p}{p-1}\right)^p E\left[|M(T)|^p\right].$$

Proof. The inequalities follow from the analogue result for martingales in discrete time, see [18, Th. II.1.7]. □

2.3. Optional Stopping Theorem

A stopping time τ defines the *σ-algebra prior to the stopping time* by

$$\mathscr{F}_\tau := \{A \in \mathscr{A} : A \cap \{\tau \leqslant t\} \in \mathscr{F}_t \quad \text{for all } t \geqslant 0\},$$

see Exercise 1.3.7. One can think of $\mathscr{F}_\tau$ as the information which is described by the random time τ. A stopping time is called *bounded* if there exists a constant $c > 0$ such that $\tau(\omega) \leqslant c$ for all $\omega \in \Omega$.

Theorem 2.3.1. *(Optional Stopping Theorem)*
Let $(M(t) : t \geqslant 0)$ be a continuous stochastic process with $E[|M(t)|] < \infty$ for all $t \geqslant 0$ and adapted to a filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$. Then the following are equivalent:

(a) *M is a martingale w.r.t. $\{\mathscr{F}_t\}_{t\geqslant 0}$;*

(b) *$(M(\tau \wedge t) : t \geqslant 0)$ is a martingale w.r.t. $\{\mathscr{F}_t\}_{t\geqslant 0}$ for all stopping times τ;*

(c) *$E[M(\tau)] = E[M(0)]$ for all bounded stopping times τ;*

(d) *$E[M(\tau)|\mathscr{F}_\sigma] = M(\sigma)$ for all bounded stopping times σ and τ with $\sigma \leqslant \tau$.*

Often, only the implication $(a) \Rightarrow (d)$ is referred to as the Option Sampling Theorem. In fact, we collect in Theorem 2.3.1 several results.

Note, that the assumption of a bounded stopping time is essential.

Example 2.3.2. Let W be a Brownian motion (see Section 3) and define the stopping time

$$\tau := \inf\{t \geqslant 0 : W(t) = 1\}.$$

Here Proposition 1.2.7 guarantees that τ is a stopping time. Since W is a continuous martingale according to Theorem 3.2.3 , the optional sampling Theorem 2.3.1 ((a) $\Rightarrow$ (b)) implies $(W(t \wedge \tau) : t \in [0,T])$ is a martingale and thus, we have

$$E[W(t \wedge \tau)] = E[W(0)] \qquad \text{for all } t \geqslant 0.$$

However, we cannot conclude $E[W(\tau)] = E[W(0)]$, since the implication (a) $\Rightarrow$ (c) in Theorem 2.3.1 requires that the stopping time is bounded, which is not true although we have $P(\tau < \infty) = 1$, see Proposition 3.3.1.

In fact, the equation $E[W(\tau)] = E[W(0)]$ would imply $E[W(\tau)] = 0$ since $W(0) = 0$ which contradicts the fact that $W(\tau) = 1$.

2.4. Local Martingales

In later sections we will consider stochastic processes which satisfy the martingale property only *locally*, that is only the stopped process is a martingale. This generalisation of martingales will enrich our theory significantly and is fundamental for some applications in financial mathematics.

Definition 2.4.1. *A stochastic process $(X(t) : t \geqslant 0)$ adapted to a filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ is called a* local martingale w.r.t. $\{\mathscr{F}_t\}_{t\geqslant 0}$ *if there exists a non-decreasing sequence $\{\tau_n\}_{n\in\mathbb{N}}$ of stopping times such that*

(a) $P\left(\lim_{n\to\infty} \tau_n = \infty\right) = 1$;

(b) $\left(X(t \wedge \tau_n) : t \geqslant 0\right)$ *is a martingale w.r.t.* $\{\mathscr{F}_t\}_{t\geqslant 0}$ *for each* $n \in \mathbb{N}$.

The sequence $\{\tau_n\}_{n\in\mathbb{N}}$ is called localising for X.

We also call a stochastic process $(M(t) : t \in [0,T])$ defined on a bounded interval a *local martingale* if it satisfies the same definition. In this case it seems to be pointless to require part (a) of the definition. In fact, in this case one can replace Condition (a) by

(a′) $P\left(\lim_{n\to\infty} \tau_n \geqslant T\right) = 1.$

Example 2.4.2. A martingale $(M(t) : t \geqslant 0)$ is a local martingale. This can be seen by taking the constant mappings $\tau_n : \Omega \to [0,\infty]$, $\tau_n(\omega) := n$ for $n \in \mathbb{N}$, which are stopping times by Example 1.2.2.

The localising sequence for a local martingale is not unique. However, for a large class of local martingales there exists a canonical choice for the localising sequence.

Lemma 2.4.3. *Let* $(X(t) : t \in [0,T])$ *be a continuous local martingale. Then*

$$\tau_n := \inf\{t \geqslant 0 : |X(t)| \geqslant n\}$$

defines a localising sequence $\{\tau_n\}_{n\in\mathbb{N}}$ *for* X*.*

In Financial Mathematics the problem occurs that one can construct the hedging strategy as a local martingale but one does not know if it is a martingale.

Lemma 2.4.4. *Every local martingale* $(X(t) : t \geqslant 0)$ *with* $\sup_{t,\omega} |X(t)(\omega)| < \infty$ *is a martingale.*

There is a necessary and sufficient condition known which guarantees that a continuous local martingale is a martingale; see uniform integrability. Later we can easily construct specific examples of local martingales which are not martingales. Here we give an explicit example originating from [14] and [24].

Example 2.4.5. A local martingale which is not a martingale
We define a stochastic process $(X(t) : t \geqslant 0)$ which attains only values in the integers and it spends an exponentially distributed random time in each value.

Let $\{p_k\}_{k\in\mathbb{Z}}$ be a probability distribution on the integers $\mathbb{Z}$ which satisfies additionally

$$\sum_{k=-\infty}^{\infty} k^2 p_k < \infty.$$

This assumption means that the probabilities p_k of large numbers k or $-k$ decrease rapidly as $k \to \pm\infty$.

Let the process X starts from 0 where it stays for an exponentially distributed time T_0 which has expectation p_0. At time T_0, with equal probability it jumps up or down by the value of a random variable Y_1 which attains the values ± 1 with equal probability. At the level Y_1 the stochastic process X stays for the exponentially distributed time T_1 which is independent of T_0 and has expectation p_{Y_1}. At time $T_0 + T_1$ the stochastic process jumps up or down which is determined by an independent random variable Y_2 which attains the values ± 1 with equal probability. At the level $Y_1 + Y_2$ the stochastic process X stays for an exponentially distributed time T_3 with expectation $p_{Y_1+Y_2}$. If the stochastic process X has jumped to a value y in the k-th jump then it stays there for the exponentially distributed time T_k which has expectation p_y and which is independent of all prior waiting times T_j for $j = 0, \dots, k-1$. Since the moments of the waiting times depend on the values of the stochastic process X it is constructed in such a way that it spends a long time at the same value if this value is near 0 and the time is short at levels away from 0. With some basic arguments one can show that in this way the stochastic process X is defined on $[0,\infty)$ (and that $T_1 + T_2 + \dots$ does not accumulate at a finite time).

The stochastic process X is a local martingale with the localising stopping times $\tau_k := T_0 + \dots + T_{k-1}$ which is the time of the k-th jump. Since

$$X(t \wedge \tau_k) = \sum_{i=1}^{k} Y_i \mathbb{1}_{\{\tau_i \leqslant t\}}$$

and Y_i is independent of $B \cap \{\tau_i \in (s,t]\}$ for every $B \in \mathscr{F}_s^X$ we obtain for $0 \leqslant s \leqslant t$ that

$$E[(X(t\wedge\tau_k) - X(s\wedge\tau_k))\mathbb{1}_B] = \sum_{i=1}^k E\left[Y_i \mathbb{1}_{B\cap\tau_i\in(s,t]}\right] = \sum_{i=1}^k E[Y_i]E\left[\mathbb{1}_{B\cap\tau_i\in(s,t]}\right] = 0,$$

which shows that X is a local martingale.

You can convince yourself that X is not a martingale by the following heuristic argument: if for $s < t$ the value $X(s)$ is given then the expected value of $X(t)$ is $X(s)$. However, if the latter is very large then this contradicts our intuition that the expected value of $X(t)$ is rather small since X spends most of the time near 0. A formal argument can be found in [24].

2.5. Exercises

In this section W denotes a Brownian motion as introduced in Section 3.

1. Let $(N(t) : t \geqslant 0)$ denote the Poisson process, constructed in Example 1.1.2. Show that the following stochastic processes are martingales w.t.r $\{\mathscr{F}_t^N\}_{t\geqslant 0}$:
   (a) $(N(t) - \lambda t : t \geqslant 0)$;
   (b) $\big((N(t) - \lambda t)^2 - \lambda t : t \geqslant 0\big)$.
   (c) $\big((N(t) - \lambda t)^2 - N(t) : t \geqslant 0\big)$.
   (Hint: use Exercise 1.3.8).

2. Let X be a random variable with $E[|X|] < \infty$ and let $\{\mathscr{F}_t\}_{t\in I}$ be a filtration. Then $Y(t) := E[X|\mathscr{F}_t]$ defines a martingale $(Y(t) : t \in [0,T])$ with respect to $\{\mathscr{F}_t\}_{t\in I}$,

3. Check whether the following stochastic processes $(X(t) : t \geqslant 0)$ are martingales with respect to the generated filtration $\{\mathscr{F}_t^W\}_{t\geqslant 0}$ where
   (a) $X(t) = W(t)$;
   (b) $X(t) = W^2(t)$;
   (c) $X(t) = \exp(cW(t) - \frac{c^2}{2}t)$ for every constant $c \in \mathbb{R}$;
   (d) $X(t) = W^3(t) - 3tW(t)$;
   (e) $X(t) = t^2W(t) - 2\int_0^t sW(s)\,ds$;
   (f) $X(t) = W^4(t) - 4t^2W(t)$.

4. For a constant $a > 0$ define $\tau := \inf\{t \geqslant 0 : W(t) \notin (-a,a)\}$.
   (a) Give a reason that τ is a stopping time with respect to $\{\mathscr{F}_t^W\}_{t\geqslant 0}$.
   (b) Show that
   $$M(t) := \exp\left(-\tfrac{c^2}{2}t\right)\cosh(cW(t))$$
   defines a martingale $(M(t) : t \geqslant 0)$ with respect to $\{\mathscr{F}_t^W\}_{t\geqslant 0}$.

(c) Show that $E[\exp(-\lambda\tau)] = (\cosh(a\sqrt{2\lambda}\,))^{-1}$ for every $\lambda > 0$. (∗)

5. For constants a, $b > 0$ define $\tau := \inf\{t \geqslant 0 : W(t) = a + bt\}$. Show that for each $\lambda > 0$ we have

$$E\left[e^{-\lambda\tau}\right] = \exp^{-a(b+\sqrt{b^2+2\lambda}\,)}.$$

(Hint: use part (c) in question 4 with $c = b + \sqrt{b^2 + 2\lambda}\,$). (∗)

6. Let $(X_k : k \in \mathbb{N}_0)$ be a submartingale with respect to a filtration $\{\mathscr{F}_k\}_{k\in\mathbb{N}_0}$. Then there exists a unique martingale $(M_k : k \in \mathbb{N}_0)$ with respect to $\{\mathscr{F}_k\}_{k\in\mathbb{N}_0}$ and an increasing sequence $(A_k)_{k\in\mathbb{N}_0}$ of random variables A_k, where A_k is $\mathscr{F}_{k-1}$-measurable and $A_0 = 0$, such that

$$X_k = M_k + A_k \qquad \text{for all } k \in \mathbb{N}_0\,.$$

This is the so-called *Doob-Meyer decomposition*, which is also true for submartingales in continuous times but there much harder to prove.

7. Let the share price be modeled by $(S(t) : t \geqslant 0)$ where

$$S(t) = s_0 \exp\left(\sigma W(t) + \left(r - \tfrac{1}{2}\sigma^2\right) t\right) \qquad \text{for all } t \geqslant 0,$$

where $s_0 \in \mathbb{R}_+$ and $r, \sigma > 0$. Calculate the value process $(V(t) : t \geqslant 0)$ of an European call option which is given by

$$V(t) = \frac{1}{e^{r(T-t)}} E[(S(T) - K)^+ | \mathscr{F}_t^W] \qquad \text{for all } t \geqslant 0,$$

and where $K > 0$ denotes the strike price $K > 0$ and $T > 0$ the maturity. (♮)

8. Let $(X_k : k \in \mathbb{N})$ be a local martingale with $E[|X_k|] < \infty$ for all $k \in \mathbb{N}$. Then it follows that X is a martingale.
Note, this is only true because it is a local martingale in discrete time.

3 Brownian Motion

The most important example of a stochastic process is the Brownian motion which is also called a Wiener process. The botanist Robert Brown (1773 -1858) observed an example of a two-dimensional Brownian motion as the diffusion of pollen of different plants in water in 1827. However, Brown was not able to explain his observations. Later, the one-dimensional Brownian motion was used by Louis Bachelier in his PhD thesis (Théorie de la spéculation, Ann. Sci. École Norm. Sup. 17, 1900) to model a financial market. In 1905, Albert Einstein published a theory to explain the motion of pollen observed by Brown. He observed that the kinetic energy of fluids makes the molecules of water to move randomly. Thus, a pollen grain is suspended to a random number of impacts of random strength and from random directions. This random bombardment by the molecules of the fluid causes a small particle to move as it was described by Brown. However, Einstein did not provide a mathematical proof of the existence of a Browian motion. This was done in 1923 by the American mathematician Norbert Wiener who used newly developed methods from measure theory. Finally, the work by the Japanese Mathematician Kiyoshi Itô in the '40s plays a fundamental role in the application of Brownian motion in a wide spectrum of sciences such as biology, economics, finance and physics.

Definition 3.0.1. *A stochastic process* $(W(t) : t \geqslant 0)$ *with values in* $\mathbb{R}^d$ *is called a* d-dimensional Brownian motion *if*

(a) $W(0) = 0$ *P-a.s.*

(b) W *has independent increments, i.e.*

$$W(t_2) - W(t_1), \dots, W(t_n) - W(t_{n-1})$$

are independent for all $0 \leqslant t_1 < t_2 < \dots < t_n$ *and all* $n \in \mathbb{N}$*;*

(c) *the increments are normally distributed, i.e.*

$$W(t) - W(s) \stackrel{\mathscr{D}}{=} N(0, (t-s)\,\mathrm{Id}_d)$$

for all $0 \leqslant s \leqslant t$.

(d) W *has continuous trajectories.*

Condition (c) implies for every $h \geqslant 0$

$$W(t+h) - W(s+h) \stackrel{\mathscr{D}}{=} W(t) - W(s) \qquad \text{for all } 0 \leqslant s < t.$$

Together with Condition (b) we can conclude for all $0 \leqslant t_1 < t_2 < \cdots < t_n$ and all $n \in \mathbb{N}$ that the random vector

$$\Big(W(t_2) - W(t_1), \dots, W(t_n) - W(t_{n-1})\Big)$$

has the same joint distribution as the random vector

$$\Big(W(t_2+h) - W(t_1+h), \dots, W(t_n+h) - W(t_{n-1}+h)\Big).$$

This property is called *stationary increments.*

In contrast to the case of a Poisson process in Example 1.1.2 a Brownian motion is only formally defined. Thus, from a mathematical point of view, we should convince ourself now that there exists a Brownian motion in a probability space. In fact, there are three common ways to construct a Brownian motion, but in this course, we skip the mathematical proof and assume its existence.

Since the Brownian motion W maps to the d-dimensional Euclidean space $\mathbb{R}^d$ we can represent it componentwise:

$$W = ((W_1(t), \dots, W_d(t)) : t \geqslant 0).$$

Let k be in $\{1, \dots, d\}$. By property (c) it follows that for each $0 \leqslant s < t$ the random variable $W_k(t) - W_k(s)$ is a normally distributed random variable with expectation 0 and variance $t - s$. Moreover, property (b) implies that the one-dimensional stochastic process $W_k := (W_k(t) : t \geqslant 0)$ has independent increments and therefore, W_k is also a Brownian motion but with values in $\mathbb{R}$. Furthermore, for all $t \geqslant 0$ the random variables $W_1(t), \dots, W_d(t)$ are independent since $\operatorname{Cov}(W_i(t)W_j(t)) = 0$ for each $t \geqslant 0$ which together with (b) implies that the stochastic processes $W_1, \dots, W_d$ are independent. For these reasons, we will consider in the remaining part of this chapter only one-dimensional Brownian motion, i.e. $d = 1$, and the generalisation to the multi-dimensional setting is in most cases obvious.

Recall that for a given stochastic process $W = (W(t) : t \geqslant 0)$ the generated filtration is denoted by $\{\mathscr{F}_t^W\}_{t \geqslant 0}$, i.e.

$$\mathscr{F}_t^W = \sigma(W(s) : s \in [0,t]) \qquad \text{for all } t \geqslant 0.$$

The property (b) in Definition 3.0.1 is equivalent to

$$\text{(b}'\text{)}\ W(t) - W(s) \text{ is independent of } \mathscr{F}_s^W \text{ for all } 0 \leqslant s < t. \tag{3.0.1}$$

We often use this formulation instead of Condition (b).

3.1. Brownian filtration

A mathematically more precise definition of Brownian motions includes the underlying filtration into the definition, similarly as in the definition of a stopping time or of a martingale.

Definition 3.1.1. *A stochastic process* $(W(t): t \geqslant 0)$ *with values in* $\mathbb{R}^d$ *and adapted to a filtration* $\{\mathscr{F}_t\}_{t\geqslant 0}$ *is called a* d-dimensional Brownian motion with respect to a filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ *if*

(a) $W(0) = 0$ *P-a.s.*

(b) *for every* $0 \leqslant s < t$ *the increment* $W(t) - W(s)$ *is independent of* $\mathscr{F}_s$.

(c) *the increments are normally distributed, i.e.*

$$W(t) - W(s) \stackrel{\mathscr{D}}{=} N(0, (t-s)\,\mathrm{Id}_d)$$

for all $0 \leqslant s \leqslant t$.

(d) W *has continuous trajectories.*

The main difference between Definition 3.0.1 and 3.1.1 is part (b). If $\{\mathscr{F}_t^W\}_{t\geqslant 0}$ denotes the filtration generated by a Brownian motion W but not augmented, then Condition (b) in Definition 3.0.1 implies that $W(t) - W(s)$ is independent of $\mathscr{F}_s$ for every $0 \leqslant s < t$, see (3.0.1). Consequently, a Brownian motion W satisfying Definition 3.0.1 is a Brownian motion with respect to $\{\mathscr{F}_t^W\}_{t\geqslant 0}$ in the sense of Definition 3.1.1. The problem of this subsection comes now from the fact that instead of the filtration $\{F_t^W\}_{t\geqslant 0}$ we want to consider the larger augmented filtration $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$. Considering the augmented filtration has the important advantage that it satisfies the usual conditions, see Definition 1.1.9, whereas this is not true for the generated, non-augmented filtration.

Proposition 3.1.2. *The augmented filtration* $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$ *of a Brownian motion satisfies the usual conditions, i.e.*

(a) $\mathscr{N} \subseteq \mathfrak{F}_0^W$.

(b) $\mathfrak{F}_t^W = \bigcap_{s>t} \mathfrak{F}_s^W$ *for all* $t \geqslant 0$.

Theorem 3.1.3. *A Brownian motion* W *in the sense of Definition 3.0.1 is a Brownian motion with respect to the augmented filtration* $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$ *in the sense of Definition 3.1.1.*

3.2. Properties

Proposition 3.2.1. *A Brownian motion* $(W(t): t \geqslant 0)$ *satisfies*

(a) $W(t) - W(s) \stackrel{\mathscr{D}}{=} W(t-s)$ *for all* $0 \leqslant s < t$.

(b) $E[W(t) - W(s)] = 0 \quad$ *for all* $0 \leqslant s \leqslant t$,

$\mathrm{Var}[W(t) - W(s)] = t - s \quad$ *for all* $0 \leqslant s \leqslant t$.

$E[W(s)W(t)] = s \wedge t \quad$ *for all* $s, t \geqslant 0$.

(c) $E[\exp(uW(t))] = \exp\left(\frac{1}{2}u^2 t\right) \qquad$ *for all* $u \in \mathbb{R}$, $t \geqslant 0$.

Proof. Part (a) and (b) follow easily from the definition. Part (c) follows from a short calculation. □

If X and Y are two real-valued random variables the notation $X \stackrel{\mathscr{D}}{=} Y$ means that X and Y are *equal in distribution*, i.e.

$$P(X \in B) = P(Y \in B) \quad \text{for all } B \in \mathfrak{B}(\mathbb{R}).$$

In particular, it follows that the moments coincide

$$E[X^k] = E[Y^k] \quad \text{for all } k \in \mathbb{N}\,.$$

However, equality in distribution does not mean very much; for example if X is a normally distributed random variable with $E[X] = 0$, then $X \stackrel{\mathscr{D}}{=} -X$. Thus, property (a) says only that $P(W(t)-W(s) \in A) = P(W(t-s) \in A)$ for each Borel set $A \in \mathfrak{B}(\mathbb{R})$. But this implies for example, that $E[W(t) - W(s)] = E[W(t-s)]$.

Proposition 3.2.2. *Let $(W(t) : t \geqslant 0)$ be a Brownian motion and let $c > 0$. Then we have:*

(a) $X(t) := cW(t/c^2)$, $t \geqslant 0$, *defines a Brownian motion* $(X(t) : t \geqslant 0)$.

(b) $Y(t) := \begin{cases} 0, & \text{if } t = 0, \\ tW(1/t), & \text{if } t > 0, \end{cases}$ *defines a Brownian motion* $(Y(t) : t \geqslant 0)$.

Proof. [1] (b) (follows the proof of Th.1.9 in [15]): By definition of Y we have $Y(0) = 0$. For every $0 < s \leqslant t$ we obtain

$$E[Y(t)] = tE[W(1/t)] = 0, \qquad \operatorname{Cov}(Y(s), Y(t)) = stE\left[W(\tfrac{1}{s})W(\tfrac{1}{t})\right] = st\tfrac{1}{t} = s. \tag{3.2.2}$$

Fix some $0 \leqslant t_1 < \cdots < t_n$. Since W is a Gaussian process according to Exercise 3.4.3, it follows that $(Y(t) : t \geqslant 0)$ is a Gaussian process, which implies that the random vector

$$U := \big(Y(t_1) \dots, Y(t_n)\big)$$

is normally distributed. The equalities in (3.2.2) show that

$$\begin{aligned} E\big[\big(Y(t_1), \dots, Y(t_n)\big)\big] &= E\big[\big(W(t_1), \dots, W(t_n)\big)\big], \\ \Big(\operatorname{Cov}(Y(t_k), Y(t_l))\Big)_{k,l=1,\dots,n} &= \Big(\operatorname{Cov}(W(t_k), W(t_l))\Big)_{k,l=1,\dots,n}. \end{aligned}$$

Since the normal distribution in $\mathbb{R}^n$ is characterised by the expectations and covariances, the random vector U has the same distribution as $V := \big(W(t_1), \dots, W(t_n)\big)$. Since

$$\begin{pmatrix} Y(t_1) \\ Y(t_2) - Y(t_1) \\ \vdots \\ Y(t_n) - Y(t_{n-1}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix} U \stackrel{\mathscr{D}}{=} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix} V,$$

the independent and stationary increments of W imply the independent and stationary increments of Y.

[1] this proof is only included since I had it typed already.

The paths $t \mapsto Y(t)$ are continuous on $(0,\infty)$. Since the above shows that $(Y(t) : t \in \mathbb{Q} \cap [0,\infty))$ and $(W(t) : \mathbb{Q} \cap [0,\infty))$ have the same finite-dimensional distributions, it follows

$$\lim_{\substack{t \searrow 0 \\ t \in \mathbb{Q}}} Y(t) = \lim_{\substack{t \searrow 0 \\ t \in \mathbb{Q}}} W(t) = 0 \qquad P\text{-a.s.}$$

Since $\mathbb{Q} \cap (0,\infty)$ is dense in $(0,\infty)$ and Y is continuous on $(0,\infty)$ it follows

$$\lim_{t \searrow 0} Y(t) = 0,$$

which completes the prove.

For the last (hard) part, the continuity in zero, there is another argument, based on the so-called law of the iterated logarithm of the Brownian motion, see Theorem 3.3.8. □

Property (a) in Proposition 3.2.2 is called *Brownian scaling* and property (b) is called *time inversion*. Further transformations which again lead to Brownian motions can be found in Exercise 3.4.1.

Theorem 3.2.3. *A Brownian motion W is a continuous martingale with respect to the generated filtration $\{\mathscr{F}_t^W\}_{t \geqslant 0}$.*

Corollary 3.2.4. *Let $(W(t) : t \geqslant 0)$ be a Brownian motion and $c \in \mathbb{R}$ a non-zero constant. Then we have:*

(a) $\big(W^2(t) - t : t \geqslant 0\big)$ *is a continuous martingale with respect to $\{\mathscr{F}_t^W\}_{t \geqslant 0}$.*

(b) $\Big(\exp(cW(t) - \frac{c^2}{2}t) : t \geqslant 0\Big)$ *is a continuous martingale with respect to $\{\mathscr{F}_t^W\}_{t \geqslant 0}$.*

Remark 3.2.5. In fact the two previous results, Theorem 3.2.3 and Corollary 3.2.4, are even true if the generated filtration $\{\mathscr{F}_t^W\}_{t \geqslant 0}$ is replaced by the larger augmented filtration $\{\mathfrak{F}_t^W\}_{t \geqslant 0}$. This can be easily seen by taking into account the result of Theorem 3.1.3 and repeating our proofs with the augmented filtration.

Financial Mathematics 3. Recall that the share prices under the equivalent risk-neutral measure in the Black-Scholes model are given by a geometric Brownian motion $(S(t) : t \geqslant 0)$ which is of the form

$$S(t) = \exp\left(\sigma W(t) + \left(r - \tfrac{\sigma^2}{2}\right) t\right) \qquad \text{for all } t \geqslant 0,$$

where $\sigma > 0$ is the volatility and r is the interest rate of a risk-less bond in the market. It follows by Corollary 3.2.4 that the discounted share prices $(\exp(-rt)S(t) : t \geqslant 0)$ define a martingale under the equivalent risk-neutral measure.

3.3. Path Properties

In order to develop an idea of the behaviour of the trajectories of Brownian motions we begin with studying the exit time of an interval around zero.

Proposition 3.3.1. *For $a, b > 0$ define*

$$\tau := \inf\{t \geqslant 0 : W(t) = -b \text{ or } W(t) = a\}.$$

Then the stopping time τ obeys the following:

(a) $P(\tau < \infty) = 1$.

(b) $P(W(\tau) = a) = \frac{b}{a+b}$.

(c) $E[\tau] = ab$.

Definition 3.0.1 requires the paths of a Brownian motion to be continuous. In fact, the following theorem, which is called *Kolmogorov's continuity theorem*, enables us to conclude that if a stochastic process satisfies (a), (b) and (c) in Definition 3.0.1 then there exists a modification of this process with continuous paths. Thus, it is not necessary to require condition (d) in Definition 3.0.1, which we do for simplicity.

Kolmogorov's continuity theorem gives not only a sufficient condition for guaranteeing the existence of a modification with continuous paths, but even with Hölder continuous paths. **Hölder continuous:** a function $f : [a, b] \to \mathbb{R}$ is called *Hölder continuous of order $h \in \mathbb{R}_+$*, if there exists a constant $c > 0$ such that

$$|f(x) - f(y)| \leqslant c\,|x - y|^h \qquad \text{for all } a \leqslant x \leqslant y \leqslant b.$$

A function which is Hölder continuous of order 1 is also called Lipschitz continuous.

Example 3.3.2. A function $f : [a, b] \to \mathbb{R}$ which is continuously differentiable is Hölder continuous of order 1. This follows immediately from the mean value theorem of calculus.

Theorem 3.3.3. *(Kolmogorov's continuity theorem)*
Let $(X(t) : t \in [0, T])$ be a stochastic process on a probability space $(\Omega, \mathscr{A}, P)$ which satisfies that for each $T > 0$ there exist constants $\alpha, \beta, \gamma > 0$ such that

$$E\left[|X(t) - X(s)|^\alpha\right] \leqslant \gamma\,|t - s|^{1+\beta} \qquad \text{for all } 0 \leqslant s \leqslant t \leqslant T.$$

Then there exists a modification of X with Hölder continuous paths on $[0, T]$ of any order $h \in [0, \beta/\alpha]$.

Proof. See [9, Thm. II.2.8]. □

In order to apply Theorem 3.3.3 to a Brownian motion we have to calculate its moments.

Lemma 3.3.4. *Let X be a normally distributed random variable, i.e. $X \stackrel{\mathscr{D}}{=} N(0, \sigma^2)$ for some $\sigma^2 > 0$. Then*

$$E\left[X^k\right] = \begin{cases} 0, & \text{if } k \text{ is odd}, \\ \frac{k!}{2^{k/2}(k/2)!}\sigma^k, & \text{if } k \text{ is even}. \end{cases}$$

Lemma 3.3.4 guarantees that a Brownian motion satisfies for each $\alpha \in \mathbb{N}$

$$E[|W(t) - W(s)|^{2\alpha}] \leqslant c\,|t-s|^{\alpha} \qquad \text{for all } 0 \leqslant s \leqslant t, \tag{3.3.3}$$

for a constant $c > 0$, depending on α. Theorem 3.3.3 implies that there exists a modification of the Brownian motion with Hölder continuous paths of every order $\beta < \frac{1}{2}$:

Corollary 3.3.5. *For any $h < \frac{1}{2}$, there exists a modification of a Brownion motion with Hölder continuous paths of order h on $[0,T]$ for every $T > 0$.*

After Definition 1.1.1 it is mentioned that there are at least two ways to consider a stochastic process. There is a third one (closely related to the view of trajectories) if one knows something about the trajectories, as for example for a Brownian motion W. The continuity of the paths imply that we can define

$$Z : \Omega \to C([0,T]), \qquad Z(\omega) := (W(t)(\omega) : t \in [0,T]),$$

where $C([0,T])$ denotes the space of (deterministic) continuous functions on the interval $[0,T]$. If we assume that we can equip $C([0,T])$ with a σ-algebra then we can ask if Z is a random variable. (Yes, it is!).

Next we show the opposite result that a Brownian motion does not have trajectories which are Hölder continuous of any order larger than $\frac{1}{2}$, where I follow the notes [21] by T. Seppäläinen.

Theorem 3.3.6. *Let $(W(t) : t \geqslant 0)$ be a Brownian motion and define for every ε, β, $c > 0$ the set*

$$R_\varepsilon(\beta, c) := \Big\{\omega \in \Omega : \exists s > 0 : |W(t)(\omega) - W(s)(\omega)| \leqslant c\,|t-s|^{\beta} \text{ for all } t \in [s-\varepsilon, s+\varepsilon]\Big\}.$$

Then we obtain for each $\beta > \frac{1}{2}$

$$P\big(R_\varepsilon(\beta, c)\big) = 0 \qquad \text{for all } \varepsilon, c > 0.$$

By the mean-value theorem, any continuous, differentiable function $f : [0,T] \to \mathbb{R}$ with a bounded derivative is Hölder continuous of order 1. As a consequence, we obtain from Theorem 3.3.6 that the paths of a Brownian motion cannot be differentiable.

Corollary 3.3.7. *With probability one, the trajectories of a Brownian motion $(W(t) : t \geqslant 0)$ are not differentiable at any time $t \geqslant 0$.*

The trajectories of a Brownian motion are Hölder continuous of any order less than 1/2 but not of any order larger than 1/2. This can be made precisely:

Theorem 3.3.8. *Every Brownian motion $(W(t) : t \geqslant 0)$ satisfies*

$$\limsup_{\delta \searrow 0} \sup_{\substack{0 \leqslant s < t \\ t-s \leqslant \delta}} \frac{W(t) - W(s)}{\sqrt{2\delta \log \delta^{-1}}} = 1 \quad P\text{-a.s.}$$

Another characterisation of the irregularity of a function is its total variation which we introduce in the following short excursion to calculus of deterministic functions.

Finite Variation: for an interval $[a,b]$ a partition is any sequence $\{t_k\}_{k=0,\dots,n}$ for $n \in \mathbb{N}$ with

$$a = t_0 < t_1 < \cdots < t_n = b.$$

The *mesh of a partition* $\pi = \{t_k\}_{k=0,\dots,n}$ is defined as

$$|\pi| := \max_{k=1,\dots,n}(t_k - t_{k-1}).$$

The set of all partitions of the interval $[a,b]$ is denoted by $P[a,b]$. For a function $f : [a,b] \to \mathbb{R}$ the *total variation* $TV_f([a,b])$ is defined by

$$TV_f([a,b]) := \sup\left\{\sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| : \{t_k\}_{k=0,\dots n} \in P[a,b],\, n \in \mathbb{N}\right\}.$$

If $\mathrm{TV}_f([a,b])$ is finite the function f is *of finite variation on* $[a,b]$. The total variation of a function f does not give the length of the function, consider for example $f(t) := t$ for $t \in [0,1]$. The total variation might be described as the trace of the function f projected to the y-axis.

Example 3.3.9. If $f : [a,b] \to \mathbb{R}$ is an increasing function, then the sum in the definition of the total variation is just a telescope sum and we obtain

$$TV_f([a,b]) = f(b) - f(a).$$

A similar result holds if f is decreasing.

Example 3.3.10. If $f : [a,b] \to \mathbb{R}$ is differentiable and the derivative $f' : [a,b] \to \mathbb{R}$ is continuous then f has finite variation and

$$TV_f([a,b]) = \int_a^b |f'(s)|\, ds.$$

This follows from the mean value theorem.

Example 3.3.11. Define a function

$$f : [0,1] \to \mathbb{R}, \qquad f(x) := \begin{cases} x \sin\left(\frac{1}{x}\right), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Show, that f is continuous but not of finite variation; see Exercise 3.4.8.

In order to be consistent with other notations in particular with books on stochastic analysis we mention the following: for a function $f : [a,b] \to \mathbb{R}$ define for every partition $\pi = \{t_k\}_{k=0,\dots,n}$ of $[a,b]$

$$s_\pi := \sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)|\,.$$

If f is continuous then

$$TV_f([a,b]) = \lim_{|\pi|\to 0} s_\pi, \tag{3.3.4}$$

which is shorthand that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $\pi \in P(a,b)$:

$$|\pi| \leqslant \delta \;\Rightarrow\; TV_f([a,b]) - \varepsilon \leqslant s_\pi \leqslant TV_f([a,b]).$$

The equation (3.3.4) is noteworthy, since on the left hand side the supremum is taken over all partitions whereas on the right hand the limit is taken only over partitions with vanishing mesh.

In the sequel we show that the paths of a Brownian motion are not of finite variation, which will imply later that we cannot apply standard integration theory to define an integral with respect to Brownian motion. Furthermore, the result also shows that the paths of a Brownian motion are quite rough and oscillate rapidly. We begin with a kind of a positive result which we will need later:

Proposition 3.3.12. *Every Brownian motion $(W(t) : t \geqslant 0)$ satisfies for each $T > 0$:*

$$\lim_{\substack{|\pi|\to 0\\ P[0,T]\ni\pi=\{t_k\}_{k=1,\dots,n}}} E\left[\left|\sum_{k=1}^{n-1}\Big(W(t_{k+1}) - W(t_k)\Big)^2 - T\right|^2\right] = 0.$$

We express the property that the trajectories of a Brownian motion W are not of finite variation by replacing the deterministic function f in the definition of $TV_f([a,b])$ by the Brownian motion W. Obviously, $TV_W([a,b])$ depends then on $\omega \in \Omega$ and we do assume that it is a random variable.

Theorem 3.3.13. *With probability one, a Brownian motion is of infinite variation on every interval $[a,b]$, that is*

$$P\Big(TV_W([0,T]) = \infty \quad \textit{for every } T > 0\Big) = 1.$$

There are much more known about the trajectories of a Brownian motion, but there are even some questions open. For us the most important result is that the trajectories are not of bounded variation.

3.4. Exercises

In this section $(W(t) : t \geqslant 0)$ denotes a one-dimensional Brownian motion.

1. Show that the following stochastic processes are Brownian motions:
   (a) $(-W(t) : t \geqslant 0)$.
   (b) $(W(t+t_0) - W(t_0) : t \geqslant 0)$ for any fixed constant $t_0 \geqslant 0$.
   (c) $(cW(t/c^2) : t \geqslant 0)$ for any fixed constant $c > 0$.

2. Show that a normally distributed random variable X with expectation 0 and variance $\sigma^2 > 0$ satisfies

$$E\left[X^k\right] = \begin{cases} 0, & \text{if } k \text{ is odd,} \\ \frac{k!}{2^{k/2}(k/2)!}\sigma^k, & \text{if } k \text{ is even,} \end{cases}$$

for $k \in \mathbb{N}$ (Proof of Lemma 3.3.4).

(Hint: apply partial integration and then induction.) $(*)$

3. Show that a Brownian motion is a Gaussian process by using the definition given in Example 1.1.2.

(Hint: use the result that the image of a normal random vector under an affine function is again normally distributed.)

4. Define a stochastic process $(B(t) : t \in [0,1])$ by $B(t) = W(t) - tW(1)$ for all $t \in [0,1]$. Show that the covariance is given by

$$\operatorname{Cov}(B(t_i)B(t_j)) = \min\{t_i, t_j\} - t_i t_j$$

for all $t_i, t_j \in [0,1]$. Conclude from this that $(B(t) : t \in [0,1])$ is a Brownian bridge: a stochastic process $(B(t) : t \geqslant 0)$ is called a *Brownian bridge* if B is a continuous, Gaussian stochastic process which starts at 0 P-a.s. and satisfies

$$E\left[B(t)\right] = 0 \qquad \text{for all } t \geqslant 0,$$
$$E\left[B(t)B(s)\right] = \min\{s,t\} - st \qquad \text{for all } s,t \in [0,1].$$

5. Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a continuous function. Define for each $t \geqslant 0$ the random variable

$$Y(t) := \int_0^t W(s) f(s)\, ds.$$

   (a) Calculate expectation and variance of $Y(t)$ and $\operatorname{Cov}(Y(s), Y(t))$ for $s,t \geqslant 0$. Calculate the explicit values for $f(t) = 1$ and $f(t) = t^2$, $t \geqslant 0$. $(*)$

   (b) Show that $Y(t)$ is a Gaussian random variable (you need the definition of Riemann integrals) $(*)$

6. (a) If X is a non-negative random variable show that

$$E[X] = \int_0^\infty P(X \geqslant y)\, dy.$$

   (b) Show that if a random variable X satisfies $E[|X|] < \infty$ then $P(X < \infty) = 1$.

If you don't know measure theory you can assume in both questions that the distribution of X has a density f.

These arguments are used in the proof of Proposition 3.3.1.

7. For some $H \in (0,1)$, let $(W^H(t) : t \geqslant 0)$ be a Gaussian stochastic process with $E[W^H(t)] = 0$ for all $t \geqslant 0$ and with covariance

$$r(s,t) := E[W^H(s)W^H(t)] = \frac{1}{2}\Big(t^{2H} + s^{2H} - |t-s|^{2H}\Big) \qquad \text{for all } s,t \geqslant 0.$$

   (a) Show that W^H has stationary increments.

   (b) Show that W^H has a modification with Hölder continuous paths of every order $\alpha < H$.

   This exercise introduces fractional Brownian motions. A stochastic process $(W^H(t) : t \geqslant 0)$ is called a *fractional Brownian motion with Hurst index* $H \in (0,1)$ if W^H is a continuous Gaussian stochastic process which starts at 0 P-a.s. and satisfies

$$E\left[W^H(t)\right] = 0 \qquad \text{for all } t \geqslant 0,$$
$$E\left[W^H(t)W^H(s)\right] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right) \qquad \text{for all } s,t \geqslant 0.$$

8. (a) Show that a Hölder continuous function $f : [a,b] \to \mathbb{R}$ of order $\alpha = 1$ is of finite variation.

   (b) (Example 3.3.11) Define a function

$$f : [0,1] \to \mathbb{R}, \qquad f(x) := \begin{cases} x \sin\left(\frac{1}{x}\right), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

   Show that f is continuous but not of bounded variation. (♯)

4 Stochastic Integration

In this chapter we define a stochastic integral for a large class of integrands with respect to Brownian motion. The main challenge is to overcome the fact that the Brownian motion is not of bounded variation, see Theorem 3.3.13. This prevents us to apply the integration theory from calculus in a pathwise sense (see the first section). A way to circumvent this difficulty was introduced by K. Itô in the 1940s by defining the stochastic integral as a limit in $L^2_P(\Omega)$. We finish the chapter by deriving an analogue of the fundamental theorem of calculus for this new kind of integration.

In this chapter $(\Omega, \mathscr{A}, P)$ is a probability space with a filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ which satisfies the usual conditions. Furthermore, $(W(t) : t \geqslant 0)$ denotes a Brownian motion and we assume that

(i) $W(t)$ is $\mathscr{F}_t$-adapted for all $t \geqslant 0$;

(ii) $W(t) - W(s)$ is independent of $\mathscr{F}_s$ for all $0 \leqslant s \leqslant t$.

These assumptions are nothing else than to require that W is a Brownian motion with respect to the filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ in the sense of Definition 3.1.1. The augmented filtration $\{\mathfrak{F}^W_t\}_{t\geqslant 0}$ always satisfies these assumptions however it might be beneficiary to include larger filtrations.

4.1. Why do we need a new kind of integration?

There are different ways and generalities to define an integral for a deterministic function $f\colon [a,b] \to \mathbb{R}$. The standard approach is the *Riemann integral* which is defined by

$$\int_a^b f(s)\,ds := \lim_{|\pi_n|\to 0} \sum_{i=0}^{m_n-1} f\left(\zeta_i^{(n)}\right)\left(t_{i+1}^{(n)} - t_i^{(n)}\right), \tag{4.1.1}$$

where $\pi_n = \{t_i^{(n)}\}_{i=0,\dots,m_n}$ is a partition of $[a,b]$ for each $n \in \mathbb{N}$ and $\zeta_i^{(n)} \in [t_i^{(n)}, t_{i+1}^{(n)}]$. Of course, this definition makes only sense if the right hand side converges for every sequence $\{\pi_n\}_{n\in\mathbb{N}}$ of partitions of $[a,b]$ with $|\pi_n| \to 0$ and for every choice of $\zeta_i^{(n)} \in [t_i^{(n)}, t_{i+1}^{(n)}]$. In this case, the function f is called *Riemann integrable* and the left hand side is the *Riemann integral*. One can show that at least every continuous function $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable.

The fundamental theorem of calculus states that if $f\colon [a,b] \to \mathbb{R}$ is differentiable and the derivative f' is Riemann integrable then

$$f(b) - f(a) = \int_a^b f'(s)\,ds. \tag{4.1.2}$$

The Riemann integral can be generalised to the *Riemann-Stieltjes integral* which assigns various weights to the integrand f. The weights are described by a function $g : [a,b] \to \mathbb{R}$. Then the Riemann-Stieltjes integral is defined by

$$\int_a^b f(s)\,g(ds) := \lim_{|\pi_n|\to 0} \sum_{i=0}^{m_n-1} f\left(\zeta_i^{(n)}\right)\left(g\left(t_{i+1}^{(n)}\right) - g\left(t_i^{(n)}\right)\right), \tag{4.1.3}$$

where $\pi_n = \{t_i^{(n)}\}_{i=0,\dots,m_n}$ is a partition of $[a,b]$ for each $n \in \mathbb{N}$ and $\zeta_i^{(n)} \in [t_i^{(n)}, t_{i+1}^{(n)}]$. As before, the definition makes only sense if the right hand side converges for every sequence $\{\pi_n\}_{n\in\mathbb{N}}$ of partitions and every intermediate argument $\zeta_i^{(n)} \in [t_i^{(n)}, t_{i+1}^{(n)}]$. In this situation, the function f is called *Riemann-Stieltjes integrable with respect to* g and the unique limit is called the *Riemann-Stieltjes integral of* f *with respect to* g. One can show that at least every continuous function $f : [a,b] \to \mathbb{R}$ is Riemann-Stieltjes integrable with respect to every function g of finite variation. In this situation it follows easily that

$$\left|\int_a^b f(s)\,g(ds)\right| \leqslant \sup_{s\in[a,b]} |f(s)|\, TV_g([a,b]).$$

If a function $f\colon [a,b] \to \mathbb{R}$ is Riemann-Stieltjes integrable with respect to some $g : [a,b] \to \mathbb{R}$, then g is Riemann-Stieltjes integrable with respect to f and they satisfy

$$f(b)g(b) - f(a)g(a) = \int_a^b f(s)\,g(ds) + \int_a^b g(s)\,f(ds).$$

This formula is called *integration by parts*.

If the function $f\colon [a,b] \to \mathbb{R}$ and the derivative $g'\colon [a,b] \to \mathbb{R}$ of a function $g\colon [a,b] \to \mathbb{R}$ are both Riemann integrable then f is Riemann-Stieltjes integrable with respect to g and it satisfies the relation:

$$\int_a^b f(s)\,g(ds) = \int_a^b f(s)g'(s)\,ds. \tag{4.1.4}$$

Remark 4.1.1. In general, the integrand f can level out irregularities of the integrator g such that the limit in (4.1.3) exists and vice versa. However, if we ask for the largest class of integrators g such that at least every continuous function is Riemann-Stieltjes integrable with respect to g, then g must be of finite variation (see [17, Th. 56]). This result rules out the possibility to define the stochastic integral of a stochastic process $(\Phi(t) : t \in [0,T])$ with respect to a Brownian motion pathwise. That is, for fixed $\omega \in \Omega$ define an integral of the function $t \mapsto f_\omega(t) := \Phi(t)(\omega)$ with respect to the function $t \mapsto g_\omega(t) := W(t)(\omega)$, which turns out to be impossible since the function g_ω is of infinite variation for P-a.a. $\omega \in \Omega$ according to Theorem 3.3.13.

In calculus the Riemann integral is not the best approach to define an integral since it lacks some important properties concerning the interchangeability of taking the limit and integration. A much better approach is *Lebesgue integration* which is applied, for example in probability theory, in order to define the expectation of a random variable in a measure-theoretic way.

For all who are familiar with Lebesgue integration: in the book [10] by F. C. Klebaner I have found the following description:

In folklore the following analogy is used. Imagine that money is spread out on a floor. In the Riemann method of integration, you collect the money as you progress in the room. In the Lebesgue method, first you collect \$100 bills everywhere you can find them, then \$50, etc.

4.2. The Construction

Definition 4.2.1. *A stochastic process $(H(t) : t \in [0,T])$ is called* simple *if it is of the form*

$$H(t)(\omega) = \mathbb{1}_{\{0\}}(t)X_0(\omega) + \sum_{k=0}^{n-1} \mathbb{1}_{(t_k,t_{k+1}]}(t)X_k(\omega) \qquad \text{for all } t \in [0,T],\ \omega \in \Omega, \tag{4.2.5}$$

where $X_k : \Omega \to \mathbb{R}$ are $\mathscr{F}_{t_k}$-measurable random variables with $E[|X_k|^2] < \infty$, and $0 = t_0 < t_1 < \cdots < t_n = T$ and $n \in \mathbb{N}$. The space of all simple stochastic processes is denoted by $\mathscr{H}_0 := \mathscr{H}_0([0,T])$.

Note that each simple stochastic process $H \in \mathscr{H}_0$ is adapted to the given filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$. For a simple stochastic process $H \in \mathscr{H}_0$ of the form (4.2.5), we define the stochastic integral $I(H)$ with respect to a Brownian motion $(W(t) : t \geqslant 0)$ by

$$I(H) := \sum_{k=0}^{n-1} X_k\big(W(t_{k+1}) - W(t_k)\big). \tag{4.2.6}$$

Instead of the notion $I(H)$, one can use the more striking symbol $\int_0^T H(s)\,dW(s)$. It is easy to see that I is a linear mapping on $\mathscr{H}_0$, that is

$$I(\alpha H + \beta G) = \alpha I(H) + \beta I(G)$$

for all $H, G \in \mathscr{H}_0$ and $\alpha, \beta \in \mathbb{R}$.

Definition (4.2.6) is analogue to the sum on the right hand side in (4.1.3). For partitions $\pi_n = \{t_i^{(n)}\}_{i=0,\dots,m_n} \in P[0,T]$ and a function $f\colon [0,T] \to \mathbb{R}$ choose $\zeta_i^{(n)} := t_i^{(n)}$ for all

$i = 0, \dots, m_n - 1$ and $n \in \mathbb{N}$. Then the sum in the right hand side in (4.1.3) becomes

$$\sum_{i=0}^{m_n-1} f\left(\zeta_i^{(n)}\right)\left(g\left(t_{i+1}^{(n)}\right) - g\left(t_i^{(n)}\right)\right).$$

Financial Mathematics 4. Assume that $(W(t) : t \geqslant 0)$ models some share prices (with negative values) and that an agent trades X_k units of this share at the times $0, t_1, \dots, t_n$. Then her yield at time T is $I(H)$. We will make this idea more accurate later.

Of course, we are not satisfied with only integrating simple stochastic processes. For the extension of the class of admissible integrands the following result is essential.

Lemma 4.2.2. *(simple Itô's isometry)*
Each $H \in \mathscr{H}_0$ satisfies

$$E\left[|I(H)|^2\right] = \int_0^T E\left[|H(s)|^2\right]\, ds. \tag{4.2.7}$$

Example 4.2.3. Let Y be a $\mathscr{F}_1$-measurable random variable with $E[Y^2] = 2$ and define a stochastic process $(\Phi(t) : t \in [0,3])$ by

$$\Phi(t) = \begin{cases} 4, & \text{if } t \in [0,1], \\ Y, & \text{if } t \in (1,2], \\ 0, & \text{if } t \in (2,3]. \end{cases}$$

Write the stochastic integral

$$\int_0^3 \Phi(s)\, dW(s)$$

as the sum of two random variables and calculate its mean and variance.

Equation (4.2.7) is the crucial result which enables us to extend the space of integrands. Firstly, recall that $L_P^2(\Omega)$ is a Hilbert space with scalar product and corresponding norm

$$\langle X, Y\rangle_{L_P^2} := E\left[XY\right], \qquad \|Y\|_{L_P^2} = \left(E[|Y|^2]\right)^{1/2}.$$

Thus, we can rewrite the left hand side of (4.2.7) as

$$\|I(H)\|_{L_P^2}^2 = E\left[|I(H)|^2\right].$$

Also the right hand side can be interpreted as a norm. For this purpose, it is fruitful to consider a stochastic process $(\Phi(t) : t \in [0,T])$ as a mapping

$$\Phi : [0,T] \times \Omega \to \mathbb{R}.$$

We require that Φ is in the space[1]

$$L^2_{ds\otimes P}([0,T]\times\Omega) := \left\{ \Phi\colon [0,T]\times\Omega \to \mathbb{R} \text{ measurable}, \int_{[0,T]} E\left[|\Phi(s)|^2\right]\, ds < \infty \right\}.$$

This space can be equipped with a scalar product

$$\langle \Phi, \Psi\rangle_{L^2_{ds\otimes P}} := \int_0^T E[\Phi(s)\Psi(s)]\, ds,$$

and also with the corresponding norm

$$\|\Phi\|_{L^2_{ds\otimes P}} := \left(\int_0^T E[|\Phi(s)|^2]\, ds\right)^{1/2}.$$

It turns out that $L^2_{ds\otimes P}([0,T]\times\Omega)$ is a Hilbert space. Note, that the space $\mathscr{H}_0$ of simple stochastic processes is a subspace of $L^2_{ds\otimes P}([0,T]\times\Omega)$. Equation (4.2.7) can be rewritten as

$$\|I(H)\|_{L^2_P} = \|H\|_{L^2_{ds\otimes P}}. \tag{4.2.8}$$

This equation (4.2.8) is precisely the reason why the result of Lemma 4.2.2 is called *isometry:* it shows that if we understand the definition of the stochastic integral $I(H)$ as a mapping

$$I : \mathscr{H}_0 \to L^2_P(\Omega),$$

then this mapping preserves the norm. Here we use the fact that $\mathscr{H}_0$ is a subspace of $L^2_{ds\otimes P}([0,T]\times\Omega)$ and thus, it can be equipped with the same norm.

The recipe: Assume that for a stochastic process $\Phi \in L^2_{ds\otimes P}([0,T]\times\Omega)$ there exists a sequence $\{H_n\}_{n\in\mathbb{N}}$ of simple stochastic processes in $\mathscr{H}_0$ converging to Φ:

$$\lim_{n\to\infty} \|H_n - \Phi\|^2_{L^2_{ds\otimes P}} = \lim_{n\to\infty} \int_0^T E\left[|H_n(s) - \Phi(s)|^2\right]\, ds = 0.$$

By the triangle inequality of norms, it follows

$$\lim_{m,n\to\infty} \|H_m - H_n\|^2_{L^2_{ds\otimes P}} = 0.$$

Since (4.2.8) yields

$$\|I(H_m) - I(H_n)\|^2_{L^2_P} = \|H_m - H_n\|^2_{L^2_{ds\otimes P}} \qquad \text{for all } m,n\in\mathbb{N},$$

we obtain

$$\lim_{m,n\to\infty} \|I(H_m) - I(H_n)\|^2_{L^2_P} = 0,$$

[1] Do not bother the strange notation I use for denoting the space, see Remark 4.2.4.

that is $\{I(H_n)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in the Hilbert space $L^2_P(\Omega)$. A crucial property of Hilbert spaces is that each Cauchy sequence converges to an element in the Hilbert space. In our case this means, that there exists a random variable $Y \in L^2_P(\Omega)$ such that

$$\lim_{n\to\infty} \|I(H_n) - Y\|^2_{L^2_P} = 0.$$

Clearly, we call Y the stochastic integral of Φ and define

$$I(\Phi) := \int_0^T \Phi(s)\,dW(s) := \lim_{n\to\infty} \int_0^T H_n(s)\,dW(s) \qquad \text{in } L^2_P(\Omega).$$

Remark 4.2.4. If we define the product measure $\mu := ds \otimes P$ (where ds denotes the Lebesgue measure on $[0,T]$) then μ is a measure on the product σ-algebra $\mathfrak{B}([0,T]) \otimes \mathscr{A}$ over the space $S := [0,T] \times \Omega$. In this case the space of square-integrable functions is defined as

$$L^2_\mu(S) := \left\{\Phi\colon S \to \mathbb{R} \text{ measurable}, \int_S |\Phi(s)|^2\,\mu(ds) < \infty\right\}.$$

By writing out the definition of S and μ and writing the integral with respect to the probability measure P as expectation (using Fubini's theorem) we obtain

$$\begin{aligned} L^2_\mu(S) &= \left\{\Phi\colon [0,T]\times\Omega \to \mathbb{R} \text{ measurable}, \int_{[0,T]\times\Omega} |\Phi(s)(\omega)|^2\ (ds\otimes P)(ds,d\omega) < \infty\right\} \\ &= \left\{\Phi\colon [0,T]\times\Omega \to \mathbb{R} \text{ measurable}, \int_{[0,T]} E\left[|\Phi(s)|^2\right]\,ds < \infty\right\}. \end{aligned}$$

Thus, the space $L^2_{ds\otimes P}([0,T]\times\Omega)$, introduced above, is actually the space $L^2_\mu(S)$ which is well known to be a Hilbert space equipped with the norm

$$\|\Phi\|_{L^2_{ds\otimes P}} := \left(\int_S |\Phi(s)|^2\,\mu(ds)\right)^{1/2}$$

In even smaller print: in order that $\|\cdot\|_{L^2_{ds\otimes P}}$ is a norm, we must have the property that $\|\Psi\| = 0$ if and only if $\Psi = 0$. This can only be achieved by considering two stochastic processes Ψ and Φ as equal if

$$\int_{[0,T]} P\big(\Psi(t) \neq \Phi(t)\big)\,dt = 0.$$

In particular, stochastic processes which are modifications or versions of each other are considered to be equal in this interpretation.

We summarise the conclusions from the recipe in the following result with a small addition on uniqueness.

Theorem 4.2.5. *Assume that for $\Phi \in L^2_{ds\otimes P}([0,T]\times\Omega)$ there exists a sequence $(H_n)_{n\in\mathbb{N}}$ of simple stochastic processes $H_n \in \mathscr{H}_0$ such that*

$$\lim_{n\to\infty} \|H_n - \Phi\|^2_{L^2_{ds\otimes P}} = \lim_{n\to\infty} \int_0^T E\left[|H_n(s) - \Phi(s)|^2\right]\,ds = 0.$$

Then there exists a random variable $Y \in L^2_P(\Omega)$ which obeys

$$\lim_{n\to\infty} E\left[|I(H_n) - Y|^2\right] = 0.$$

The random variable Y does not depend on the approximating sequence $(H_n)_{n\in\mathbb{N}}$.

Due to Theorem 4.2.5 the question of extending the definition of the stochastic integral to a larger space of integrands becomes which stochastic processes $\Phi \in L^2_{ds\otimes P}([0,T]\times\Omega)$ can be approximated by a sequence of simple stochastic processes in $\mathscr{H}_0$. Note that the stochastic processes in $L^2_{ds\otimes P}([0,T]\times\Omega)$ are not required to be adapted whereas in the derivation of Lemma 4.2.2 adaptness of the simple stochastic process is essential. In fact it turns out that exactly all stochastic processes in $\Phi \in L^2_{ds\otimes P}([0,T]\times\Omega)$ which are adapted can be approximated by a sequence of simple stochastic processes in $\mathscr{H}_0$, that is stochastic processes in the space

$$\mathscr{H} := \left\{ \Phi : [0,T]\times\Omega \to \mathbb{R} \text{ measurable, adapted and } \int_0^T E\left[|\Phi(s)|^2\right] ds < \infty \right\}.$$

The space $\mathscr{H}$ is a subspace of $L^2_{ds\otimes P}([0,T]\times\Omega)$ and it is equipped with the same norm, but for ease of notation we define for all $\Phi \in \mathscr{H}$

$$\|\Phi\|_{\mathscr{H}} := \|\Phi\|_{L^2_{ds\otimes P}} = \left(\int_0^T E\left[|\Phi(s)|^2\right] ds \right)^{1/2}. \tag{4.2.9}$$

The set $\mathscr{H}$ is a linear space, that is,

$$\Phi, \Psi \in \mathscr{H},\ \alpha, \beta \in \mathbb{R} \ \Rightarrow\ \alpha\Phi + \beta\Psi \in \mathscr{H}$$

and the norm satisfies (as each norm does by definition) for Φ, $\Psi \in \mathscr{H}$ and $\alpha \in \mathbb{R}$:

$$\|\alpha\Phi\|_{\mathscr{H}} \leqslant |\alpha|\, \|\Phi\|_{\mathscr{H}},$$
$$\|\Phi + \Psi\|_{\mathscr{H}} \leqslant \|\Phi\|_{\mathscr{H}} + \|\Psi\|_{\mathscr{H}}$$

The approximation of elements in $\mathscr{H}$ by simple stochastic process is given in the following result:

Proposition 4.2.6. *For each $\Phi \in \mathscr{H}$ there exists a sequence $\{H_n\}_{n\in\mathbb{N}} \subseteq \mathscr{H}_0$ such that*

$$\lim_{n\to\infty} \|\Phi - H_n\|_{\mathscr{H}} = 0.$$

In order to complete our construction we define the stochastic integral according to our recipe:

Definition 4.2.7. *For every $\Phi \in \mathscr{H}$ the stochastic integral $I(\Phi)$ is defined as the unique random variable in $L^2_P(\Omega)$ which satisfies*

$$I(\Phi) := \int_0^T \Phi(s)\, dW(s) := \lim_{n\to\infty} I(H_n) \qquad \text{in } L^2_P(\Omega),$$

where $(H_n)_{n\in\mathbb{N}} \subseteq \mathscr{H}_0$ is a sequence satisfying $\|H_n - \Phi\|_{\mathscr{H}} \to 0$ as $n \to \infty$.

Remark 4.2.8. (pathwise vs. not pathwise)
As we pointed out in Remark 4.1.1, we can not define an integral with respect to a Brownian motion pathwise for a suitable large class of integrands by following the Riemann-Stieltjes approach. However, the definition of $I(H)$ for a simple stochastic process $H \in \mathscr{H}_0$ in (4.2.6) is pathwise, since you apply this definition for every $\omega \in \Omega$. But for $\Phi \in \mathscr{H}$ we define the integral $I(\Phi)$ as a limit in $L^2_P(\Omega)$, which is a space of functions where the limit does not necessarily coincide with a pointwise limit.

Theorem 4.2.9. *(Properties of the stochastic integral)*
Let Φ and Ψ be in $\mathscr{H}$ and α, β in $\mathbb{R}$. Then the stochastic integral satisfies:

(a) $\displaystyle\int_0^T (\alpha\Phi(s) + \beta\Psi(s))\, dW(s) = \alpha \int_0^T \Phi(s)\, dW(s) + \beta \int_0^T \Psi(s)\, dW(s).$

(b) $\displaystyle E\left[\left|\int_0^T \Phi(s)\, dW(s)\right|^2\right] = \int_0^T E\left[|\Phi(s)|^2\right] ds.$ *(Itô's isometry)*

By using the norms Itô's isometry can be expressed shorthand as

$$\|I(\Phi)\|_{L^2_P} = \|\Phi\|_{\mathscr{H}} \qquad \text{for all } \Phi \in \mathscr{H}.$$

Example 4.2.10. By approximating the identity on $\mathbb{R}$ by step functions we obtain

$$\int_0^T W(s)\, dW(s) = \tfrac{1}{2}W^2(T) - \tfrac{1}{2}T.$$

Note, that this formula does NOT correspond to the case if the integral on the left hand side would be a Stieltjes integral.

In a more abstract way the construction of the stochastic integral can be described in the the following way: firstly, we define a linear operator $I\colon \mathscr{H}_0 \to L^2_P(\Omega)$ by (4.2.6). Lemma 4.2.2 shows that this linear operator I is continuous. Since the set $\mathscr{H}_0$ is dense in $\mathscr{H}$ according to Proposition 4.2.6 and $L^2_P(\Omega)$ is complete, the continuous linear operator can be uniquely extended to a continuous linear operator I on $\mathscr{H}$. This is a classical argument often applied in analysis and related areas to linear operators and called *principle of extension of uniformly continuous mappings*.

4.3. The Integral Process

From a well defined integral we expect that if $\int_0^T$ exists then $\int_0^t$ exists for every $t \leqslant T$. This is satisfied by the stochastic integral since $\Phi \in \mathscr{H}$ implies $\Phi(\cdot)\mathbb{1}_{[0,t]}(\cdot) \in \mathscr{H}$ for each $t \leqslant T$ and we introduce the notation

$$I_t(\Phi) := \int_0^t \Phi(s)\, dW(s) := \int_0^T \Phi(s)\mathbb{1}_{[0,t]}(s)\, dW(s) \qquad \text{for each } t \in [0, T].$$

That is, for every $t \in [0,T]$ we obtain a random variable $I_t(\Phi)$. By collecting all $t \in [0,T]$ we want the family $(I_t(\Phi) : t \in [0,T])$ to fit in our framework of stochastic processes and martingales. Recall that for each $t \in [0,T]$ the random variable $I_t(\Phi)$ is defined as a limit in $L^2_P(\Omega)$. However, random variables in $L^2_P(\Omega)$ might differ on a set $N \in \mathscr{A}$ with $P(N) = 0$ but they are equal as elements in $L^2_P(\Omega)$. Consequenlty, for each $t \in [0,T]$ the random variable $I_t(\Phi)$ can be specified arbitrarily on a set $A_t \in \mathscr{A}$ with $P(A_t) = 0$. Since the set $[0,T]$ is uncountable and the union of uncountable null sets might equal the whole set Ω, it could happen that the whole family $(I_t(\Phi) : t \in [0,T])$ is ambiguous on Ω. We circumvent this problem by passing to a continuous modification of the stochastic integral.

Theorem 4.3.1. *If $\Phi \in \mathscr{H}$ then there exists a modification of the stochastic process*

$$\left(\int_0^t \Phi(s)\, dW(s) : t \in [0,T] \right),$$

which is a continuous martingale with respect to $\{\mathscr{F}_t\}_{t\geqslant 0}$ and with finite second moments; in particular, expectation and variance are given for all $t \in [0,T]$ by

$$E\left[\int_0^t \Phi(s)\, dW(s) \right] = 0, \tag{4.3.10}$$

$$E\left[\left| \int_0^t \Phi(s)\, dW(s) \right|^2 \right] = \int_0^t E\left[|\Phi(s)|^2 \right] ds, \tag{4.3.11}$$

Proof. The main part of this theorem is the existence of a continuous modification. The claimed equalities for the moments are either easily proved or just Itô's isometry and are mentioned here only for completeness. □

A stochastic process Φ in $\mathscr{H}$ defines the new stochastic process

$$\left(\int_0^t \Phi(s)\, dW(s) : t \in [0,T] \right)$$

which we often call the *integral process of* Φ and which we always assume to be a continuous martingale. Alternatively, we use the notation $(I_t(\Phi) : t \in [0,T])$ for the integral process of Φ.

Example 4.3.2. The calculation in Example 4.2.10 can be repeated for each $t \in [0,T]$ such that we obtain

$$\int_0^t W(s)\, dW(s) = \tfrac{1}{2} W^2(t) - \tfrac{1}{2} t \qquad \text{for all } t \in [0,T].$$

Theorem 4.3.1 implies that $(W^2(t) - t : t \in [0,T])$ is a martingale, which recovers the result from Corollary 3.2.4.

The next result considers the integral process $(I_t(\Phi) : t \in [0,T])$ stopped by a stopping time τ. Note that since the integral process is a stochastic process one can consider the stopped process $(I_{t\wedge\tau}(\Phi) : t \in [0,T])$ introduced in Proposition 1.2.5.

Theorem 4.3.3. *If τ is a stopping time and Φ is in $\mathscr{H}$ then*

$$I_{t\wedge\tau}(\Phi) = I_t\Big(\Phi(\cdot)\mathbb{1}_{\{[0,\tau]\}}(\cdot)\Big) \qquad P\text{-a.s. for all } t \in [0,T].$$

If we replace the stopping time τ by a constant $c \in [0,T]$ then the result of Theorem 4.3.3 is obvious. Even for a stopping time one might be misled to follow a pathwise argument in order to show the result. However, the truth is that the proof requires some more work.

4.4. Localising

Although the space $\mathscr{H}$ provides a rich class of admissible integrands, some natural examples are excluded. For example, even for a continuous function $f : \mathbb{R} \to \mathbb{R}$ the stochastic process $\Phi := (f(W(t)) : t \in [0,T])$ might not be in $\mathscr{H}$.

Example 4.4.1. For the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \exp(x^2)$ it follows that

$$E[|f(W(t))|^2] = E[\exp(2W^2(t))] = \begin{cases} \frac{1}{\sqrt{1-4t}}, & \text{if } t \in [0, \frac{1}{4}), \\ \infty, & \text{if } t \geqslant \frac{1}{4}. \end{cases}$$

Thus, if $T \geqslant \frac{1}{4}$, the process Φ is not in $\mathscr{H}$.

By the price that the integral will not be a martingale, we enlarge the space of admissible integrands:

$$\mathscr{H}_{\text{loc}} := \left\{ \Phi : [0,T] \times \Omega \to \mathbb{R} \text{ measurable, adapted, } P\left(\int_0^T |\Phi(s)|^2 \, ds < \infty\right) = 1 \right\}$$

Remark 4.4.2. Note, that since every stochastic process $\Phi \in \mathscr{H}$ satisfies

$$\int_0^T E\left[|\Phi(s)|^2\right] ds < \infty,$$

we can conclude by Exercise 3.4.6.(b) and Fubini's theorem that

$$\int_0^T |\Phi(s)|^2 \, ds < \infty \quad P\text{-a.s.}$$

Consequently, we obtain $\mathscr{H} \subseteq \mathscr{H}_{loc}$.

Example 4.4.3. The integrand $\big(f(W(t)) : t \in [0,T]\big)$ in Example 4.4.1 is in the space $\mathscr{H}_{\text{loc}}$. In fact, for every continuous function $f : \mathbb{R} \to \mathbb{R}$, the stochastic process $\big(f(W(t)) : t \in [0,T]\big)$ is in $\mathscr{H}_{\text{loc}}$:

$$\int_0^T |f(W(s))|^2 \, ds \leqslant T \sup_{s\in[0,T]} |f(W(s))|^2 < \infty \qquad P\text{-a.s.}$$

as f is continuous and W has continuous trajectories

By localising, that is by applying a good sequence of stopping times, we make elements from $\mathscr{H}_{\rm loc}$ treatable.

Definition 4.4.4. *An increasing sequence of stopping times* $\{\tau_n\}_{n\in\mathbb{N}}$ *is called* localising *for* $\Phi \in \mathscr{H}_{\rm loc}$ *if*

(i) $P\left(\bigcup_{n=1}^{\infty}\{\tau_n = T\}\right) = 1$;

(ii) $\big(\Phi(t)\mathbb{1}_{[0,\tau_n]}(t) : t \in [0,T]\big) \in \mathscr{H}$ *for each* $n \in \mathbb{N}$.

In fact, the space $\mathscr{H}_{\rm loc}$ is well chosen, since every element in this space has a localising sequence:

Proposition 4.4.5. *For any* $\Phi \in \mathscr{H}_{\rm loc}$, *the random elements*

$$\tau_n : \Omega \to [0,T], \qquad \tau_n(\omega) := T \wedge \inf\left\{t \geqslant 0 : \int_0^t |\Phi(s)(\omega)|^2 \, ds \geqslant n\right\}$$

define a localising sequence $\{\tau_n\}_{n\in\mathbb{N}}$ *for* Φ.

The recipe: for $\Phi \in \mathscr{H}_{\rm loc}$ let $\{\tau_n\}_{n\in\mathbb{N}}$ be a localising sequence. Due to condition (ii) in Definition 4.4.4, Theorem 4.3.1 implies that for each $n \in \mathbb{N}$ the random variable

$$X_n(t) := \int_0^t \Phi(s)\mathbb{1}_{[0,\tau_n]}(s)\, dW(s)$$

is well defined and forms a continuous martingale $(X_n(t) : t \in [0,T])$. Theorem 4.3.3 suggests that for $\Phi \in \mathscr{H}_{\rm loc}$ we define

$$\int_0^t \Phi(s)\, dW(s) := X_n(t) \qquad \text{for all } t \leqslant \tau_n. \tag{4.4.12}$$

By Condition (i) in Definition 4.4.4, one can always choose n large enough such that the left hand side in (4.4.12) is defined for each $t \in [0,T]$. In order that this definition makes sense it must be consistent, i.e. $X_m = X_n$ for all $0 \leqslant m \leqslant n$ and it must not depend on the localising sequence $\{\tau_n\}_{n\in\mathbb{N}}$. This is summarised in the following result:

Theorem 4.4.6. *For* Φ *in* $\mathscr{H}_{\rm loc}$ *let* $\{\tau_n\}_{n\in\mathbb{N}}$ *be a localising sequence. Then there exists an adapted continuous stochastic process* $Y = (Y(t) : t \in [0,T])$ *such that*

$$Y(t) = \lim_{n\to\infty} \int_0^t \Phi(s)\mathbb{1}_{[0,\tau_n]}(s)\, dW(s) \quad P\text{-a.s. for all } t \in [0,T]. \tag{4.4.13}$$

The limit Y *does not depend on the localising sequence* $\{\tau_n\}_{n\in\mathbb{N}}$ *and it is unique in the sense of indistinguishability.*

Definition 4.4.7. *The stochastic integral for $\Phi \in \mathscr{H}_{loc}$ is defined for all $t \in [0, T]$ by*

$$I_t(\Phi) := \int_0^t \Phi(s)\, dW(s) := Y(t),$$

where $(Y(t) : t \in [0, T])$ is the uniquely determined adapted continuous stochastic process satisfying (4.4.13).

The stochastic integral of a general integrand in $\mathscr{H}_{\text{loc}}$ might not have finite expectation, i.e. property (b) from Theorem 4.2.9 cannot hold. But we have the following:

Theorem 4.4.8. *(Properties of the stochastic integral)*
For Φ, $\Psi \in \mathscr{H}_{loc}$ and α, β in $\mathbb{R}$, the stochastic integral satisfies:

(a) $\displaystyle\int_0^T (\alpha\Phi(s) + \beta\Psi(s))\, dW(s) = \alpha \int_0^T \Phi(s)\, dW(s) + \beta \int_0^T \Psi(s)\, dW(s).$

(b) *no analogue.*

Similarly, we can not expect that the integral process is a martingale. However, we obtain:

Theorem 4.4.9. *If $\Phi \in \mathscr{H}_{loc}$ then*

$$\left(\int_0^t \Phi(s)\, dW(s) : t \in [0, T]\right)$$

is a continuous local martingale with respect to $\{\mathscr{F}_t\}_{t \geqslant 0}$.

4.5. Itô's formula

If $f\colon \mathbb{R} \to \mathbb{R}$ is a differentiable function then the fundamental theorem of calculus states (see (4.1.2)) that

$$f(t) - f(0) = \int_0^t f'(s)\, ds \qquad \text{for all } t \geqslant 0.$$

If furthermore $x\colon [0, T] \to \mathbb{R}$ is another differentiable function then by the chain rule, the fundamental theorem of calculus can be generalised to

$$f(x(t)) - f(x(0)) = \int_0^t f'(x(s))x'(s)\, ds \qquad \text{for all } t \in [0, T].$$

Rewriting this integral as a Riemann-Stieltjes integral (see (4.1.4)) results in

$$f(x(t)) - f(x(0)) = \int_0^t f'(x(s))\, x(ds) \qquad \text{for all } t \in [0, T]. \tag{4.5.14}$$

Naturally the question arises whether there is also a fundamental theorem for the stochastic integral.

Theorem 4.5.1. *(Itô's formula)*
If $f\colon \mathbb{R}\to\mathbb{R}$ is a twice continuously differentiable function then it obeys

$$f(W(t)) = f(0) + \int_0^t f'(W(s))\,dW(s) + \frac{1}{2}\int_0^t f''(W(s))\,ds$$

for all $t \geqslant 0$.

Remark 4.5.2.

(a) In comparison with the formula for the Riemann-Stieltjes integral in (4.5.14) there appears an additional term in the case of the Itô integral. This additional term is a consequence of the infinite variation of the Brownian motion.

(b) The generalisation of the stochastic integration to the space $\mathscr{H}_{\text{loc}}$ is the first time fruitful: we immediately know that the stochastic integral appearing in Theorem 4.5.1 is well defined since the integrand $(f'(W(t)) : t \in [0,T])$ is in $\mathscr{H}_{\text{loc}}$, see Example 4.4.3.

For the proof of Theorem 4.5.1 we need another result, which has its own interest.

Lemma 4.5.3. *A continuous function $g\colon \mathbb{R}\to\mathbb{R}$ satisfies*

$$\sum_{k=1}^n g\Big(W\big(t_{k-1}^{(n)}\big)\Big)\Big(W\big(t_k^{(n)}\big) - W\big(t_{k-1}^{(n)}\big)\Big) \xrightarrow{P} \int_0^T g\big(W(s)\big)dW(s) \qquad \text{as } n\to\infty,$$

where $t_k^{(n)} := \frac{k}{n}T$, $k = 0,\dots,n$.

Example 4.5.4. For $f\colon \mathbb{R}\to\mathbb{R}$, $f(x) = x^2$ we obtain

$$W^2(t) = t + 2\int_0^t W(s)\,dW(s) \qquad \text{for all } t \geqslant 0.$$

Compare with the efforts required in Example 4.2.10 in order to achieve the same result. Moreover since $W \in \mathscr{H}$, Theorem 4.3.1 implies that $(W^2(t) - t : t \geqslant 0)$ is a martingale with respect to $\{\mathscr{F}_t\}_{t\geqslant 0}$, a result which we derived in Corollary 3.2.4 by hand.

Example 4.5.5. For $f\colon \mathbb{R}\to\mathbb{R}$, $f(x) = \exp(x)$ we obtain

$$\exp(W(t)) = 1 + \int_0^t \exp(W(s))\,dW(s) + \frac{1}{2}\int_0^t \exp(W(s))\,ds \qquad \text{for all } t \geqslant 0.$$

Similarly as Example 4.5.4 recovers the result of Corollary 3.2.4.(a), we can obtain part (b) of the same Corollary by a modest generalisation of Itô's formula:

Theorem 4.5.6. *If $f\colon [0,T]\times\mathbb{R}\to\mathbb{R}$ is a function in $C^{1,2}([0,T]\times\mathbb{R})$ then it obeys*

$$\begin{aligned} &f(t, W(t)) \\ &\quad = f(0,0) + \int_0^t f_t(s,W(s))\,ds + \int_0^t f_x(s,W(s))\,dW(s) + \frac{1}{2}\int_0^t f_{xx}(s,W(s))\,ds \end{aligned}$$

for all $t \in [0,T]$.

The notion $C^{1,2}([0,T]\times\mathbb{R})$ denotes the space of all functions $f\colon [0,T]\times\mathbb{R}\to\mathbb{R}$ which are continuously differentiable in the first variable and twice continuously differentiable in the second variable. Accordingly, for $f\in C^{1,2}([0,T]\times\mathbb{R})$ we use the following abbreviation for the partial derivatives:

$$f_t(t,x) := \frac{\partial f}{\partial t}(t,x), \qquad f_x(t,x) := \frac{\partial f}{\partial x}(t,x), \qquad f_{xx}(t,x) := \frac{\partial^2 f}{\partial x^2}(t,x)$$

for all $t\in[0,T]$ and $x\in\mathbb{R}$. These are still deterministic functions! Only by applying these functions to a random argument they become random as well:

$$f_t(t,W(t)), \qquad f_x(t,W(t)), \qquad f_{xx}(t,W(t)).$$

Note, from a fussy point of view the expression $\frac{d}{dt}f(t,W(t))$ does not exist or does not make any sense, since it denotes the derivative of the function $t\mapsto f(t,W(t))$ which is not differentiable due to Corollary 3.3.7. For that reason make clear which function you differentiate.

Example 4.5.7. For a constant $c>0$ define the function

$$f\colon [0,T]\times\mathbb{R}\to\mathbb{R}, \qquad f(t,x)=\exp\left(cx-\tfrac{c^2}{2}t\right).$$

By applying Theorem 4.5.6 we conclude

$$f(t,W(t)) = 1 + c\int_0^t f(s,W(s))\,dW(s) \qquad \text{for all } t\in[0,T].$$

Moreover since $(f(t,W(t)) : t\in[0,T])\in\mathscr{H}$, Theorem 4.3.1 implies that the stochastic process

$$\left(\exp\left(cW(t)-\tfrac{c^2}{2}t\right) : t\in[0,T]\right)$$

is a martingale. Thus as promised, we also could recover the result of Corollary 3.2.4.(b).

Financial Mathematics 5. In the Black-Scholes model the share prices are modeled by a geometric Brownian motion $(S(t) : t\geqslant 0)$ which is defined by

$$S(t) := s_0\exp\left(\left(\mu-\tfrac{1}{2}\sigma^2\right)t+\sigma W(t)\right) \qquad \text{for all } t\geqslant 0,$$

where $s_0,\mu\in\mathbb{R}$ and $\sigma>0$ are constants. Theorem 4.5.6 implies

$$S(t) = s_0 + \mu\int_0^t S(s)\,ds + \sigma\int_0^t S(s)\,dW(s) \qquad \text{for all } t\geqslant 0.$$

Example 4.5.8. Define the functions

$$f\colon [0,T]\times\mathbb{R}\to\mathbb{R}, \qquad f(t,x)=e^{\frac{1}{2}t}\sin(x),$$
$$g\colon [0,T]\times\mathbb{R}\to\mathbb{R}, \qquad g(t,x)=e^{\frac{1}{2}t}\cos(x).$$

By applying Theorem 4.5.6 we obtain

$$e^{\frac{1}{2}t}\sin(W(t)) = \int_0^t e^{\frac{1}{2}s}\big(\cos W(s)\big)\,dW(s) \qquad \text{for all } t\in[0,T],$$
$$e^{\frac{1}{2}t}\cos(W(t)) = 1-\int_0^t e^{\frac{1}{2}s}\big(\sin W(s)\big)\,dW(s) \qquad \text{for all } t\in[0,T].$$

Let $\imath$ denote the unit on the imaginary line in the complex plane $\mathbb{C}$, i.e. $\imath=\sqrt{-1}$. Then we have the nice formula

$$e^{x+\imath y} = e^x\cos y + \imath e^x \sin y \qquad \text{for all } x,\, y\in\mathbb{R}\,. \tag{4.5.15}$$

By applying this formula to the function f and g above we obtain

$$\exp\big(\tfrac{1}{2}t+\imath W(t)\big) = g(t,W(t)) + \imath f(t,W(t)) \qquad \text{for all } t\in[0,T].$$

In analogy to (4.5.15), the stochastic processes $(f(t,W(t)) : t\in[0,T])$ and $(g(t,W(t)) : t\in[0,T])$ might be considered as the counterparts of the sin and cos function in the Itô calculus.

4.6. Itô Processes

Itô's formula in Theorem 4.5.6 considers only the transformation of a Brownian motion. The next step in a generalisation is to consider the transformation of a larger class of stochastic processes:

Definition 4.6.1. *A stochastic process $(X(t) : t\in[0,T])$ is called an* Itô process*, if it is of the form*

$$X(t) = X(0) + \int_0^t \Upsilon(s)\,ds + \int_0^t \Phi(s)\,dW(s) \qquad \textit{for all } t\in[0,T], \tag{4.6.16}$$

where $(\Upsilon(t) : t\in[0,T])$ and $(\Phi(t) : t\in[0,T])$ are adapted, measurable stochastic processes with

$$\int_0^T |\Upsilon(s)|\,ds<\infty \;\; P\text{-a.s. and} \qquad \int_0^T |\Phi(s)|^2\,ds<\infty \;\; P\text{-a.s.}$$

and $X(0)$ is an $\mathscr{F}_0$-measurable random variable.

Note, that the families $(\Upsilon(t) : t\in[0,T])$ and $(\Phi(t) : t\in[0,T])$ are assumed to be stochastic processes. The integral for Υ is understood in the pathwise sense, that is fix $\omega\in\Omega$ and then the condition on Υ guarantees that the ordinary integral for the function $s\mapsto\Upsilon(s)(\omega)$ exists:

$$Z(t)(\omega) := \int_0^t \Upsilon(s)(\omega)\,ds \qquad \text{for all } t\in[0,T].$$

In this way, for each $t\in[0,T]$ a random variable $Z(t):\Omega\to\mathbb{R}$ is defined. The second integral in (4.6.16) is the stochastic integral which is well defined since the assumptions on the stochastic process Φ are nothing else than to require $\Phi\in\mathscr{H}_{loc}$.

Example 4.6.2. In Example 4.5.4 we derive that

$$W^2(t) = t + 2\int_0^t W(s)\,dW(s) \qquad \text{for all } t \geqslant 0.$$

Consequently, $(W^2(t) : t \in [0,T])$ is an Itô process since

$$W^2(t) = 0 + \int_0^t \underbrace{1}_{=:\Upsilon(s)}\,ds + \int_0^t \underbrace{2W(s)}_{=:\Phi(s)}\,dW(s) \qquad \text{for all } t \geqslant 0.$$

Example 4.6.3. If f is a function in $C^{1,2}([0,T]\times\mathbb{R})$ then $(f(t,W(t)) : t \in [0,T])$ is an Itô process since Itô's formula in Theorem 4.5.6 implies for all $t \in [0,T]$

$$f(t,W(t)) = f(0,0) + \int_0^t \underbrace{\left(f_t(s,W(s))\,ds + \frac{1}{2}f_{xx}(s,W(s))\right)}_{=:\Upsilon(s)}\,ds + \int_0^t \underbrace{f_x\big(s,W(s)\big)}_{=:\Phi(s)}\,dW(s).$$

For an Itô process of the form (4.6.16) one often uses the shorthand notation

$$dX(t) = \Upsilon(t)\,dt + \Phi(t)\,dW(t). \tag{4.6.17}$$

This has no meaning on its own, it is only shorthand for (4.6.16). By means of Example 4.6.3 the same shorthand notation describes Itô's formula in Theorem 4.5.6 by

$$df(t,W(t)) = \Big(f_t(t,W(t)) + \tfrac{1}{2}f_{xx}(t,W(t))\Big)\,dt + f_x(t,W(t))\,dW(t).$$

Again, this has no meaning on its own.

The notation of an Itô process motivates to define an integral with respect to an Itô process. This does not require any efforts since we only use known integrals:

Definition 4.6.4. *Let X be an Itô process of the form (4.6.16) and $(\Psi(t) : t \in [0,T])$ be an adapted measurable stochastic process with*

$$\int_0^T |\Psi(s)\Upsilon(s)|\,ds < \infty \quad P\text{-a.s. and} \quad \int_0^T |\Psi(s)\Phi(s)|^2\,ds < \infty \ P\text{-a.s.}$$

Then define the random variable

$$\int_0^T \Psi(s)\,dX(s) := \int_0^T \Psi(s)\Upsilon(s)\,ds + \int_0^T \Psi(s)\Phi(s)\,dW(s). \tag{4.6.18}$$

The assumptions on the stochastic process Ψ yield that both integrals in (4.6.18) exist, the first one as an ordinary integral and the second one as a stochastic integral for integrands in $\mathscr{H}_{\text{loc}}$. Let X be an Itô process and let Ψ be as in Definition 4.6.4. Analogously to (4.6.17) a stochastic process $(Z(t) : t \in [0,T])$ defined by the stochastic integral

$$Z(t) := \int_0^t \Psi(s)\,dX(s)$$

is often described by the shorthand notation

$$dZ(t) = \Psi(t)\,dX(t). \tag{4.6.19}$$

Financial Mathematics 6. The geometric Brownian motion $(S(t) : t \geqslant 0)$ in Financial Mathematics 5 is an Itô process since

$$S(t) = s_0 + \int_0^t \underbrace{\mu S(s)}_{=:\Upsilon(s)} \, ds + \int_0^t \underbrace{\sigma S(s)}_{=:\Phi(s)} \, dW(s) \qquad \text{for all } t \geqslant 0$$

for some constants s_0, $\mu \in \mathbb{R}$ and $\sigma > 0$. If S describes the share prices in the Black-Scholes model and $(\Psi(t) : t \geqslant 0)$ denotes the number of shares held at time t, then $\int_0^t \Psi(s) \, dS(s)$ is the gain from trading the shares during the time interval $[0, t]$. Compare with Financial Mathematics 4.

Theorem 4.6.5. *(Itô's formula)*
Let X be an Itô process of the form (4.6.16) and $f\colon [0,T] \times \mathbb{R} \to \mathbb{R}$ be a function in $C^{1,2}([0,T] \times \mathbb{R})$. Then, for $t \in [0,T]$ we have P-a.s.:

$$\begin{aligned} f(t, X(t)) = f(0, X(0)) &+ \int_0^t f_t(s, X(s)) \, ds + \int_0^t f_x(s, X(s)) \Phi(s) \, dW(s) \\ &+ \int_0^t f_x(s, X(s)) \Upsilon(s) \, ds + \frac{1}{2} \int_0^t f_{xx}(s, X(s)) \Phi^2(s) \, ds. \end{aligned}$$

Remark 4.6.6. Theorem 4.6.5 implies that the process $(f(t, X(t)) : t \in [0, T])$ is an Itô process with the representation:

$$\begin{aligned} &f(t, X(t)) \\ &\quad = f(0, X(0)) + \int_0^t \left(f_t(s, X(s)) + f_x(s, X(s)) \Upsilon(s) + \frac{1}{2} f_{xx}(s, X(s)) \Phi^2(s) \right) ds \\ &\qquad + \int_0^t f_x(s, X(s)) \Phi(s) \, dW(s). \end{aligned}$$

By using the shorthand notation (4.6.17) for an Itô process one can write

$$\begin{aligned} df(t, X(t)) = &\left(f_t(t, X(t)) + f_x(t, X(t)) \Upsilon(t) + \tfrac{1}{2} f_{xx}(t, X(t)) \Phi^2(t) \right) dt \\ &+ f_x(t, X(t)) \Phi(t) \, dW(t). \end{aligned}$$

Also, by using the shorthand notation (4.6.19) for a stochastic integral with respect to an Itô process this reads as

$$df(t, X(t)) = f_t(t, X(t)) \, dt + f_x(t, X(t)) \, dX(t) + \tfrac{1}{2} f_{xx}(t, X(t)) \Phi^2(t) \, dt.$$

Note again, both shorthand notations have no meaning on their own, only if they are understood in the integrated version.

Example 4.6.7. Let $(X(t) : t \geqslant 0)$ be an Itô process of the form (4.6.16). For the function $f : [0, T] \times \mathbb{R} \to \mathbb{R}$, $f(t, x) = x^2$ we obtain that

$$X^2(t) = X^2(0) + 2 \int_0^t X(s) \, dX(s) + \int_0^t \Phi^2(s) \, ds \qquad \text{for all } t \in [0, T].$$

If $\Upsilon(s)=0$ for all $s\in[0,T]$ in the representation of X, this formula shows that

$$\left(X^2(t)-\int_0^t \Phi^2(s)\,ds:\,t\in[0,T]\right)$$

is a local martingale, compare with Corollary 3.2.4.(a).

Example 4.6.8. Let $\varphi\colon[0,T]\to\mathbb{R}$ be a deterministic continuous function satisfying $\int_0^T \varphi^2(s)\,ds<\infty$ and define[2] a stochastic process $(X(t):t\in[0,T])$ by

$$X(t)=-\int_0^t \varphi(s)\,dW(s)\qquad\text{for all } t\in[0,T].$$

Define the function

$$f:[0,T]\times\mathbb{R}\to\mathbb{R},\qquad f(t,x)=\exp x.$$

Since X is an Itô process, Theorem 4.6.5 implies

$$f(t,X(t))=1-\int_0^t f(s,X(s))\,\varphi(s)\,dW(s)\qquad\text{for all } t\in[0,T],$$

which in differential form has the nice appearance

$$df(t,X(t))=-f(t,X(t))\,dX(t)\qquad\text{for all } t\in[0,T].$$

Moreover, since $(f(t,X(t)):t\in[0,T])\in\mathscr{H}$ the stochastic process $(f(t,X(t)):t\in[0,T])$ is a continuous martingale. This martingale is called the *exponential martingale*. Compare with Corollary 3.2.4.(b).

Example 4.6.9.

(a) We define stochastic processes $(Y(t):t\in[0,T])$ and $(Z(t):t\in[0,T])$ by

$$Y(t):=\cos(W(t)),\qquad Z(t):=\sin(W(t)).$$

Then we obtain that Y and Z are Itô processes of the form

$$\begin{aligned} dY(t)&=-\tfrac{1}{2}\cos(W(t))\,dt-\sin(W(t))\,dW(t),\\ dZ(t)&=-\tfrac{1}{2}\sin(W(t))\,dt+\cos(W(t))\,dW(t). \end{aligned}$$

(b) Let $(Y(t):t\in[0,T])$ and $(Z(t):t\in[0,T])$ be the stochastic processes defined in part (a). Then we obtain

$$dY^2(t)=\Big(-(\cos W(t))^2+(\sin W(t))^2\Big)\,dt-2(\cos W(t))(\sin W(t))\,dW(t)$$

[2] The multiplication by (-1) is for later convenience.

and analogously

$$dZ^2(t) = \Big(-(\sin W(t))^2 + (\cos W(t))^2\Big)\,dt + 2(\sin W(t))(\cos W(t))\,dW(t).$$

It follows from $\sin^2(x) + \cos^2(x) = 1$ for all $x \in \mathbb{R}$ that $Y^2(t) + Z^2(t) = 1$ for all $t \in [0,T]$ which motivates us to call the stochastic process

$$((Y(t), Z(t)) : t \in [0,T])$$

Brownian motion on the circle.

4.7. The multidimensional Itô calculus

In this section we generalise the stochastic integration to $\mathbb{R}^d$ for a fixed dimension $d \in \mathbb{N}$. For distinction to the one-dimensional situation we denote here[3] the Euclidean norm in $\mathbb{R}^d$ by $\|\cdot\|_d$, i.e. $\|a\|_d = (a_1^2 + \cdots + a_d^2)^{1/2}$ for $a = (a_1, \dots, a_d) \in \mathbb{R}^d$. Let $(W(t) : t \geqslant 0)$ be a d-dimensional Brownian motion, i.e. $W(t) = (W_1(t), \dots, W_d(t))$ for all $t \geqslant 0$ and each $(W_i(t) : t \geqslant 0)$ is a one-dimensional Brownian motion and W_i and W_j are independent for $i \neq j$. An $n \times d$-dimensional stochastic process $((\Phi_{ij}(t))_{i=1,\dots,n;j=1,\dots,d} : t \geqslant 0)$ is said to be in $\mathscr{H}$ if each $(\Phi_{i,j}(t) : t \in [0,T])$ is in $\mathscr{H}$ for every $i = 1, \dots, n$, $j = 1, \dots, d$ and analogously for $\mathscr{H}_{\text{loc}}$. Then we define

$$\begin{aligned}
\int_0^T \Phi(s)\,dW(s) :&= \int_0^T \begin{pmatrix} \Phi_{1,1}(s) & \Phi_{1,2}(s) & \dots & \dots & \Phi_{1,d}(s) \\ \Phi_{2,1}(s) & \Phi_{2,2}(s) & \dots & \dots & \Phi_{2,d}(s) \\ \vdots & \vdots & & & \vdots \\ \Phi_{n,1}(s) & \Phi_{n,2}(s) & \dots & \dots & \Phi_{n,d}(s) \end{pmatrix} d \begin{pmatrix} W_1(s) \\ W_2(s) \\ \vdots \\ W_d(s) \end{pmatrix} \\
&:= \begin{pmatrix} \sum_{j=1}^d \int_0^T \Phi_{1,j}(s)\,dW_j(s) \\ \sum_{j=1}^d \int_0^T \Phi_{2,j}(s)\,dW_j(s) \\ \vdots \\ \sum_{j=1}^d \int_0^T \Phi_{n,j}(s)\,dW_j(s) \end{pmatrix}
\end{aligned}$$

Note that the stochastic integral is an n-dimensional random vector. The Itô's isometry has the following analogue:

Corollary 4.7.1.
If $(W(t) : t \geqslant 0)$ is a d-dimensional Brownian motion and $((\Phi_{i,j}(t))_{i=1,\dots,n;j=1,\dots,d} : t \geqslant 0)$ is an $n \times d$-dimensional stochastic process in $\mathscr{H}$ then

$$E\left[\left\|\int_0^t \Phi(s)\,dW(s)\right\|_n^2\right] = E\left[\int_0^T \|\Phi(s)\|_{HS}^2\,ds\right],$$

[3] this might be not consistent with other parts of the lecture notes where the Euclidean norm is just denoted by $|\cdot|$.

where for any matrix $c = (c_{ij})_{i=1,\dots,n;j=1,\dots,d} \in \mathbb{R}^{n\times d}$ *the* Hilbert-Schmidt norm *is defined by*

$$\|c\|_{HS} := \left(\sum_{i=1}^{n} \sum_{j=1}^{d} |c_{ij}|^2 \right)^{1/2}. \tag{4.7.20}$$

An n-dimensional stochastic process $(X(t) : t \geqslant 0)$ is called an *Itô process*, if it is of the form

$$X(t) = X(0) + \int_0^t \Upsilon(s)\, ds + \int_0^t \Phi(s)\, dW(s) \qquad \text{for all } t \geqslant 0, \tag{4.7.21}$$

where $(\Upsilon(t) : t \geqslant 0)$ is an n-dimensional, adapted stochastic process of the form $\Upsilon(t) = (\Upsilon_1(t), \dots, \Upsilon_n(t))$ for all $t \in [0,t]$ such that each component Υ_i is measurable and satisfies

$$\int_0^T |\Upsilon_i(s)|\, ds < \infty \qquad P\text{-a.s. for all } i = 1, \dots, n$$

and $(\Phi(t) : t \geqslant 0)$ is an $n \times d$-dimensional stochastic process such that each component $\Phi_{i,j}$ is measurable and satisfies

$$\int_0^T |\Phi_{ij}(s)|^2\, ds < \infty \qquad P\text{-a.s. for all } i = 1, \dots n, j = 1, \dots, d.$$

Theorem 4.7.2. *Let X be an Itô process of the form (4.7.21) and $f\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}$ be a function such that the partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x_i}$ and $\frac{\partial^2 f}{\partial x_i x_j}$ exist and that they are continuous. Then, for $t \in [0,T]$ we have P-a.s.:*

$$\begin{aligned} f(t, X(t)) = f(0, X(0)) &+ \int_0^t \frac{\partial f}{\partial t}(s, X(s))\, ds + \sum_{i=1}^{n} \int_0^t \Upsilon_i(s) \frac{\partial f}{\partial x_i}(s, X(s))\, ds \\ &+ \frac{1}{2} \sum_{i,j=1}^{n} \sum_{k=1}^{d} \int_0^t \Phi_{ik}(s)\Phi_{jk}(s) \frac{\partial^2 f}{\partial x_i \partial x_j}(s, X(s))\, ds \\ &+ \sum_{i=1}^{n} \sum_{j=1}^{d} \int_0^t \Phi_{ij}(s) \frac{\partial f}{\partial x_i}(s, X(s))\, dW_j(s). \end{aligned}$$

It is assumed in Theorem 4.7.2 that the components W_i and W_j of the d-dimensional Brownian motion W are independent for $i \neq j$, which corresponds to our Definition (3.0.1). However, a multi-dimensional Brownian motion can be defined more general by allowing that the components might be dependent. In this case the formula in 4.7.2 has some additional terms.

Recall the standard notion of multidimensional calculus; the gradient is defined by

$$\nabla f := \left(\tfrac{\partial f}{\partial x_1}, \dots, \tfrac{\partial f}{\partial x_n} \right),$$

and the Hessian matrix by

$$\nabla^2 f := \left(\tfrac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1}^{n}.$$

Let $\langle \cdot, \cdot \rangle$ denote the scalar product in the Euclidean space $\mathbb{R}^n$. Then we can rewrite Itô's formula in Theorem 4.7.2 as

$$\begin{aligned}
f(t, X(t)) &= f(0, X(0)) + \int_0^t \frac{\partial f}{\partial t}(s, X(s))\, ds \\
&\quad + \int_0^t \Big(\Big\langle \Upsilon(s),\, (\nabla f)(s, X(s)) \Big\rangle + \frac{1}{2}\mathrm{Tr}\Big[\Phi(s)\Phi^T(s)\big((\nabla^2 f)(s, X(s))\big)\Big]\Big)\, ds \\
&\quad + \sum_{j=1}^{d} \int_0^t \Big\langle (\nabla f)(s, X(s)),\, \Phi_{\bullet j}(s) \Big\rangle\, dW_j(s),
\end{aligned}$$

with $\Phi_{\bullet j}(s) := (\Phi_{1j}(s), \dots, \Phi_{nj}(s))$ for all $j = 1, \dots, d$. Often one uses the shorthand notation

$$\begin{aligned}
df(t, X(t)) = \Big(&f_t(t, X(t)) + \big\langle \Upsilon(t),\, (\nabla f)(t, X(t))\big\rangle + \tfrac{1}{2}\mathrm{Tr}\big[\Phi(t)\Phi^T(t)\big((\nabla^2 f)(t, X(t))\big)\big]\Big)\, dt \\
&+ \big\langle (\nabla f)(s, X(s)),\, \Phi(s)\, dW(s)\big\rangle.
\end{aligned}$$

Remark 4.7.3. Suppose that the Itô process X takes values only in an open set $U \subseteq \mathbb{R}^n$. Then Theorem 4.7.2 can be generalised to a function $f\colon [0,T] \times U \to \mathbb{R}$ by requiring that its derivatives $\frac{\partial f}{\partial t}$ in $[0,T]$, $\frac{\partial f}{\partial x_i}$ and $\frac{\partial^2 f}{\partial x_i x_j}$ in U exist and that they are continuous. This can be derived from Theorem 4.7.2 by localising. An example for the case $n = d = 1$ is the geometric Brownian motion S in Financial Mathematics 5, which takes only values in $(0, \infty)$ if the initial value s_0 is positive. Then we can use Itô's formula to calculate the form of $(\sqrt{S(t)}\, : t \in [0,T])$.

Example 4.7.4. (Product formula)
For independent Brownian motions $(W_1(t) : t \geqslant 0)$ and $(W_2(t) : t \geqslant 0)$ let X and Y be of the form

$$\begin{aligned}
X(t) &= X(0) + \int_0^t \Upsilon_1(s)\, ds + \int_0^t \Phi_1(s)\, dW_1(s) \\
Y(t) &= Y(0) + \int_0^t \Upsilon_2(s)\, ds + \int_0^t \Phi_2(s)\, dW_2(s).
\end{aligned}$$

Then their product satisfies

$$d(X(t)Y(t)) = Y(t)\, dX(t) + X(t)\, dY(t).$$

4.8. Exercises

1. Let Y_1 be a $\mathscr{F}_1^W$-measurable random variable with $E[Y_1^2] = 2$ and let Y_2 be a $\mathscr{F}_2^W$-measurable random variable with $E[Y_2^2] = 4$. Define a stochastic process $(\Phi(t) : t \in$

$[0,4]$) by

$$\Phi(t) = \begin{cases} 2, & \text{if } t \in [0,1], \\ Y_1, & \text{if } t \in (1,2], \\ Y_2, & \text{if } t \in (2,3], \\ 0, & \text{if } t \in (3,4]. \end{cases}$$

(a) Write the stochastic integral

$$\int_0^4 \Phi(s)\, dW(s)$$

as the sum of three random variables.

(b) Calculate the mean and variance of the stochastic integral

$$\int_0^4 \Phi(s)\, dW(s).$$

2. Find for $0 \leqslant a \leqslant b$ the expectation and variance of the following random variables:
   (a) $X = \int_a^b |W(t)|\, dW(t)$.
   (b) $X = \int_a^b \sqrt{t}\, e^{W(t)}\, dW(t)$.
   (c) $X = \int_a^b \operatorname{sgn}(W(t))\, dW(t)$ where $\operatorname{sgn}(x) = \begin{cases} 0, & \text{if } x = 0, \\ \frac{x}{|x|}, & \text{if } x \neq 0. \end{cases}$

   The stochastic process $(\int_0^t \operatorname{sgn}(W(s))\, dW(s) : t \geqslant 0)$ is a Brownian motion.

3. Show at least in two different ways that $M(t) := \exp(W(t) - \frac{1}{2}t)$ defines a martingale $(M(t) : t \geqslant 0)$.

4. Let $f : [0,T] \to \mathbb{R}$ be a continuous function. For a partition $\pi = \{t_i\}_{i=0,\dots,m}$ of $[0,T]$ define the sums

$$L(\pi) = \sum_{k=0}^{m-1} f(t_k)\big(f(t_{k+1}) - f(t_k)\big),$$
$$R(\pi) = \sum_{k=0}^{m-1} f(t_{k+1})\big(f(t_{k+1}) - f(t_k)\big).$$

   (a) Show that ($\natural$)

$$L(\pi) = \frac{1}{2}\left(f^2(T) - f^2(0) - \sum_{k=0}^{m-1} \big(f(t_{k+1}) - f(t_k)\big)^2\right),$$
$$R(\pi) = \frac{1}{2}\left(f^2(T) - f^2(0) + \sum_{k=0}^{m-1} \big(f(t_{k+1}) - f(t_k)\big)^2\right).$$

(b) Conclude from part (a) that for the existence of the Riemann-Stieltjes integral $\int_0^T f(s)\,df(s)$ it is necessary that the quadratic variation vanishes, that is

$$\lim_{|\pi|\to 0} \sum_{k=0}^{m-1} |f(t_{k+1}) - f(t_k)|^2 \to 0,$$

where the limit is taken over all partitions $\pi \in P(0,T)$ with $|\pi| \to 0$. (♮)

5. Show that for all $t \geqslant 0$ and $m \geqslant 2$, the Brownian motion satisfies

$$W^m(t) = m\int_0^t W^{m-1}(s)\,dW(s) + \frac{m(m-1)}{2}\int_0^t W^{m-2}(s)\,ds.$$

6. Let $(X(t) : t \in [0,T])$ be a stochastic process in $\mathscr{H}_0$ and $0 \leqslant s \leqslant t \leqslant T$. Show the following

   (a) $E\left[\int_s^t X(u)\,dW(u)|\mathscr{F}_s\right] = 0.$

   (b) $E\left[\left(\int_s^t X(u)\,dW(u)\right)^2 |\mathscr{F}_s\right] = E\left[\int_s^t X^2(u)\,du|\mathscr{F}_s\right].$

   (c) How can you generalise the results in (a) and (b) to $X \in \mathscr{H}$?

7. Define the random variable

$$X(t) := \int_0^t W^2(s)\,dW(s)$$

   (a) Show that

$$X(t) = \frac{1}{3}W^3(t) - \int_0^t W(s)\,ds \qquad \text{for all } t \geqslant 0$$

   by the same method as in Example 4.2.10 and verify directly that $(X(t) : t \geqslant 0)$ is a martingale. (∗∗)

   (b) Repeat part (a) by another arguments.

8. (a) Verify that

$$\int_0^t s\,dW(s) = tW(t) - \int_0^t W(s)\,ds \qquad P\text{-a.s. for all } t \geqslant 0.$$

   (b) Show that the stochastic process $(X(t) : t \geqslant 0)$ defined by

$$X(t) = \exp\left(\tfrac{1}{2}t\right)\cos W(t)$$

   satisfies the equation

$$X(t) = 1 - \int_0^t \exp\left(\tfrac{1}{2}s\right)\sin W(s)\,dW(s) \qquad \text{for all } t \geqslant 0.$$

(c) Show by calculating the differential dY that the process $(Y(t) : t \geqslant 0)$ defined by

$$Y(t) = \frac{1}{3}(W(t))^3 - tW(t),$$

is a martingale.

9. (a) Show that if $\Psi \in \mathscr{H}$ is such that the mapping $(s,t) \mapsto E[\Psi(s)\Psi(t)]$ is continuous then

$$\int_0^T \Psi(s)\, dW(s) = \lim_{|\pi|\to 0} \sum_{k=0}^{n-1} \Psi(t_k)(W(t_{k+1}) - W(t_k)) \qquad \text{in } L^2(\Omega),$$

where the limit is taken over all partitions $\pi = \{t_k\}_{k=0,\dots,n}$ of $[0,T]$.

(b) One can use (a) to define a λ-*integral* for $\lambda \in [0,1]$ by

$$(\lambda)\!-\!\int_0^T \Phi(s)\, dW(s) := \lim_{|\pi|\to 0} \sum_{k=0}^{n-1} \Psi(\zeta_k)(W(t_{k+1}) - W(t_k)) \qquad \text{in } L^2(\Omega),$$

where $\zeta_k = (1-\lambda)t_k + \lambda t_{k+1}$ for all adapted stochastic processes $(\Psi(t) : t \in [0,t])$ for which this limit exists. If $\lambda = 0$ the λ-integral coincide with the Itô integral due to part (a). For $\lambda = \frac{1}{2}$ this integral is called the *Stratonovich integral.*

(i) If the adapted stochastic process $(\Psi(t) : t \in [0,T])$ has Lipschitz continuous paths then

$$(\lambda)\!-\!\int_0^T \Psi(s)\, dW(s) = \int_0^T \Psi(s)\, dW(s),$$

i.e. the λ-integral and the Itô integral coincide. ($\natural$)

(ii) Show that ($\natural$)

$$(\lambda)\!-\!\int_0^T W(s)\, dW(s) = \tfrac{1}{2}W^2(t) - \left(\tfrac{1}{2} - \lambda\right)t.$$

(iii) Show that if $f : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function then ($\natural$)

$$(\tfrac{1}{2})\!-\!\int_0^T f(W(s))\, dW(s) = \int_0^T f(W(s))\, dW(s) + \tfrac{1}{2}f'(W(s))\, dWs.$$

10. The Laplace operator of a function $f \in C^2(\mathbb{R}^n)$ is defined by

$$\Delta f\colon\ \mathbb{R}^n \to \mathbb{R}, \qquad \Delta f := \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2}.$$

The function f is called *harmonic on an open set* $D \subseteq \mathbb{R}^n$ if $\Delta f = 0$ on D. Let W be an n-dimensional Brownian motion.

(a) Show that Itô's formula for f in $C^2(\mathbb{R}^n)$ is given by

$$f(W(t)) = f(W(0)) + \int_0^t \nabla f(W(s))\, dW(s) + \frac{1}{2}\int_0^t \Delta f(W(s))\, ds \qquad \text{for all } t \in [0,T].$$

(b) Let $D \subseteq \mathbb{R}^d$ be an open set and define for an open set $S \subseteq D$ with $0 \in S$ and $\overline{S} \subseteq D$ the stopping time

$$\tau := \inf\{t \geqslant 0 : W(t) \notin S\}.$$

Show that if f is harmonic, then $(f(W(\tau \wedge t)) : t \geqslant 0)$ is a local martingale.

This result can be used to show that Brownian motion in $\mathbb{R}^d$ is recurrent only if $d = 1$ or $d = 2$. A Brownian process is recurrent if it returns almost surely to any neighborhood of the origin infinitely often.

5 Stochastic Differential Equations

An ordinary differential equation is of the form

$$x'(t) = f(x(t)) \qquad \text{for all } t \in [0,T] \qquad \text{and } x(0) = x_0, \tag{5.0.1}$$

where $x_0 \in \mathbb{R}$ is a given initial value and $f\colon \mathbb{R} \to \mathbb{R}$ is a given function. A solution is a function $x\colon [0,T] \to \mathbb{R}$ which satisfies this equation. A simple example is

$$x'(t) = \alpha x(t) \qquad \text{for all t} \in \text{[0, T]} \qquad \text{and } x(0) = x_0,$$

for a constant $\alpha \in \mathbb{R}$, which is solved by the function $x(t) = x_0 \exp(\alpha t)$ for all $t \in [0,T]$. The differential equation (5.0.1) can be written as an integral equation:

$$x(t) = x(0) + \int_0^t f(x(s))\, ds \qquad \text{for all t} \in \text{[0, T]} \tag{5.0.2}$$

and it has the same solution. In this chapter we introduce analogously stochastic differential equations based on the Itô calculus.

5.1. The Equation

From a modelling point of view one can think of perturbing the deterministic dynamic described in equation (5.0.1) by a random noise modelled by a Brownian motion:

$$X(t) = X(0) + \int_0^t f(X(s))\, ds + \int_0^t g(X(s))\, dW(s) \qquad \text{for all } t \in [0,T]$$

for some functions $f,\, g\colon \mathbb{R}\to\mathbb{R}$. Since the Brownian motion is not differentiable we use the integrated version (5.0.2). The first integral can be understood pathwise, that is one fix $\omega\in\Omega$ and considers $s\mapsto f(X(s)(\omega))$ as a deterministic function. If this function is integrable for each $\omega\in\Omega$ then the integral

$$Z(\omega) := \int_0^t f(X(s)(\omega))\,ds$$

exists and $Z\colon\Omega\to\mathbb{R}$ is a random variable. The second integral is the stochastic integral defined in the previous chapter. It exists if the stochastic process $(g(X(t)) : t\in[0,T])$ is in $\mathscr{H}$ or more general in $\mathscr{H}_{loc}$.

In fact, these kind of equations are called *stochastic differential equations*. If we also allow the coefficients to depend on time we arrive at the following stochastic differential equation which we write in the differential form:

$$\begin{aligned} dX(t) &= f(t,X(t))\,dt + g(t,X(t))\,dW(t) \qquad \text{for all } t\in[0,T],\\ X(0) &= X_0, \end{aligned} \tag{5.1.3}$$

where the coefficients are given by measurable functions

$$f,\, g\colon [0,T]\times\mathbb{R}\to\mathbb{R}$$

and the initial value is a given $\mathscr{F}_0$-measurable, random variable $X_0\colon\Omega\to\mathbb{R}$ with $E[|X_0|^2]<\infty$. The term f is called the *drift coefficient* and the term g is called the *diffusion coefficient*[1]. The underlying probability space $(\Omega,\mathscr{A},P)$ is equipped with a filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ which satisfies the usual conditions. The noise is modelled by a Brownian motion $(W(t) : t\geqslant 0)$ and we assume the same conditions as mentioned in the very beginning of Chapter 4. As before the equation (5.1.3) does not have a meaning on its own, but the following definition explains what we understand as a solution.

Definition 5.1.1. *A stochastic process $X=(X(t) : t\in[0,T])$ is a solution of the stochastic differential equation (5.1.3) if*

(a) *X is adapted;*

(b) *$X(0)=X_0$ P-a.s.*

(c) $P\left(\int_0^T |f(s,X(s))|\,ds + \int_0^T |g(s,X(s))|^2\,ds<\infty\right)=1.$

(d) *For all $t\in[0,T]$, the stochastic process X satisfies*

$$X(t) = X(0) + \int_0^t f(s,X(s))\,ds + \int_0^t g(s,X(s))\,dW(s) \qquad P\text{-a.s.}$$

[1]Some authors call g dispersion coefficient and g^2 diffusion coeeficient. In Financial Mathematics g is called volatility.

Condition (c) guarantees that both integrals appearing in (d) are well defined. In particular, since the solution X is required to be adapted due to Condition (a) and $g\colon [0,T]\times\mathbb{R}\to\mathbb{R}$ is measurable, the stochastic process $(g(t,X(t)) : t\in[0,T])$ is adapted and in $\mathscr{H}_{\text{loc}}$ because of Condition (c).

Often the initial condition is just a deterministic value $x_0\in\mathbb{R}$ instead of an $\mathscr{F}_0$-measurable random variable $X_0\colon\Omega\to\mathbb{R}$. A further modification of the considered equation (5.1.3) is the situation where the solution is required to exist on the entire time axis $[0,\infty)$ and not only on an interval $[0,T]$. This can easily be achieved by taking Definition (5.1.1) *locally in time*, i.e. a stochastic process $(X(t) : t\geqslant 0)$ is called a *solution* if $(X(t) : t\in[0,T])$ satisfies Definition 5.1.1 for all $T>0$. In fact, our definition of a solution depends on the considered time interval $[0,T]$, since not for all $T>0$ a solution must exist.

Example 5.1.2. Consider the example of a deterministic differential (integral) equation:

$$x(t) = 1 + \int_0^t x^2(s)\,ds \qquad \text{for all t} \in [0, \text{T}]\,.$$

The solution is given by $x(t)=(1-t)^{-1}$ but only for $T<1$.

Recall that a solution of a deterministic differential equation need not to be unique. A simple example of this fact is the following:

Example 5.1.3. The deterministic differential (integral) equation

$$x(t) = \int_0^t 2\sqrt{x(s)}\;ds \qquad \text{for all t} \in [0, \text{T}],$$

has infinitely many solutions. For example, $x(t)=0$ and $x(t)=t^2$ for $t\in[0,T]$ are both solutions.

For stochastic differential equations there are different notions of uniqueness for a solution of (5.1.3). We use the strongest one and say that a solution is unique if all other solutions are indistinguishable.

Definition 5.1.4. *A solution* $(X(t) : t\in[0,t])$ *of* (5.1.3) *is called* unique *if for any other solution* $(Y(t) : t\in[0,T])$ *of* (5.1.3) *it follows*

$$P\big(X(t)=Y(t) \quad \textit{for all } t\in[0,T]\big)=1.$$

The standard conditions to guarantee the existence of a unique solution of a deterministic differential equation are the Lipschitz and linear growth conditions. The same applies to stochastic differential equations, and even the proof is very similar based on Picard's iteration scheme.

Theorem 5.1.5. *Suppose that there exists a constant $c>0$ such that the following are satisfied for all $t\in[0,T]$ and all $x,y\in\mathbb{R}$:*

$$|f(t,x)-f(t,y)|+|g(t,x)-g(t,y)| \leqslant c\,|x-y|\,, \tag{5.1.4}$$

$$|f(t,x)|^2+|g(t,x)|^2 \leqslant c^2(1+|x|^2). \tag{5.1.5}$$

Then there exists a unique solution $(X(t) : t\in[0,T])$ of the stochastic differential equation (5.1.3)*. Moreover, the solution X has the following properties:*

(i) *X has continuous trajectories;*

(ii) *there exists a constants $\alpha \geqslant 0$ such that*

$$E\left[\sup_{t\in[0,T]} |X(t)|^2\right] \leqslant \alpha\left(1+E\left[\,|X_0|^2\,\right]\right).$$

Proof. See Theorem 5.2.9 in [9]. □

Condition (5.1.4) is called *Lipschitz condition* and (5.1.5) *linear growth condition.* If the coefficients f and g do not depend on t, i.e. if they are of the form $f\colon \mathbb{R}\to\mathbb{R}$ and $g\colon \mathbb{R}\to\mathbb{R}$, then the Lipschitz condition (5.1.4) is satisfied if and only if f and g are Hölder continuous of order 1; see Section 3.3 for the definition of Hölder continuity. Due to the inequalities

$$1+a^2 \leqslant (1+a)^2 \leqslant 2(1+a^2) \qquad \text{for all } a\in\mathbb{R}_+,$$

Condition (5.1.5) is satisfied if and only if there exists a constant $c' > 0$ such that for all $t\in[0,T]$ and all $x,y\in\mathbb{R}$:

$$|f(t,x)|+|g(t,x)| \leqslant c'(1+|x|).$$

This is the reason for the term linear in linear growth condition.

Example 5.1.6. The stochastic differential equation

$$dX(t) = -\frac{X(t)}{1+t}\,dt + \frac{1}{1+t}\,dW(t) \qquad \text{for all t} \in [0,\text{T}], \quad X(0)=0,$$

has a unique solution, since the coefficients

$$\begin{aligned} f\colon [0,T]\times\mathbb{R}\to\mathbb{R}, &\qquad f(t,x) := \frac{-x}{1+t},\\ g\colon [0,T]\times\mathbb{R}\to\mathbb{R}, &\qquad g(t,x) := \frac{1}{1+t}, \end{aligned}$$

satisfy the conditions in Theorem 5.1.5. Itô's formula in Theorem 4.6.5 shows that the solution $(X(t): t\in[0,T])$ is given by

$$X(t) = \frac{W(t)}{1+t}.$$

Financial Mathematics 7. Many models in financial mathematics are based on stochastic differential equations. By far the most important one is the Black-Scholes model where the share prices $(S(t): t\in[0,T])$ are modelled by the stochastic differential equation

$$dS(t) = \mu S(t)\,dt + \sigma S(t)\,dW(t) \qquad \text{for all t} \in [0,\text{T}],$$

where $\mu\in\mathbb{R}$ and $\sigma>0$ are constants.

Various models for interest rates are described by stochastic differential equations, for example the Vasicek model

$$\begin{aligned} dR(t) &= -a(R(t)-b)\,dt + \sigma\,dW(t) \qquad \text{for all } t \in [0,T],\\ R(0) &= r_0, \end{aligned}$$

where $r_0, a, b, \sigma > 0$ are constants. The draw back of the rather simple Vasicek model is that the solution R can be negative which does not fit to the purpose of modelling interest rates. This can be fixed by considering the Cox-Ingersoll-Ross model, where the interst rates R are modelled by

$$\begin{aligned} dR(t) &= -a(R(t)-b)\,dt + \sigma\sqrt{R(t)}\;dW(t) \qquad \text{for all } t \in [0,T],\\ R(0) &= r_0, \end{aligned}$$

where $r_0, a, b, \sigma > 0$ are some constants. Once can show that if the parameters a , b and σ are in a specific area then the solution R is positive. Note, that the existence of a solution of this equation does not follow from Theorem 5.1.5 but it can be established by some other arguments, see Exercise 5.7.7.

5.2. Example: Ornstein-Uhlenbeck Process

As an example we consider a linear stochastic differential equation with additive noise. Although this equation is rather simple it was the starting point of the idea of stochastic differential equations and nevertheless, it is applied in various models.

Definition 5.2.1. *Let $\mu \in \mathbb{R}$, $u_0 \in \mathbb{R}$ and $\sigma^2 > 0$ be given constants. The unique solution $(U(t) : t \geqslant 0)$ of the stochastic differential equation*

$$\begin{aligned} dU(t) &= \mu U(t)\,dt + \sigma\,dW(t) \qquad \textit{for all } \mathrm{t} \geqslant 0,\\ U(0) &= u_0, \end{aligned} \tag{5.2.6}$$

is called Ornstein-Uhlenbeck process*.*

The coefficients of the stochastic differential equation (5.2.6) are given by

$$\begin{aligned} f&\colon [0,T]\times\mathbb{R}\to\mathbb{R}, \qquad f(t,x) := \mu x,\\ g&\colon [0,T]\times\mathbb{R}\to\mathbb{R}, \qquad g(t,x) := \sigma. \end{aligned}$$

Since these functions satisfy the conditions in Theorem 5.1.5 the definition above makes sense, i.e. the solution of the considered stochastic differential equation exists and is unique. In this case the solution can even be written explicitly.

Proposition 5.2.2. *The solution of* (5.2.6) *is given by*

$$U(t) = u_0 e^{\mu t} + \sigma \int_0^t e^{\mu(t-s)}\,dW(s) \qquad \textit{for all } \mathrm{t} \geqslant 0\,.$$

Moreover, for each $t \geqslant 0$ the random variable $U(t)$ is normally distributed with moments and variance governed by

$$
\begin{aligned}
E[U(t)] &= u_0 \exp(\mu t), \qquad E\big[(U(t))^2\big] = \tfrac{-\sigma^2}{2\mu} + \Big(u_0^2 + \tfrac{\sigma^2}{2\mu}\Big) e^{2\mu t}, \\
\mathrm{Var}[U(t)] &= \tfrac{\sigma^2}{2\mu}\left(e^{2\mu t} - 1\right).
\end{aligned}
$$

The claim in the last Proposition that the random variable $U(t)$ is normally distributed for each $t \geqslant 0$ is based on the following general result:

Lemma 5.2.3. *Let $f : [0,T] \to \mathbb{R}$ be a function such that $\int_0^T |f(s)|^2 \, ds < \infty$. Then the random variables*

$$X(t) := \int_0^t f(s)\, dW(s) \qquad \text{for } t \in [0,T],$$

are normally distributed with $E[X(t)] = 0$ and $Var[X(t)] = \int_0^t f^2(s)\, ds$ and

$$Cov(X(s), X(t)) = \int_0^{s\wedge t} f^2(u)\, du.$$

for $s, t \in [0,T]$.

Lemma 5.2.4. *If $f : [0,T] \to \mathbb{R}$ is a differentiable function then*

$$\int_0^t f(s)\, dW(s) = f(t)W(t) - \int_0^t f'(s)W(s)\, ds \qquad \text{for all } t \in [0,T].$$

5.3. Example: Geometric Brownian Motion

The most popular model in Financial Mathematics is based on the Geometric Brownian Motion, which we define as the solution of a stochastic differential equation:

Definition 5.3.1. *Let $\mu \in \mathbb{R}$, $g_0 > 0$ and $\sigma^2 > 0$ be given constants. The unique solution $(G(t) : t \geqslant 0)$ of the stochastic differential equation*

$$
\begin{aligned}
dG(t) &= \mu G(t)\, dt + \sigma G(t)\, dW(t) \qquad \text{for all t} \geqslant 0, \\
G(0) &= g_0,
\end{aligned}
\tag{5.3.7}
$$

is called Geometric Brownian Motion.

The coefficients for the stochastic differential equation (5.3.7) are given by the functions

$$f\colon [0,T]\times\mathbb{R} \to \mathbb{R}, \qquad f(t,x) := \mu x, \tag{5.3.8}$$

$$g\colon [0,T]\times\mathbb{R} \to \mathbb{R}, \qquad g(t,x) := \sigma x, \tag{5.3.9}$$

Obviously, these functions satisfy the linear growth condition and the Lipschitz condition in Theorem 5.1.5 and thus, the definition above makes sense. By applying Itô's formula, we can even represent the solution explicitly.

Proposition 5.3.2. *The solution of* (5.3.7) *is given by*

$$G(t) = g_0 \exp\Big(\big(\mu - \tfrac{1}{2}\sigma^2\big)\, t + \sigma W(t)\Big) \qquad \textit{for all } \mathrm{t} \geqslant 0\,.$$

Moreover, for each $t \geqslant 0$ the random variable $G(t)$ is lognormally distributed[2] *with moments and variance given by*

$$E[G(t)] = g_0 \exp(\mu t), \qquad E[(G(t))^2] = g_0^2 \exp\big((2\mu + \sigma^2)t\big),$$
$$\mathrm{Var}[G(t)] = g_0^2 e^{2\mu t}\Big(e^{\sigma^2 t} - 1\Big)\,.$$

The model of a Geometric Brownian motion can easily be generalised; let $(\Upsilon(t) : t \geqslant 0)$ and $(\Phi(t) : t \geqslant 0)$ be adapted stochastic processes which satisfy

$$\int_0^t |\Upsilon(s)|\, ds < \infty \quad \text{P-a.s.,} \qquad \int_0^t |\Phi(s)|^2\,, ds < \infty \quad \text{P-a.s.}$$

for all $t \geqslant 0$. Then we can define a stochastic process $(X(t) : t \geqslant 0)$ by

$$X(t) = \int_0^t \big(\Upsilon(s) - \tfrac{1}{2}\Phi^2(s)\big)\, ds + \int_0^t \Phi(s)\, dW(s) \qquad \text{for all t} \geqslant 0\,.$$

Let the stochastic process $(Y(t) : t \geqslant 0)$ be defined by $Y(t) := y_0 \exp(X(t))$ where $y_0 \in \mathbb{R}$ is a given constant. Since X is an Itô process we can apply Itô's formula in Theoerm 4.6.5 to derive

$$dY(t) = \Upsilon(t)Y(t)\, dt + \Phi(t)Y(t)\, dW(t) \qquad \text{for all t} \geqslant 0\,. \tag{5.3.10}$$

This equation differs from the stochastic differential equation (5.1.3) by the fact that the coefficients in (5.3.10) cannot be described by purely deterministic functions f and g, as for example in (5.3.8), due to the randomness of Υ and Φ. Nevertheless we can understand Y intuitively as a solution of a "stochastic differential equation with random coefficients" although we do not provide the corresponding theory[3], e.g a formal definition of a solution.

Equation (5.3.10) is much more general than (5.3.7) due to the coefficients depending on time t and on the random outcome ω. This yields that the solution Y is much less regular than the classical Geometric Brownian motion. In the next section we will use equation (5.3.10) to introduce a model of a share price much more general than the classical Black-Scholes model.

5.4. Application: Modelling the Share Prices

One of the reasons of the popularity of the Black-Scholes model is its simplicity and the fact that many terms can be calculated explicitly. However in order to be able to model

[2] A random variable X is called *lognormally distributed* if $\ln X$ is normally distributed.
[3] More advanced books often consider stochastic differential equation with random coefficients from the very beginning.

financial derivatives and markets it is important to understand which phenomena in the Black-Scholes model are intrinsic and which might be implied by a much broader view on the financial market.

One assumes that there is a risk-free asset $(B(t) : t \in [0,T])$ which is given by

$$\begin{aligned} dB(t) &= R(t)B(t)\,dt \qquad \text{for all t} \in [0, \mathrm{T}], \\ B(0) &= 1, \end{aligned} \tag{5.4.11}$$

where $R = (R(t) : t \in [0,T])$ is an adapted stochastic process. The process R might be interpreted as an instantaneous risk-free interest rate which is possibly random. If R is a deterministic constant, i.e. $R(t) = r$ for all $t \in [0,T]$, then we can interpret B as the price of a bond, say a savings account. If we fix $\omega \in \Omega$ equation (5.4.11) reduces to an ordinary differential equation of the form

$$\begin{aligned} \dot{B}_\omega(t) &= R_\omega(t)B_\omega(t) \qquad \text{for all t} \in [0, \mathrm{T}], \\ B_\omega(0) &= 1, \end{aligned}$$

where we use momentarily the notation $B_\omega(t) := B(t)(\omega)$ and $R_\omega(t) := R(t)(\omega)$. This differential equation can be solved explicitly and the solution is given by

$$B(t)(\omega) = \exp\left(\int_0^t R(s)(\omega)\,ds\right) \qquad \text{for all t} \in [0, \mathrm{T}] \text{ and all } \omega \in \Omega.$$

The share prices $(S(t) : t \in [0,T])$ is assumed to evolve according to the dynamic

$$dS(t) = \Upsilon(t)S(t)\,dt + \Phi(t)S(t)\,dW(t) \qquad \text{for all t} \in [0, \mathrm{T}], \tag{5.4.12}$$

where the *mean rate of return process* $(\Upsilon(t) : t \in [0,T])$ and the *volatility process* $(\Phi(t) : t \in [0,T])$ are adapted, measurable stochastic processes. The only constraint on these adapted, stochastic processes are that they are required to satisfy

$$\int_0^T |\Upsilon(s)|\,ds < \infty \qquad P\text{-a.s.}, \qquad \int_0^T |\Phi(s)|^2\,ds < \infty \qquad P\text{-a.s.}$$

Although equation (5.4.12) looks like a stochastic differential equation we do not call it so since the coefficients depend on $\omega \in \Omega$. However, the interpretation of the dynamic of the stochastic process S which satisfies the equation (5.4.12) is the same. In the end of Section 5.3 we derive its representation; the stochastic process $(S(t) : t \in [0,T])$ which satisfies equation (5.4.12) is given by

$$S(t) = S(0)\exp\left(\int_0^t \left(\Upsilon(s) - \tfrac{1}{2}\Phi^2(s)\right)\,ds + \int_0^t \Phi(s)\,dW(s)\right) \tag{5.4.13}$$

for all $t \in [0,T]$. Recall that we do not assume that the filtration is generated by the Brownian motion W. Thus, the condition on R, Υ and Φ to be adapted is very weak. For example, the interest rate R could be driven by another process.

Example 5.4.1. The (standard) Black-Scholes model is included in this setting. Just set $R(t) = r$, $\Upsilon(t) = \mu$ and $\Phi(t) = \sigma$ for all $t \in [0, T]$ and for some given constants $\mu \in \mathbb{R}$, $r \in \mathbb{R}_+$ and $\sigma > 0$.

The generalised Black-Scholes model captures a variety of models which can be found in the modern literature. A particular important case, which has gained much interest recently, are the so-called volatility models.

Example 5.4.2. Empirical studies suggest that the volatility in security time markets itself should be modelled by a stochastic processes, typically described by a stochastic differential equation. These are the so-called *stochastic volatility models* and they are included in the framework of this section. For example, consider the following model of the share prices S:

$$dS(t) = \mu S(t)\,dt + \Phi(t)S(t)\,dW_1(t) \qquad \text{for all t} \in \text{[0, T]}, \tag{5.4.14}$$
$$d\Phi(t) = f(t, \Phi(t))\,dt + g(t, \Phi(t))\,dW_2(t) \qquad \text{for all t} \in \text{[0, T]}, \tag{5.4.15}$$

where $\mu \in \mathbb{R}$ is a given constant and $f, g : [0, T] \times \mathbb{R} \to \mathbb{R}$ are given deterministic functions, $(W_1(t) : t \in [0, T])$ and $(W_2(t) : t \in [0, T])$ are Browian motions. In this situation, $(\Phi(t) : t \in [0, T])$ is an adapted stochastic process, defined as the solution of the stochastic differential equation (5.4.15), and the share price S is modelled by a generalised Black-Scholes model as described by equation (5.4.12).

Note, that since Φ is adapted to the filtration $\{\mathscr{F}_t^{W_2}\}_{t\geqslant 0}$ generated by W_2, the share price S is not necessarily adapted to the filtration $\{\mathscr{F}_t^{W_1}\}_{t\geqslant 0}$ generated by the Brownian motion W_1 which drives the equation of the share price. This fact has important consequences for these models as we will see later. The idea of volatility models goes back to Merton (1977).

We finish this short introduction to stochastic modelling of share prices by completing the model. Once the share prices is described we want also to model the trading of this share.

Definition 5.4.3.

(a) *A* trading strategy (Γ, Δ) *is an adapted stochastic process* $((\Gamma(t), \Delta(t)) : t \in [0, T])$ *with values in* $\mathbb{R}^2$.

(b) *the* value process $V = V(\Gamma, \Delta)$ *of the trading strategy* (Γ, Δ) *is given by*

$$V(t) = \Gamma(t)B(t) + \Delta(t)S(t) \qquad \textit{for all } \text{t} \in \text{[0, T]}\,.$$

(c) *a trading strategy* (Γ, Δ) *with*

$$\int_0^T |\Gamma(s)|\,ds < \infty, \qquad \int_0^T |\Delta(s)|^2\,ds < \infty \qquad P\text{-a.s.},$$

is called self-financing *if the value process* $V = V(\Gamma, \Delta)$ *satisfies*

$$dV(t) = \Gamma(t)\,dB(t) + \Delta(t)\,dS(t) \qquad \textit{for all } \text{t} \in \text{[0, T]}\,. \tag{5.4.16}$$

The interpretation of a trading strategy $((\Gamma(t), \Delta(t)) : t \in [0, T])$ is the following: at time t the agent who is trading according to this strategy keeps $\Gamma(t)$ units of the bond and $\Delta(t)$ units of the share. The bond has at this time the value $B(t)$ and the share price is written as $S(t)$. Negative values of $\Gamma(t)$ or $\Delta(t)$ correspond to short positions, respectively. This leads immediately to the value process $(V(t) : t \in [0, T])$ whose entry $V(t)$ represents the value of a portfolio at the time t if $\Gamma(t)$ units of the bond and $\Delta(t)$ units of the share are kept in the portfolio.

The interpretation of the definition of a self-financing portfolio is slightly more subtle. The self-financing condition (5.4.16) requires that every change in the value of the portfolio is only due to changes in the share price and in the bond.

5.5. Systems of stochastic differential equations

We also can consider stochastic differential equations in $\mathbb{R}^n$. For distinction to the one-dimensional situation we denote here[4] the Euclidean norm in $\mathbb{R}^n$ by $\|\cdot\|_n$, i.e. $\|a\|_n = (a_1^2 + \dots + a_n^2)^{1/2}$ for $a = (a_1, \dots, a_n) \in \mathbb{R}^n$; the Hilbert-Schmidt norm $\|\cdot\|_{\mathrm{HS}}$ of a matrix is defined in (4.7.20).

Let $(W(t) : t \geqslant 0)$ be a d-dimensional Brownian motion, i.e. $W(t) = (W_1(t), \dots, W_d(t))$ for all $t \geqslant 0$, and denote some measurable functions by

$$f\colon [0, T] \times \mathbb{R}^n \to \mathbb{R}^n \quad \text{and} \quad g\colon [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times d}.$$

Then a *system of stochastic differential equation* or a *multi-dimensional stochastic differential equation* is of the form

$$\begin{aligned} dX(t) &= f\big(t, X(t)\big)\, dt + g\big(t, X(t)\big)\, dW(t) \qquad \text{for all t} \in [0,\mathrm{T}], \\ X(0) &= X_0, \end{aligned} \tag{5.5.17}$$

where the initial condition X_0 is an $\mathscr{F}_0$-measurable random vector $X_0\colon \Omega \to \mathbb{R}^n$ with $E[\|X_0\|_n^2] < \infty$. We can write this stochastic differential equation in vector form, since the functions f and g are of the form

$$f(t,x) = \begin{pmatrix} f_1(t,x) \\ f_2(t,x) \\ \vdots \\ f_n(t,x) \end{pmatrix}, \qquad g(t,x) = \begin{pmatrix} g_{1,1}(t,x) & g_{1,2}(t,x) & \dots & \dots & g_{1,d}(t,x) \\ g_{2,1}(t,x) & g_{2,2}(t,x) & \dots & \dots & g_{2,d}(t,x) \\ \vdots & \vdots & & & \vdots \\ g_{n,1}(t,x) & g_{n,2}(t,x) & \dots & \dots & g_{n,d}(t,x) \end{pmatrix}$$

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

$$f_i\colon [0, T] \times \mathbb{R}^n \to \mathbb{R}, \qquad g_{i,j}\colon [0, T] \times \mathbb{R}^n \to \mathbb{R},$$

[4] this might be not consistent with other parts of the lecture notes where the Euclidean norm is just denoted by $|\cdot|$.

are some measurable functions for $i = 1, \dots, n$ and $j = 1, \dots, d$. Then the stochastic differential equation (5.5.17) can be written as

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ \vdots \\ dX_n(t) \end{pmatrix} = \begin{pmatrix} f_1\big(t, X(t)\big)\, dt \\ f_2\big(t, X(t)\big)\, dt \\ \vdots \\ f_n\big(t, X(t)\big)\, dt \end{pmatrix} + \begin{pmatrix} g_{1,1}\big(t, X(t)\big) & g_{1,2}\big(t, X(t)\big) & \dots & g_{1,d}\big(t, X(t)\big) \\ g_{2,1}\big(t, X(t)\big) & g_{2,2}\big(t, X(t)\big) & \dots & g_{2,d}\big(t, X(t)\big) \\ \vdots & \vdots & & \vdots \\ g_{n,1}\big(t, X(t)\big) & g_{n,2}\big(t, X(t)\big) & \dots & g_{n,d}\big(t, X(t)\big) \end{pmatrix} \begin{pmatrix} dW_1(t) \\ dW_2(t) \\ \vdots \\ dW_d(t) \end{pmatrix}$$

$$\begin{pmatrix} X_1(0) \\ X_2(0) \\ \vdots \\ X_n(0) \end{pmatrix} = \begin{pmatrix} X_0^{(1)} \\ X_0^{(2)} \\ \vdots \\ X_0^{(n)} \end{pmatrix}$$

for an initial condition $X_0 = (X_0^{(1)}, \dots, X_0^{(n)}) \colon \Omega \to \mathbb{R}^n$.

The definition of a solution of (5.5.17) is the same as in the one dimensional situation in Definition 5.1.1, only part (c) is adjusted in order to guarantee the existence of the multi-dimensional integrals in part (d).

Definition 5.5.1. *A stochastic process $X = (X(t) : t \in [0, T])$ in $\mathbb{R}^n$ is a solution of the stochastic differential equation (5.5.17) if*

(a) *X is adapted;*

(b) *$X(0) = X_0$ P-a.s.*

(c) $P\left(\int_0^T \|f(s, X(s))\|_n \, ds + \int_0^T \|g(s, X(s))\|_{\mathrm{HS}}^2 \, ds < \infty \right) = 1.$

(d) *For all $t \in [0, T]$, the process X satisfies*

$$X(t) = X(0) + \int_0^t f(s, X(s))\, ds + \int_0^t g(s, X(s))\, dW(s) \qquad P\text{-a.s.}$$

Theorem 5.5.2. *Suppose that there exists a constant $c > 0$ such that the following are satisfied for all $t \in [0, T]$ and all $x, y \in \mathbb{R}^n$:*

$$\|f(t, x) - f(t, y)\|_n + \|g(t, x) - g(t, y)\|_{\mathrm{HS}} \leqslant c \, \|x - y\|_n \,, \tag{5.5.18}$$

$$\|f(t, x)\|_n^2 + \|g(t, x)\|_{\mathrm{HS}}^2 \leqslant c^2 (1 + \|x\|_n^2). \tag{5.5.19}$$

Then there exists a unique solution $(X(t) : t \in [0, T])$ of the stochastic differential equation (5.1.3). Moreover, the solution X has the following properties:

(i) *X has continuous trajectories;*

(ii) *there exists a constants $\alpha \geqslant 0$ such that*

$$E\left[\sup_{t \in [0,T]} \|X(t)\|_n^2 \right] \leqslant \alpha \left(1 + E\big[\, \|X_0\|_n^2 \,\big] \right).$$

Example 5.5.3. The result from part (a) in Example 4.6.9 can be summarised as

$$\begin{aligned} dY(t) &= -\tfrac{1}{2}Y(t)\,dt - Z(t)\,dW(t) \qquad \text{for all t} \in [0,\text{T}], \\ dZ(t) &= -\tfrac{1}{2}Z(t)\,dt + Y(t)\,dW(t) \qquad \text{for all t} \in [0,\text{T}], \end{aligned}$$

where $Y(t) = \cos(W(t))$ and $Z(t) = \sin(W(t))$. Thus, the 2-dimensional stochastic process $((Y(t), Z(t)) : t \in [0,T])$ is a solution of the stochastic differential equation in $\mathbb{R}^2$:

$$\begin{aligned} dX(t) &= -\frac{1}{2}X(t)\,dt + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} X(t)\,dW(t) \qquad \text{for all t} \in [0,\text{T}], \\ X(0) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Here W denotes a 1-dimensional Brownian motion, i.e. $d = 1$ and $n = 2$ in (5.5.17). The coefficients here are very simple functions, i.e.

$$f\colon [0,T] \times \mathbb{R}^2 \to \mathbb{R}^2, \qquad f\big(t,(x_1,x_2)\big) = \begin{pmatrix} -\frac{1}{2}x_1 \\ -\frac{1}{2}x_2 \end{pmatrix},$$

$$g\colon [0,T] \times \mathbb{R}^2 \to \mathbb{R}^2, \qquad g\big(t,(x_1,x_2)\big) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

5.6. Numerical approximation

In most cases one can not explicitly give the solution of a stochastic differential equation although its existence is guaranteed. For this reason it is important to have numerical methods to simulate and to approximate the solutions. Since our target is a stochastic process or at least a random variable there are different possibilities which object we actually simulate: the paths of a solution, the probability distribution of the solution at a specific time or the expectation of the solution at a specific time or something else. In this section we only introduce shortly a method to simulate the paths of a solution. There are much more sophisticated methods for this aim but the ideas are similar.

The considered numerical scheme is called the *Euler-Maruyama method* and it is very close to the classical Euler scheme which is used to approximate the solutions of ordinary differential equations. Divide the time interval $[0,T]$ into subinterval of length T/h by the partition $\{t_k\}_{k=0,\dots,h}$ with $t_k = kT/h$. If $(X(t) : t \in [0,T])$ denotes a solution of (5.1.3) then it satisfies

$$X(t_k) = X(t_{k-1}) + \int_{t_{k-1}}^{t_k} f(s,X(s))\,ds + \int_{t_{k-1}}^{t_k} g(s,X(s))\,dW(s) \qquad \text{for all } k = 0,\dots,h.$$

Since for each $k = 0,\dots,h$ the right hand side depends not only on the value of $X(t_{k-1})$ but also on the values $(X(s) : s \in [t_{k-1},t_k])$ we can not use this representation to approximate $X(t_k)$. However, if the length T/h of the interval $[t_{k-1},t_k]$ is small then, if we require f and g to be continuous in both arguments, one can expect that $f(s,X(s))$ for $s \in [t_{k-1},t_k]$

does not differ too much from $f(t_{k-1}, X(t_{k-1}))$ since f and the solution X is continuous and analogously for g. This results in the approximating scheme $(Y_{t_k}^{(h)} : k = 0, \dots, h)$ defined recursively by $Y_0^{(h)} = x_0$ and

$$Y_{t_k}^{(h)} := Y_{t_{k-1}}^{(h)} + f(t_{k-1}, Y_{t_{k-1}}^{(h)})(t_k - t_{k-1}) + g(t_{k-1}, Y_{t_{k-1}}^{(h)})(W(t_k) - W(t_{k-1})) \tag{5.6.20}$$

for all $k = 1, \dots, h$. In each step the only unknown term is $W(t_k) - W(t_{k-1})$ which can be easily simulated since it is a normally distributed random variable with expectation 0 and variance T/h.

The scheme $(Y_{t_k}^{(h)} : k = 0, \dots, h)$ will not coincide with the true solution $(X(t) : t \in [0, T])$ but we can expect that they are close to each other. The distance between the approximating scheme and the true solution can be measured in different ways. In the following result we measure this error by the norm in $L^2_P(\Omega$.

Theorem 5.6.1. *Let X be the solution of the stochastic differential equation (5.1.3) with the initial condition $X(0) = x_0$, where the coefficients f, $g\colon [0, T] \times \mathbb{R} \to \mathbb{R}$ satisfy the Lipschitz (5.1.4) and linear growth condition (5.1.5). If $T > 0$ is fixed and $(Y_{t_k}^{(h)} : k = 0, \dots, h)$ is defined by (5.6.20) for $h > 0$ then*

$$\max_{k=0,\dots,h} E\left[\left|X(t_k) - Y_{t_k}^{(h)}\right|^2\right] \leqslant \alpha h^{-1/2}$$

for a constant $\alpha > 0$ which depends on T, x_0, f and g.

The exponent $\gamma = 1/2$ on the right hand side in the result of Theorem 5.6.1 is the so-called *order of the Euler-Maruyama scheme*. It is important since if it is required to decrease the error 10 times one has to make the interval length 100 times smaller. There are other schemes, e.g Milstein, Runge-Kutta, to approximate the paths of the solutions with higher order but they require more efforts to compute.

5.7. Exercises

Let $(W(t) : t \geqslant 0)$ be a Brownian motion on a probability space $(\Omega, \mathscr{A}, P)$ in the following exercises.

1. Let $(S(t) : t \in [0, T])$ be the solution of

$$dS(t) = \alpha S(t)\, dt + \sigma S(t)\, dW(t)$$

for some constants $\alpha \in \mathbb{R}$ and $\sigma^2 > 0$. Find the stochastic differential equation satisfied by $(Z(t) : t \in [0, T])$ where

$$Z(t) = S^m(t) \qquad \text{for some } m \in \mathbb{N}\,.$$

2. (a) Define for a constant $a > 0$ the stochastic process $(X(t) : t \geqslant 0)$ by

$$X(t) = \begin{cases} (W(t) - a)^3, & \text{if } W(t) \geqslant a, \\ 0, & \text{else.} \end{cases}$$

Show that X solves the stochastic differential equation

$$\begin{aligned} dX(t) &= 3X^{1/3}(t)\, dt + 3X^{2/3}(t)\, dW(t) \qquad \text{for all } t \geqslant 0, \\ X(0) &= 0. \end{aligned} \tag{5.7.21}$$

(b) Conclude from (a) that there exists infinitely many solutions of the given stochastic differential equation and indicate which condition in Theorem 5.1.5 is not satisfied.

3. Assume that $X := (X(t) : t \in [0, T])$ is the unique solution of

$$\begin{aligned} dX(t) &= b\big(a - \ln(X(t))\big)X(t)\, dt + \sigma X(t)\, dW(t) \qquad \text{for all } t \in [0, T], \\ X(0) &= x_0, \end{aligned}$$

for some constants $a, b, \sigma, x_0 > 0$.

(a) Let $Y(t) := \ln(X(t))$ and derive the stochastic differential equation satisfied by $(Y(t) : t \in [0, T])$.

(b) Derive an explicit representation of $(\exp(bt)Y(t) : t \in [0, T])$ in terms of a, b, σ^2, x_0 and the Brownian motion W.

(c) Use (a) and (b) to derive an explicit representation of the stochastic process X in terms of a, b, σ^2, x_0 and the Brownian motion W.

(d) Since the representation of X in (c) is derived under the assumption that X is a solution it remains to show that X is in fact is a solution of the stochastic differential equation above.

4. Let $X := (X(t) : t \in [0, T])$ be the solution of

$$\begin{aligned} dX(t) &= \alpha X(t)\, dt + \sigma X(t)\, dW(t) \qquad \text{for all } t \in [0, T], \\ X(0) &= x_0, \end{aligned}$$

where $\alpha, \sigma \in \mathbb{R}$ and $x_0 > 0$. Derive the stochastic differential equation which is satisfied by $(X^{-1}(t) : t \in [0, T])$.

5. Let X and Y be real-valued Itô processes of the form

$$\begin{aligned} X(t) &= X(0) + \int_0^t \Upsilon_1(s)\, ds + \int_0^t \Phi_1(s)\, dW(s) \qquad \text{for all t} \in [0, \text{T}], \\ Y(t) &= Y(0) + \int_0^t \Upsilon_2(s)\, ds + \int_0^t \Phi_2(s)\, dW(s) \qquad \text{for all t} \in [0, \text{T}], \end{aligned}$$

where $(W(t) : t \geqslant 0)$ is a one-dimensional Brownian motion. Show that the product of X and Y obeys

$$d(X(t)Y(t)) = Y(t)\,dX(t) + X(t)\,dY(t) + \Phi_1(t)\Phi_2(t)\,dt.$$

Compare with Example 4.7.4.

6. Let $(Z(t) : t \in [0,T])$ be the solution of the stochastic differential equation in $\mathbb{R}^2$:

$$\begin{aligned} dZ(t) &= \alpha Z(t)\,dt + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} Z(t)\,dW(t) \qquad \text{for all t} \in [0, \text{T}], \\ Z(0) &= \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \end{aligned}$$

where $\alpha \in \mathbb{R}$ and $x_0,\ y_0 \in \mathbb{R}$ and W is a Brownian motion in $\mathbb{R}$. Denote $Z(t) = (X(t), Y(t))$ for all $t \in [0,T]$.

   (a) Show that there exists a unique solution of this stochastic differential equation.

   (b) Show that $t \mapsto X^2(t) + Y^2(t)$ is a deterministic function.

   (c) Compute $E[X(t)]$, $E[Y(t)]$ and $\operatorname{Cov}(X(t), Y(t))$.

7. Let $((W_1(t), W_2(t)) : t \geqslant 0)$ be a Brownian motion in $\mathbb{R}^2$ and define for $i = 1, 2$ the stochastic processes $(X_i(t) : t \in [0,T])$ by

$$X_i(t) := \exp\left(-\frac{1}{2}\int_0^t \alpha(u)\,du\right)\left(x_i + \int_0^t \left(\tfrac{1}{2}\sigma(s)\exp\left(\tfrac{1}{2}\int_0^s \alpha(u)\,du\right)\right)dW_i(s)\right),$$

where $\alpha, \sigma : [0,T] \to \mathbb{R}$ are continuous functions and $x_1, x_2 > 0$.

   (a) Show that X_i is the solution of the stochastic differential equation

$$\begin{aligned} dX_i(t) &= -\tfrac{1}{2}\alpha(t)X_i(t)\,dt + \frac{1}{2}\sigma(t)\,dW_i(t) \qquad \text{for all } t \in [0,T], \\ X_i(0) &= x_i. \end{aligned}$$

   (b) Determine the probability distribution of $X_i(t)$ and the values of $E[X_i(t)]$ and $\operatorname{Var}[X_i(t)]$ and $\operatorname{Cov}(X_i(s), X_i(t))$ for $s, t \in [0,T]$ and $i = 1, 2$.

   (c) Show that the stochastic process $(R(t) : t \in [0,T])$ defined by $R(t) := X_1^2(t) + X_2^2(t)$ obeys

$$dR(t) = \left(\tfrac{1}{2}\sigma^2(t) - \alpha(t)R(t)\right)dt + \sigma(t)X_1(t)\,dW_1(t) + \sigma(t)X_2(t)\,dW_2(t).$$

A well known result, Lévy's Characterisation of Brownian motion, shows immediately that

$$B(t) := \sum_{i=1}^{2}\int_0^t \frac{X_i(s)}{\sqrt{R(s)}}\,dW_i(s) \qquad \text{for all t} \in [0, \text{T}],$$

defines a Brownian motion $(B(t) : t \geqslant 0)$. Consequently, the equation in part (c) can be written as

$$dR(t) = \Big(\tfrac{1}{2}\sigma^2(t) - \alpha(t)R(t)\Big)\,dt + \sigma(t)\sqrt{R(s)}\;dB(s).$$

As a very special case, i.e. if $\alpha(t) = a$ and $\sigma(t) = b$ for some constants $a, b > 0$ we have

$$dR(t) = -a\Big(R(t) - \tfrac{b^2}{2a}\Big)\,dt + b\sqrt{R(s)}\;dB(s),$$

which is of form as in the Cox, Ingersoll and Ross interest rate model, see Financial Mathematics 7. Note, that the Brownian motion B is defined in terms of the given W_1 and W_2 and it is not an arbitrary Brownian motion, as we usually assume.

8. Find a stochastic process $(R(t) : t \in [0, T])$ which satisfies the Vasicek model

$$\begin{aligned} dR(t) &= -a(R(t) - b)\,dt + \sigma\,dW(t) \qquad \text{for all } t \in [0, T],\\ R(0) &= r_0, \end{aligned}$$

where $r_0, a, b, \sigma > 0$ are constants.

6 Girsanov's Theorem

In this section we introduce one of the fundamental results in stochastic calculus and which does not have an analogue in classical calculus. The result is about changing the measure P of the underlying probability space $(\Omega, \mathscr{A}, P)$, which might look rather awkward if one is only used to elementary probability theory where the underlying probability space is fixed. However, in many models one can interpret the probability measure P as an individual perspective on the underlying dynamics, e.g. it models the risk aversion of an agent in financial mathematics. But also in other areas of mathematics there is sometimes a need to change the underlying probability measure, e.g. simulation of rare events, statistics (Maximum Likelihood) and stochastic control theory.

We assume the same properties of the underlying probability space $(\Omega, \mathscr{A}, P)$, of the filtration $\{\mathscr{F}_t^W\}_{t\geqslant 0}$ and of the Brownian motion $(W(t) : t \geqslant 0)$, as in Chapter 4.

6.1. Girsanov's Theorem

Example 6.1.1. Let the probability space $(\Omega, \mathscr{A}, P)$ be given by $\Omega := \{\omega_1, \omega_2\}$, $\mathscr{A} := \mathscr{P}(\Omega)$ and $P(\{\omega_1\}) := p$ and $P(\{\omega_2\}) := 1 - p$ for a constant $p \in (0, 1)$. For a given constant $q \in (0, 1)$ we define a new random variable by

$$\Lambda\colon \Omega \to \mathbb{R}, \qquad \Lambda(\omega) := \begin{cases} \frac{q}{p}, & \text{if } \omega = \omega_1, \\ \frac{1-q}{1-p}, & \text{if } \omega = \omega_2. \end{cases}$$

It follows that

$$Q : \mathscr{A} \to [0, 1], \qquad Q(A) := \sum_{\omega \in A} \Lambda(\omega) P(\{\omega\}) \quad \text{for } \omega \in \Omega,$$

defines a new probability measure. Equivalently we can write

$$Q(A) = E_P[\mathbb{1}_A \Lambda] \qquad \text{for all } A \in \mathscr{A}.$$

where E_P denotes expectation under the probability measure P. The distribution of the random variable X under P is given by

$$P(X \in B) = p\mathbb{1}_B(X(\omega_1)) + (1-p)\mathbb{1}_B(X(\omega_2)) \quad \text{for all } B \in \mathfrak{B}(\mathbb{R}).$$

But the distribution of the random variable X under the new probability measure Q obeys

$$\begin{aligned} Q(X \in B) &= \Lambda(\omega_1)P(\{\omega_1\})\mathbb{1}_B(X(\omega_1)) + \Lambda(\omega_2)P(\{\omega_2\})\mathbb{1}_B(X(\omega_2)) \\ &= q\mathbb{1}_B(X(\omega_1)) + (1-q)\mathbb{1}_B(X(\omega_2)) \end{aligned}$$

for all $B \in \mathfrak{B}(\mathbb{R})$. The expectation of X under P, denoted by E_P, is given by

$$E_P[X] = X(\omega_1)P(\{\omega_1\}) + X(\omega_2)P(\{\omega_2\}) = pX(\omega_1) + (1-p)X(\omega_2),$$

and the expectation of X under Q, denoted by E_Q, is given by

$$\begin{aligned} E_Q[X] &= X(\omega_1)Q(\{\omega_1\}) + X(\omega_2)Q(\{\omega_2\}) \\ &= X(\omega_1)\Lambda(\omega_1)P(\{\omega_1\}) + X(\omega_2)\Lambda(\omega_2)X(\omega_2) = E_P[X\Lambda]. \end{aligned}$$

Example 6.1.2. Let X be a random variable, defined on the probability space $(\Omega, \mathscr{A}, P)$, which is normally distributed with expectation 0 and variance 1. Let f_μ denote the density of the $N(\mu, 1)$ distribution for $\mu \in \mathbb{R}$, that is

$$f_\mu \colon \mathbb{R} \to \mathbb{R}_+ \qquad f_\mu(u) := \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(u-\mu)^2}.$$

Since X is standard normally distributed its distribution under P is given by

$$P\big(X \in B\big) = \int_B f_0(u)\, du \qquad \text{for all } B \in \mathfrak{B}(\mathbb{R}).$$

For a constant $a \in \mathbb{R}$ we define the random variable $Y := X + a$. Clearly, the distribution of Y under the original measure P obeys

$$P\big(Y \in B\big) = P\big(X \in B - a\big) = \int_{B-a} f_0(u)\, du = \int_B f_a(u)\, du \qquad \text{for all } B \in \mathfrak{B}(\mathbb{R}).$$

Can we find a probability measure Q on $\mathscr{A}$ such that Y is normally distributed with expectation 0 and variance 1 under Q? In order to answer this question we define a new random variable by

$$\Lambda \colon \Omega \to \mathbb{R} \qquad \Lambda(\omega) := e^{-aX(\omega) - \frac{a^2}{2}},$$

and a new probability measure by

$$Q : \mathscr{A} \to [0,1], \qquad Q(A) := E_P[\mathbb{1}_A \Lambda].$$

It follows that the distribution of Y under Q is given by

$$Q\big(Y \in B\big) = Q\big(Y^{-1}(B)\big) = E_P\big[\mathbb{1}_{Y^{-1}(B)}\Lambda\big] = E_P\big[\mathbb{1}_{X^{-1}(B-a)}\Lambda\big],$$

where we use the equality $Y^{-1}(B) = X^{-1}(B-a)$. Due to the set equality

$$\{\omega \in \Omega : \omega \in X^{-1}(B-a)\} = \{\omega \in \Omega : X(\omega) \in B - a\},$$

and since $\Lambda = f(X)$ for the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = e^{-ax-\frac{a^2}{2}}$, we obtain

$$E_P\big[\mathbb{1}_{X^{-1}(B-a)}\Lambda\big] = E_P\big[\mathbb{1}_{(B-a)}(X)\, f(X)\big] = \frac{1}{\sqrt{2\pi}} \int_{B-a} e^{-au-\frac{a^2}{2}} e^{-\frac{1}{2}u^2}\, du.$$

By summarising the terms in the last integral we arrive at

$$Q(Y \in B) = \frac{1}{\sqrt{2\pi}} \int_{B-a} e^{-\frac{1}{2}(u+a)^2}\, du = \frac{1}{\sqrt{2\pi}} \int_{B} e^{-\frac{1}{2}u^2}\, du,$$

which shows that Y is normally distributed with expectation 0 and variance 1 under Q.

Theorem 6.1.3. *(Girsanov's Theorem)*
Let $(X(t) : t \in [0,T])$ be an adapted stochastic process satisfying

$$P\left(\int_0^T X^2(s)\, ds < \infty\right) = 1. \tag{6.1.1}$$

Define a stochastic process $(L(t) : t \in [0,T])$ by

$$L(t) := \exp\left(-\int_0^t X(s)\, dW(s) - \tfrac{1}{2}\int_0^t X^2(s)\, ds\right),$$

and a mapping Q by

$$Q\colon \mathscr{A} \to [0,\infty], \qquad Q(A) := E_P[\mathbb{1}_A L(T)].$$

If $(L(t) : t \in [0,T])$ is a martingale under P then the mapping Q is a probability measure and the stochastic process $(\widetilde{\mathrm{W}}(t) : t \in [0,T])$ defined by

$$\widetilde{\mathrm{W}}(t) := W(t) + \int_0^t X(s)\, ds,$$

is a Brownian motion under the new measure Q.

Notation: The stochastic process $(L(t) : t \in [0,T])$ is called *Radon-Nikodym derivative* and is denoted by

$$\frac{dP}{dQ}|_{\mathscr{F}_t} := L(t).$$

If you know some measure theory, then the probability measure Q can equivalently be written as

$$Q(A) = \int_A L(T)(\omega)\, P(d\omega) \qquad \text{for all } A \in \mathscr{A}.$$

Compare with Example (6.1.1), where both representations are mentioned and are understandable without knowing measure theory.

Remark 6.1.4.

(a) The condition (6.1.1) guarantees that both integrals in the definition of $L(t)$ exist. The main condition and often most difficult to verify is the requirement that $(L(t) : t \in [0,T])$ is a martingale under P.

(b) Girsanov's Theorem is also true under the condition that $E_P[L(T)] = 1$. Obviously, this is implied by the requirement that $(L(t) : t \in [0,T])$ is a martingale, and in fact both conditions are equivalent, see Example 6.1.9.

(c) Since by definition of L we have $P(L(t) > 0) = 1$ it follows from the definition of Q by $Q(A) = E_P[\mathbb{1}_A L(T)]$ that

$$Q(A) = 0 \;\Leftrightarrow\; P(A) = 0.$$

In other words, an event has the probability 0 under P if and only if it has probability 0 under Q and the same applies for events of probability 1. Thus, the measures P and Q differ in the probability values they assign to events but no event is neglected by becoming an event of probability 0. This is called that the measures P and Q are *equivalent*.

Lemma 6.1.5. *Assume the conditions in Theorem 6.1.3. Then a random variable Y satisfies $E_Q[|Y|] < \infty$ if and only if $E_P[|YL(T)|] < \infty$. In this situation, we have*

$$E_Q\big[Y\big] = E_P\big[YL(T)\big].$$

Proof. If the sample space Ω countable the proof is easy. □

Example 6.1.6. Let $(Y(t) : t \in [0,T])$ be given by

$$Y(t) := \mu t + \sigma W(t)$$

for some constants $\mu \in \mathbb{R}$ and $\sigma > 0$ and with $(W(t) : t \in [0,T])$ denoting a Brownian motion under P. Is there a measure such that Y is a martingale? If the stochastic process $(X(t) : t \in [0,T])$ is given by $X(t) = \mu/\sigma$, then the random variables

$$L(t) := \exp\left(-\int_0^t X(s)\,dW(s) - \tfrac{1}{2}\int_0^t X^2(s)\,ds\right) = \exp\Big(-\frac{\mu}{\sigma}W(t) - \frac{\mu^2 t}{2\sigma^2}\Big),$$

define a martingale $(L(t) : t \in [0,T])$ according to part (b) in Corollary 3.2.4. Thus, Girsanov's Theorem implies that

$$\widetilde{\mathrm{W}}(t) := W(t) + \frac{\mu t}{\sigma}$$

defines a Brownian motion $(\widetilde{\mathrm{W}}(t) : t \in [0,T])$ under the probability measure

$$Q : \mathscr{A} \to [0,1], \qquad Q(A) = E_P\left[\mathbb{1}_A \exp\left(-\tfrac{\mu}{\sigma}W(T) - \tfrac{\mu^2 T}{2\sigma^2}\right)\right].$$

One can calculate for example

$$\begin{aligned} E_P[Y^2(t)] &= E_P\left[\mu^2 t^2 + 2\mu t\sigma W(t) + \sigma^2 W^2(t)\right] = \mu^2 t^2 + \sigma^2 t, \\ E_Q[Y^2(t)] &= E_Q\left[(\sigma^2 \tilde{W}(t))^2\right] = \sigma^2 t. \end{aligned}$$

Financial Mathematics 8. A barrier option is an option whose pay off at maturity T depends whether or not the underlying asset $(S(t) : t \in [0,T])$ hits a specified barrier during the lifetime of the option. For example, an *up-and-out call option* C is of the form

$$C = \begin{cases} (S(T) - K)^+, & \text{if } \sup_{t\in[0,T]} S(t) < B, \\ 0, & \text{else.} \end{cases},$$

where $K > 0$ is the strike price, $B > 0$ the barrier. It is assumed $K < B$. In order to calculate its value process it is essential to know the distribution of the *running maximum of a Brownian motion* W, that is

$$M(t) := \sup_{s\in[0,t]} W(s)$$

The distribution of $M(t)$ can be calculated explicitly by using Girsanov's Theorem and the reflection principle for Brownian motion.

Example 6.1.7. (Importance sampling - rare events)
Let X be a normal distributed random variable with expectation 0 and variance 1. A rare event refers to an event with a very small probability, e.g. $\{X > a\}$ for a large constant $a > 0$. The standard approach (Monte Carlo) is to generate a sample $\{x_1, \dots, x_N\}$ of N independent standard normally distributed random variables. Then the law of large numbers guarantees that

$$P(X > a) \approx \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_{\{x_k > a\}}, \tag{6.1.2}$$

where $\approx$ is to understand that the larger N the better the approximation. However, for large a the probability of the event $\{X > a\}$ is extremely small and a good approximation requires a very large sample size N.

For importance sampling one changes the underlying measure P such that the random variable X has expectation a under the new probability measure Q. For that purpose, define a random variable

$$\Lambda\colon \Omega \to \mathbb{R} \qquad \Lambda(\omega) := e^{-aX(\omega) - \frac{a^2}{2}}$$

and a new probability measure by

$$Q\colon \mathscr{A} \to [0,1], \qquad Q(A) := E_P[\mathbb{1}_A \Lambda].$$

With $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \exp(-ax - \frac{a^2}{2})$ it follows for every $A \in \mathfrak{B}(\mathbb{R})$

$$Q(X \in A) = E_P[\mathbb{1}_A(X) f(X)] = \frac{1}{\sqrt{2\pi}} \int_A e^{-au - \frac{a^2}{2}} e^{-\frac{1}{2}u^2}\, du = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{1}{2}(u+a)^2}\, du,$$

which shows that X is normally distributed with expectation a and variance 1 under Q, confer with Example 6.1.2. For each $A \in \mathscr{A}$ we obtain

$$P(A) = E_P[\mathbb{1}_A] = E_P\left[\mathbb{1}_A f(X) \frac{1}{f(X)}\right] = E_P\left[\mathbb{1}_A \Lambda e^{aX + \frac{a^2}{2}}\right] = e^{\frac{a^2}{2}} E_P\left[\mathbb{1}_A \Lambda e^{aX}\right]$$

Taking into account $E_Q[Y] = E_P[\Lambda Y]$ for every random variable Y we arrive at

$$P(A) = e^{\frac{a^2}{2}} E_Q\left[\mathbb{1}_A e^{aX}\right] \qquad \text{for all } A \in \mathfrak{B}(\mathbb{R}). \tag{6.1.3}$$

Choose the set $A := \{X > a\}$. Since X is normally distributed with expectation a and variance 1 under Q, the strong law of large number (applied under Q) guarantees that a sample $\{z_1, \dots, z_N\}$ of independent normally distributed random variables with expectation a and variance 1 satisfies

$$E_Q\left[\mathbb{1}_{\{X>a\}} e^{aX}\right] \approx \frac{1}{N} \sum_{k=1}^N \mathbb{1}_{\{z_k>a\}} e^{a z_k}.$$

Equality (6.1.3) yields

$$P(X > a) \approx e^{\frac{a^2}{2}} \frac{1}{N} \sum_{k=1}^N \mathbb{1}_{\{z_k>a\}} e^{a z_k},$$

which is a much better approximation than (6.1.2), since we can expect that about half of the sample exceeds the value a.

Admittedly this example shows the benefit of changing the underlying measure but it does not need Girsanov's Theorem, since there is no dynamic in time. However one can extend this method to sampling of solutions of stochastic differential equations where Girsanov's theorem plays an essential role.

In applying Girsanov's Theorem 6.1.3 the most difficult part is to verify that $(L(t) : t \in [0,T])$ is a martingale. A powerful tool for that is the following sufficient condition:

Theorem 6.1.8. *(Novikov's condition)*
Let $(X(t) : t \in [0,T])$ be an adapted stochastic process satisfying

$$P\left(\int_0^T X^2(s)\, ds < \infty\right) = 1.$$

Define for $t \in [0,T]$ the random variable

$$L(t) := \exp\left(-\int_0^t X(s)\, dW(s) - \tfrac{1}{2}\int_0^t X^2(s)\, ds\right).$$

If

$$E\left[\exp\left(\frac{1}{2}\int_0^T X^2(s)\, ds\right)\right] < \infty$$

then the process $(L(t) : t \in [0,T])$ is a martingale under P.

Stochastic processes L of the form as considered in the above Theorem 6.1.8 have already appeared in Example 4.6.8; we continue to consider them in the following example:

Example 6.1.9. Let $(Z(t) : t \in [0,T])$ be an Itô processes of the form

$$Z(t) = \int_0^t \underbrace{-X(s)}_{=:\Phi(s)}\, dW(s) + \int_0^t \underbrace{-\frac{1}{2}X^2(s)}_{=:\Upsilon(s)}\, ds \qquad \text{for all } t \in [0,T],$$

where $(X(t) : t \in [0,T])$ is an adapted stochastic process satisfying

$$\int_0^T |X(s)|^2 \, ds < \infty \quad P\text{-a.s.}$$

Define the function

$$f : [0,T] \times \mathbb{R} \to \mathbb{R}, \qquad f(t,z) = \exp(z).$$

Itô's formula in Theorem 4.6.5 implies

$$f(t, Z(t)) = 1 - \int_0^t f(s, Z(s))\, \Phi(s)\, dW(s) \qquad \text{for all } t \in [0,T].$$

In general we only know that $(f(t,Z(t))\Phi(t) : t \in [0,T]) \in \mathscr{H}_{\mathrm{loc}}$ and thus, we only have that

$$\left(L(t) := \exp\left(- \int_0^t X(s)\, dW(s) - \tfrac{1}{2} \int_0^t X^2(s)\, ds \right) : t \in [0,T] \right)$$

is a local martingale according to theorem 4.4.9.

Let $\{\tau_n\}_{n\in\mathbb{N}}$ be a localising sequence for the local martingale $(L(t) : t \in [0,T])$. Then $(L(t \wedge \tau_n) : t \in [0,T])$ is a martingale and we obtain

$$1 = E[L(0)] = E[L(t\wedge\tau_n)] \qquad \text{for all } t \in [0,T],\ n \in \mathbb{N}.$$

Since $L(t) \geqslant 0$ for all $t \in [0,T]$, Fatou's Lemma[1] implies for all $t \in [0,T]$

$$E[L(t)] = E\left[\lim_{n\to\infty} L(t\wedge\tau_n) \right] \leqslant \lim_{n\to\infty} E\left[L(t\wedge\tau_n)\right] = 1,$$

which shows $E[|L(t)|] < \infty$ for all $t \geqslant 0$. In the same way one can establish for each $0 \leqslant s \leqslant t$

$$L(s) = \lim_{n\to\infty} L(s\wedge\tau_n) = \lim_{n\to\infty} E[L(t\wedge\tau_n)|\mathscr{F}_s] \geqslant E[\lim_{n\to\infty} L(t\wedge\tau_n)|\mathscr{F}_s] = E[L(t)|\mathscr{F}_s],$$

which shows that $(L(t) : t \in [0,T])$ is a supermartingale. In particular, since $E[L(t)] \leqslant E[L(s)] \leqslant 1$ for all $0 \leqslant s \leqslant t$ it follows that L is a martingale if and only if $E[L(t)] = 1$ for all $t \in [0,T]$.

Example 6.1.10. Girsanov's theorem can be used to transform the drift in a stochastic differential equation as in (5.1.3). Let $(X(t) : t \in [0,T])$ be the unique solution of

$$dX(t) = f(t, X(t))\, dt + g(t, X(t))\, dW(t) \qquad \text{for all } t \in [0,T], \tag{6.1.4}$$

where the coefficients are given by some measurable functions $f, g \colon [0,T] \times \mathbb{R} \to \mathbb{R}$. Let $\varphi : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a continuous function and assume that

$$E\left[\exp\left(\tfrac{1}{2} \int_0^T |\varphi(s, X(s))|^2 \, ds \right) \right] < \infty.$$

[1] If you do not know Fatou's Lemma consider it as a justification for changing the order of expectation and limit below, similarly as Lebesgue's Theorem of dominated convergence but with $\leqslant$.

Define a stochastic process $(L(t) : t \in [0,T])$ by

$$L(t) := \exp\left(-\int_0^t \varphi\big(s, X(s)\big)\, dW(s) - \tfrac{1}{2}\int_0^t \varphi\big(s, X(s)\big)^2\, ds\right).$$

Theorem 6.1.8 and Theorem 6.1.3 imply that

$$Q\colon \mathscr{A} \to [0,\infty], \qquad Q(A) := E_P[\mathbb{1}_A L(T)],$$

defines a probability measure and that the stochastic process $(\widetilde{W}(t) : t \in [0,T])$ defined by

$$\widetilde{W}(t) := W(t) + \int_0^t \varphi\big(s, X(s)\big)\, ds,$$

is a Brownian motion under the probability measure Q. The solution X of (6.1.4) obeys then

$$\begin{aligned}
dX(t) &= f(t, X(t))\, dt + g(t, X(t))\, dW(t)\\
&= f(t, X(t))\, dt + g(t, X(t))\,\big(dW(t) + \varphi(t, X(t)\, dt)\big) - g(t, X(t))\varphi(t, X(t))\, dt\\
&= \big(f(t, X(t)) - g(t, X(t))\varphi(t, X(t))\big)\, dt + g(t, X(t))\, d\widetilde{W}(t).
\end{aligned}$$

Thus, X is also the solution of the stochastic differential equation

$$dX(t) = \big(f(t, X(t)) - g(t, X(t))\varphi(t, X(t))\big)\, dt + g(t, X(t))\, d\widetilde{W}(t).$$

For example, if we can choose

$$\varphi(t,x) = \frac{f(t,x)}{g(t,x)} \qquad \text{for all } t \in [0,T],\ x \in \mathbb{R},$$

and the function φ satisfies the conditions assumed above we arrive at

$$dX(t) = g(t, X(t))\, d\widetilde{W}(t). \tag{6.1.5}$$

Many properties of the solution X which are true under the probability measure Q are also true under the probability measure P. Since (6.1.5) is often easier to analyse one can conclude in this way certain properties of the solution of the stochastic differential equation (6.1.4).

6.2. Financial mathematics: arbitrage-free models

Financial mathematics relies on the assumption of arbitrage-free markets, that is there does not exist an opportunity to make money without any risk. One can formally describe an arbitrage opportunity in terms of self-financing strategies as introduced in Definition 5.4.3. However, in continuous time this formulation is a bit tricky and we can circumvent this difficulty by motivating the following result based on our understanding that a martingale represents a fair game, i.e. the likelihood to win or to loose is equal. Thus, if the considered market is arbitrage-free then there should be at least one market such that the (discounted) share prices are a (local) martingale. More precisely, if $(S(t) : t \in [0,T])$ is described by the dynamic (5.4.12), i.e.

$$dS(t) = \Upsilon(t)S(t)\,dt + \Phi(t)S(t)\,dW(t) \qquad \text{for all t} \in [0, \text{T}],$$

and the risk-free asset $(B(t) : t \in [0,T])$ by (5.4.11), i.e.

$$dB(t) = R(t)B(t)\,dt \qquad \text{for all t} \in [0, \text{T}],$$

then the discounted share prices $(\hat{S}(t) : t \in [0,T])$ are defined by

$$\hat{S}(t) := \frac{S(t)}{B(t)} \qquad \text{for all } t \in [0,T].$$

Discounting takes into account that the value of money you will receive at a future time has a smaller value than the money you have in your pocket today. The *first fundamental theorem of asset pricing* reads as follows:

Theorem 6.2.1. *A model (S, B) of a market as described by (5.4.12) and (5.4.11) is arbitrage-free if and only if there exists an equivalent measure Q such that the discounted share prices $\hat{S}$ are a local martingale under Q.*

Proof. See Theorem 10.14 in the monograph [2] by T. Björk. □

Recall, that there is a precise definition of an arbitrage-free market but we think of it as a market without the possibility to make money without any risk. Two measures P and Q are equivalent if for each $A \in \mathscr{A}$ they satisfy

$$P(A) = 0 \iff Q(A) = 0,$$

see part (c) of Remark 6.1.4.

The fundamental Theorem 6.2.1 requires to show the existence of an equivalent measure under which the discounted share prices are local martingales. This measure is called *equivalent martingale measure*. Note, that this measure might not be unique. Girsanov's Theorem enables us to derive conditions guaranteeing the existence of an equivalent martingale measure and to construct it explicitly.

Theorem 6.2.2. *If the volatility process $(\Phi(t) : t \in [0,T])$ satisfies*

$$P\big(\Phi(t) \neq 0\big) = 1 \qquad \textit{for all } \text{t} \in [0, \text{T}],$$

and if the market price $\Big(M(t) := \frac{\Upsilon(t)-R(t)}{\Phi(t)} : t \in [0,T]\Big)$ obeys the Novikov condition

$$E_P\left[\exp\left(\tfrac{1}{2}\int_0^T M^2(s)\,ds\right)\right] < \infty, \tag{6.2.6}$$

then the discounted stock price $\hat{S}$ *satisfies*

$$d\hat{S}(t) = \Phi(t)\hat{S}(t)\, d\widetilde{\mathrm{W}}(t) \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}],$$

where $\widetilde{\mathrm{W}}(t) := W(t) + \int_0^t M(s)\, ds$ *defines a Brownian motion* $(\widetilde{\mathrm{W}}(t) : t \in [0, T])$ *under the probability measure*

$$Q : \mathscr{A} \to [0,1], \quad Q(A) = E_P\left[\mathbb{1}_A \exp\left(-\int_0^T M(s)\, dW(s) - \tfrac{1}{2}\int_0^T M^2(s)\, ds\right)\right].$$

Proof. From Theorem 6.1.3 and Theorem 6.1.8 it follows that Q is a probability measure and $\widetilde{\mathrm{W}}$ is a Brownian motion under Q. It remains to prove the representation of $\hat{S}$. For that purpose define the function

$$f\colon [0, T] \times \mathbb{R} \to \mathbb{R}, \qquad f(t, x) = S(0)e^x,$$

which is in $C^{1,2}$. Furthermore, by letting

$$X(t) := \int_0^t \left(\Upsilon(s) - R(s) - \tfrac{1}{2}\Phi^2(s)\right) ds + \int_0^t \Phi(s)\, dW(s) \qquad \text{for all } t \in [0, T],$$

we obtain an Itô process $(X(t) : t \in [0, T])$. By using the representation (5.4.13) of S it follows that $\hat{S}(t) = f(t, X(t))$ for all $t \in [0, T]$. Then, Itô's formula in Theorem 4.6.5 implies

$$\begin{aligned} d\hat{S}(t) &= df(t, X(t)) \\ &= \left(\hat{S}(t)\big(\Upsilon(t) - R(t) - \tfrac{1}{2}\Phi^2(t)\big) + \tfrac{1}{2}\hat{S}(t)\Phi^2(t)\right) dt + \hat{S}(t)\Phi(t)\, dW(t) \\ &= \hat{S}(t)\Phi(t)\left(\frac{\Upsilon(t) - R(t)}{\Phi(t)}\, dt + dW(t)\right) \\ &= \hat{S}(t)\Phi(t)\, d\widetilde{\mathrm{W}}(t), \end{aligned}$$

since $\widetilde{\mathrm{W}}(t) = W(t) + \int_0^t M(s)\, ds$ for all $t \in [0, T]$. □

Theorem 6.2.2 implies that the discounted share prices satisfy

$$\hat{S}(t) = \hat{S}(0) + \int_0^t \Phi(s)\hat{S}(s)\, d\widetilde{\mathrm{W}}(s) \qquad \text{for all } \mathrm{t} \in [0, \mathrm{T}]\,.$$

Since $\widetilde{\mathrm{W}}$ is a Brownian motion under Q Theorem 4.4.9 implies that $\hat{S}$ is a local martingale under Q. Furthermore, the measure Q is equivalent to the measure P according to Remark 6.1.4 and thus, Theorem 6.2.1 guarantees that the model is arbitrage-free under the conditions in Theorem 6.2.2, and thus we have proved:

Theorem 6.2.3. *If the model* (S, B) *of a market described by* (5.4.12) *and* (5.4.11) *satisfies the condition in Theorem 6.2.2, then the model is arbitrage-free.*

Theorem 6.2.2 gives also the explicit construction of the measure Q which is important for pricing options on the share S. Not each model satisfies Condition (6.2.6) in Theorem 6.2.2, but this is the case in the standard Black-Scholes model:

Example 6.2.4. In the Black-Scholes model, see Example 5.4.1, the market price of risk is given by

$$M(t) = \frac{\mu - r}{\sigma} \qquad \text{for all } \mathrm{t} \in [0, \mathrm{T}],$$

where all parameters μ, r and σ are deterministic and constant. Thus, $M(t)$ is also deterministic and constant which yields that (6.2.6) is satisfied and the measure

$$Q : \mathscr{A} \to [0,1], \qquad Q(A) := E_P \left[\mathbb{1}_A \exp\left(-\tfrac{\mu - r}{\sigma} W(T) - \tfrac{1}{2} \left(\tfrac{\mu-r}{\sigma}\right)^2 T \right) \right]$$

is verified as an equivalent martingale measure in the Black-Scholes model.

By using (5.4.13) one can rewrite the dynamic of the share prices under the measure Q.

Corollary 6.2.5. *Under the conditions of Theorem 6.2.2 the share price S satisfies*

$$dS(t) = R(t)S(t)\,dt + \Phi(t)S(t)\,d\widetilde{\mathrm{W}}(t) \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}]$$

where $\widetilde{\mathrm{W}}(t) := W(t) + \int_0^t M(s)\,ds$ *for all* $t \in [0,T]$. *Consequently, S can be represented by*

$$S(t) = S(0) \exp\left(\int_0^t \left(R(s) - \tfrac{1}{2}\Phi^2(s) \right) ds + \int_0^t \Phi(s)\,d\widetilde{\mathrm{W}}(s) \right) \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}]\,.$$

Corollary 6.2.5 illustrates that the change from the measure P to Q changes only the mean rate of return Υ but not the volatility Φ of the stock price S. Using the linearity of the integrals we obtain $S(t) = S(u)S(t-u)$ for all $0 \leqslant u \leqslant t \leqslant T$, that is

$$S(t) = S(u) \exp\left(\int_u^t \left(R(s) - \tfrac{1}{2}\Phi^2(s) \right) ds + \int_u^t \Phi(s)\,d\widetilde{\mathrm{W}}(s) \right) \tag{6.2.7}$$

for all $t \in [u, T]$.

6.3. Exercises

1. Calculate the expected value of each of the following random variables:
   (a) $W(T) \exp\left(- \int_0^T s^2 dW(s) \right)$.
   (b) $W(T) \exp\left(\int_0^T s^2 dW(s) \right)$.

2. Define a stochastic process $X := (X(t) : t \in [0,T])$ by
$$X(t) := \exp(W(t)) \qquad \text{for all } t \in [0,T].$$
   (a) Determine a stochastic process $(L(t) : t \in [0,T])$ such that X becomes a martingale under the probability measure
$$Q : \mathscr{A} \to [0,1] \qquad Q(A) := E[\mathbb{1}_A L(T)].$$
   (b) Compute $E_Q\left[\left(W(t) + \tfrac{1}{2}t\right)^6\right]$ for $t \geqslant 0$, where the expectation is taken with respect to the measure Q introduced in (a).

3. Find a probability measure Q which is equivalent to P such that
$$B(t) := \begin{cases} W(t), & \text{if } t \in [0, \frac{1}{2}), \\ W(t) + \left(t - \frac{1}{2}\right) W(\frac{1}{2}), & \text{if } t \in [\frac{1}{2}, 1], \end{cases}$$
forms a Brownian motion $(B(t) : t \in [0,1])$ under Q.

4. Let $(U(t) : t \in [0,T])$ be the solution of the stochastic differential equation
$$\begin{aligned} dU(t) &= aU(t)\,dt + bU(t)\,dW(t) \qquad \text{for all } t \in [0,T], \\ U(0) &= u_0, \end{aligned}$$
for some constants $a \in \mathbb{R}$ and $b, u_0 > 0$.
   (a) Let r be a non-negative constant. Use Girsanov's Theorem to derive the existence of a probability measure Q and the existence of a Brownian motion $(\widetilde{W}(t) : t \in [0,T])$ under Q such that
$$U(t) = u_0 \exp\Big(\left(r - \tfrac{1}{2}b^2\right) t + b\,\widetilde{W}(t)\Big) \qquad \text{for all } t \in [0,T].$$
   (b) Show that the stochastic process $(\hat{U}(t) : t \in [0,T])$ defined by $\hat{U}(t) := e^{-rt}U(t)$ for a constant $r \geqslant 0$ is a martingale with respect to $\{\mathscr{F}_t^W\}_{t \geqslant 0}$ under the measure Q derived in (a).
   (c) We can consider U as a model for the price of a risky-asset and $\hat{U}$ as the discounted values assuming that the interest rate of the risk-free asset in the market is given by $r \geqslant 0$.

   Calculate the value $E_Q[C]$ of a digital option
$$C = \begin{cases} 1, & \text{if } U(T) \geqslant K, \\ 0, & \text{otherwise}, \end{cases}$$
   with maturity T and strike price $K > 0$ in this model under the measure Q from (b).

5. Define a stochastic process $(\widetilde{\mathrm{W}}(t) : t \in [0,T])$ by

$$\widetilde{\mathrm{W}}(t) := W(t) + \int_0^t e^{W(s)} \mathbb{1}_{\{|W(s)| \leqslant 1\}}\, ds \qquad \text{for all } t \in [0,T].$$

Here we use the short hand notation $\mathbb{1}_{\{|W(s)| \leqslant 1\}} := \begin{cases} 1, & \text{if } |W(s)| \leqslant 1, \\ 0, & \text{else.} \end{cases}$.

(a) Determine a probability measure Q such that $\widetilde{\mathrm{W}}$ is a Brownian motion under Q.

(b) For each $t \in [0,T]$ compute

$$E_Q\left[\left|W(t) + \int_0^t e^{W(s)} \mathbb{1}_{\{|W(s)| \leqslant 1\}}\, ds\right|^6\right].$$

(c) For the stopping time $\tau := \inf\{t \geqslant 0 : |W(t)| = 1\}$ compute $E_P[\widetilde{\mathrm{W}}(\tau)]$. (*)

(d) For the stopping time $\sigma := \inf\{t \geqslant 0 : W(t) = 1 - \int_0^t e^{W(s)} \mathbb{1}_{\{|W(s)| \leqslant 1\}}\, ds\}$ compute (*)

$$E_Q[e^{-\sigma/2}].$$

7

Martingale Representation Theorem

In Theorem 4.3.1 we show that if the integrand Φ is in $\mathscr{H}$ then the stochastic integral defines a martingale. In this chapter we ask the converse question: given a martingale can it be represented by a stochastic integral? One can consider this question as an analogue of differentiation in calculus: given a function $f : [0,T] \to \mathbb{R}$ does there exist a function $g : [0,T] \to \mathbb{R}$ such that

$$f(t) = f(0) + \int_0^t g(s)\,ds \qquad \text{for all } t \in [0,T]?$$

If f is assumed to be differentiable then we can just choose $g = f'$.

As in the other chapters we assume that W is a Brownian motion defined on a probability space $(\Omega, \mathscr{A}, P)$. In the previous chapters we can consider any filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ as long as the Brownian motion satisfies

(i) $W(t)$ is $\mathscr{F}_t$-adapted for all $t \geqslant 0$;

(ii) $W(t) - W(s)$ is independent of $\mathscr{F}_s$ for all $0 \leqslant s \leqslant t$,

see the introductory part to Chapter 4. Our main result in this chapter requires that we consider only the augmented filtration $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$ generated by the Brownian motion W, which we introduce in the end of Section 1.1.

7.1. The Theorem

From Example 6.1.9 we know already a class of random variables, for which we obtain a representation by a stochastic integral.

Example 7.1.1. For a deterministic function $\varphi : [0,T] \to \mathbb{R}$ with $\int_0^T |\varphi(s)|^2 \, ds < \infty$ define

$$\mathscr{E}(t) := \exp\left(-\int_0^t \varphi(s)\, dW(s) - \tfrac{1}{2}\int_0^t \varphi^2(s)\, ds \right) \qquad \text{for all } t \in [0,T]. \tag{7.1.1}$$

In Example 6.1.9 we show

$$d\mathscr{E}(t) = -\varphi(t)\mathscr{E}(t)\, dW(t).$$

Consequently, if we define the random variable $Y := \mathscr{E}(T) : \Omega \to \mathbb{R}$ and $\Phi(t) := -\varphi(t)\mathscr{E}(t)$ for all $t \in [0,T]$ we obtain

$$Y = 1 + \int_0^T \Phi(t)\, dW(t). \tag{7.1.2}$$

Note, by the proof of Theorem 6.1.8 it follows that $\Phi \in \mathscr{H}$.

By linearity we can extend the class of random variables which enables a representation of the form (7.1.2). Let Y_1, $Y_2 : \Omega \to \mathbb{R}$ be random variables of the form

$$Y_i := \alpha_i \mathscr{E}_i(T), \qquad i = 1, 2,$$

for some $\alpha_i \in \mathbb{R}$ and with

$$\mathscr{E}_i(t) := \exp\left(-\int_0^t \varphi_i(s)\, dW(s) - \tfrac{1}{2}\int_0^t \varphi_i^2(s)\, ds \right) \qquad \text{for all } t \in [0,T],$$

where $\varphi_i : [0,T] \to \mathbb{R}$ satisfies $\int_0^T \varphi_i^2(s)\, ds < \infty$. It follows from Example 7.1.1 that the random variable $Y := Y_1 + Y_2$ can be represented as

$$Y = \alpha + \int_0^T \Phi(t)\, dW(t), \tag{7.1.3}$$

where $\alpha := \alpha_1 + \alpha_2$ and $\Phi(t) := -\big(\alpha_1\varphi_1(t)\mathscr{E}_1(t) + \alpha_2\varphi_2(t)\mathscr{E}_2(t)\big)$ for all $t \in [0,T]$. Note, that $\Phi \in \mathscr{H}$. Consequently, all random variables in the set

$$\begin{aligned}\mathscr{L} := \Bigg\{ Y = \sum_{k=1}^m \alpha_k \mathscr{E}_k(T) : \mathscr{E}_k(T) := \exp\left(-\int_0^T \varphi_k(s)\, dW(s) - \tfrac{1}{2}\int_0^T \varphi_k^2(s)\, ds \right) \\ \text{for } \varphi_k : [0,T] \to \mathbb{R} \text{ with } \int_0^T \varphi_k^2(s)\, ds < \infty,\ \alpha_k \in \mathbb{R},\ m \in \mathbb{N} \Bigg\}\end{aligned}$$

can be represented in the form (7.1.3).

In our first theorem, Itô's representation theorem, we extend the class $\mathscr{L}$ of random variables which have a representation of the form (7.1.3) to all $\mathfrak{F}_T^W$-measurable random variables X with $E[|X|^2] < \infty$. The idea of the proof is to approximate the random variable X by a sequence $\{Y_n\}_{n\in\mathbb{N}}$ of random variables $Y_n \in \mathscr{L}$. For these random variables Y_n we know already that they obey the representation (7.1.3) and thus, we can hope that this representation carries over to the limit. Fundamental for this argumentation is the following approximation result:

Lemma 7.1.2. *For every $\mathfrak{F}_T^W$-measurable random variable X with $E[|X|^2] < \infty$ there exists a sequence $\{Y_n\}_{n\in\mathbb{N}} \subseteq \mathscr{L}$ with*

$$\lim_{n\to\infty} E[|Y_n - X|^2] = 0.$$

Proof. See [16, L. 4.3.2]. □

We apply the approximation result above to obtain the following representation theorem:

Theorem 7.1.3. *(Itô's representation theorem)*
For every $\mathfrak{F}_T^W$-measurable random variable X with $E[|X|^2] < \infty$ there exists an $\{\mathfrak{F}_t^W\}_{t\in[0,T]}$-adapted stochastic process $(\Phi(t) : t \in [0,T]) \in \mathscr{H}$ such that

$$X = E[X] + \int_0^T \Phi(s)\, dW(s). \tag{7.1.4}$$

The representation is unique in the sense that if there is another adapted process $(\Psi(t) : t \in [0,T]) \in \mathscr{H}$ satisfying (7.1.4) then $\Phi(t)(\omega) = \Psi(t)(\omega)$ for a.a. $(t,\omega) \in [0,T] \times \Omega$.

The important condition in Theorem 7.1.3 is that the random variable X is $\mathfrak{F}_T^W$-measurable where $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$ is the augmented filtration generated by the Brownian motion W. The need for this can be seen by the representation (7.1.4) because the only source of uncertainty or randomness is the process Φ and the Brownian motion W both measurable with respect to the filtration generated by W; see Example 7.1.5. From Theorem 7.1.3 one can conclude the following dynamic version of this result.

Theorem 7.1.4. *(Martingale representation theorem)*
For every martingale $(M(t) : t \in [0,T])$ with respect to $\{\mathfrak{F}_t^W\}_{t\in[0,T]}$ and with $E[|M(T)|^2] < \infty$ there exists an $\{\mathfrak{F}_t^W\}_{t\in[0,T]}$-adapted stochastic process $(\Phi(t) : t \in [0,T]) \in \mathscr{H}$ such that

$$M(t) = E[M(0)] + \int_0^t \Phi(s)\, dW(s) \qquad \textit{for all } t \in [0,T]. \tag{7.1.5}$$

The stochastic process Φ is unique in the sense as in Theorem 7.1.3.

The proof of Theorem 7.1.4 is not constructive, i.e. we do not obtain an explicit formula of Φ which satisfies (7.1.5). However, with another area of modern probability theory, *Malliavin calculus*, one can express Φ in terms of X.

Example 7.1.5. Let $(N(t) : t \geqslant 0)$ be a Poisson process with intensity $\lambda > 0$ independent of the Brownian motion W. Define a filtration $\{\mathscr{F}_t^{W,N}\}_{t\geqslant 0}$ by defining

$$\mathscr{F}_t^{W,N} := \sigma\Big(\big(W(s), N(s)\big)^{-1}([a_1, b_1] \times [a_2, b_2]) : s \in [0,t],\ -\infty < a_i \leqslant b_i < \infty,\ i = 1, 2\Big).$$

Since W and N are independent, the Brownian motion W satisfies the conditions (i) and (ii) recalled in the introduction of this Chapter. The same reason implies that $M := (N(t) - \lambda t : t \geqslant 0)$ is a martingale with respect to $\{\mathscr{F}_t^{W,N}\}_{t\geqslant 0}$, see Exercise 2.5.1. Since the paths of M are not continuous whereas the martingale $(\int_0^t \Phi(s)\, dW(s) : t \in [0,T])$ has continuous paths for every $\Phi \in \mathscr{H}$ according to Theorem 4.3.1, the Poisson process N can not be represented in the form (7.1.5).

7.2. Financial mathematics: complete models

We now assume the general setting of a model of a financial market as in Section 5.4. For that purpose, let $\{\mathscr{F}_t\}_{t\geqslant 0}$ be an arbitrary filtration satisfying the usual conditions and let $(W(t) : t \in [0,T])$ be a Brownian motion with respect to $\{\mathscr{F}_t\}_{t\geqslant 0}$, i.e. it obeys the conditions mentioned in the introduction of Chapter 4. Let $S := (S(t) : t \in [0,T])$ model the prices of a risky asset and $(B(t) : t \in [0,T])$ be the values of a risk-free asset as described in Section 5.4. An *option (with maturity T)* on the underlying share S is an $\mathscr{F}_T$-measurable random variable C. A typical example is the European call option with strike price $K > 0$:

$$C = \max\{S(T) - K, 0\}.$$

The option is purchased at time $t = 0$ and exercised at time $t = T$. The seller of an option has to pay the amount C at time T. For that reason she is interested in hedging the option by trading according to a strategy which has exactly the same value as the option at time T.

Definition 7.2.1.

(a) *An option C is called* attainable *if there exists a self-financing strategy* (Γ, Δ) *such that the value process $V = V(\Gamma, \Delta)$ satisfies*

$$V(T) = C \qquad P\text{-}a.s.$$

In this case (Γ, Δ) *is called a* replicating strategy for C.

(b) *A model, that is (S, B), is called complete if every contingent claim is attainable.*

It is easy to find examples of a model such that not every option is attainable. In complete models every option has a unique arbitrage-free price which is given by the initial investment $V(0)$ required for the replicating strategy (Γ, Δ). It is easy to see that any other price leads to an arbitrage opportunity. A sufficient condition for completeness of a model is given in the next result:

Theorem 7.2.2. *Assume that Q is an equivalent local martingale measure and let C be an option with $E_Q\left[\left|\frac{1}{B(T)}C\right|\right] < \infty$. Let $(\hat{\Pi}_C(t) : t \in [0,T])$ be the arbitrage-free discounted price process of C defined by*

$$\hat{\Pi}_C(t) := E_Q\left[\frac{C}{B(T)}\middle|\mathscr{F}_t\right] \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}]\,.$$

If there exists an adapted process $(H(t) : t \in [0,T])$ such that the discounted price process $\hat{\Pi}_C$ satisfies

$$\hat{\Pi}_C(t) = \hat{\Pi}_C(0) + \int_0^t H(s)\, d\hat{S}(s) \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}],$$

then C is attainable and the replicating strategy $((\Gamma(t), \Delta(t)) : t \in [0,T])$ is given by

$$\begin{aligned} \Delta(t) &= H(t) \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}], \\ \Gamma(t) &= \hat{\Pi}_C(t) - \Delta(t)\hat{S}(t) \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}]\,. \end{aligned}$$

We call the discounted price process $\hat{\Pi}_C$ of the option C *arbitrage-free* if trading the option C does not lead to an arbitrage opportunity. This can be formally defined by extending the underlying market but we do not go into these details.

By applying the martingale representation Theorem 7.1.4 we obtain a more specific condition for the existence of a replicating strategy. Note, that it is required that the filtration $\{\mathscr{F}_t\}$ is generated by the Brownian motion.

Theorem 7.2.3. *(Second Fundamental Theorem of Asset Pricing)*
Assume that Q is an equivalent risk-neutral measure and that the volatility process $(\Phi(t) : t \in [0,T])$ satisfies

$$P\big(\Phi(t) \neq 0\big) = 1 \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}]\,.$$

Let C be an option with $E_Q\left[\left|\frac{1}{B(T)}C\right|^2\right] < \infty$ and let $(\hat{\Pi}_C(t) : t \in [0,T])$ be the arbitrage-free discounted price process of C defined by

$$\hat{\Pi}_C(t) := E_Q\left[\frac{C}{B(T)}\middle|\mathscr{F}_t\right] \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}]\,.$$

If $\mathscr{F}_t = \mathfrak{F}_t^W$ for all $t \in [0,T]$, then there exists an adapted stochastic process $(H(t) : t \in [0,T])$ such that the discounted price process $\hat{\Pi}_C$ satisfies

$$\hat{\Pi}_C(t) = \hat{\Pi}_C(0) + \int_0^t H(s)\, d\hat{S}(s) \qquad \textit{for all } \mathrm{t} \in [0, \mathrm{T}]\,.$$

Proof. (Sketch) After some arguments the martingale representation theorem 7.1.4 guarantees that there exists a stochastic process $(Z(t) : t \in [0,T]) \in \mathscr{H}$ such that

$$\hat{\Pi}_C(t) = \hat{\Pi}_C(0) + \int_0^t Z(s)\, d\widetilde{\mathrm{W}}(s) \qquad \text{for all } t \in [0,T].$$

By defining $H(t) := Z(t)\big(\Phi(t)\hat{S}(t)\big)^{-1}$ for all $t \in [0,T]$ the representation of $\hat{S}$ in Theorem 6.2.2 implies that

$$\begin{aligned}\hat{\Pi}_C(t) &= \hat{\Pi}_C(0) + \int_0^t H(t)\Phi(t)\hat{S}(t)\, d\widetilde{\mathrm{W}}(s)\\ &= \hat{\Pi}_C(0) + \int_0^t H(t) d\hat{S}(t) \qquad \text{for all } t \in [0,T].\end{aligned}$$

□

By applying Theorem 7.2.2 it follows that under the conditions in Theorem 7.2.3 the contingent claim C is attainable. Theorem 7.2.2 states the replicating strategy in terms of the stochastic process H. However, the existence of this stochastic process H is derived in Theorem 7.2.3 by applying the martingale representation theorem 7.1.4 and the latter does not give an explicit construction of this stochastic process. Thus, the existence of a replicating strategy is guaranteed but still one does not know it explicitly. A possible approach to obtain an explicit representation of the replicating strategy is to use a connection between stochastic differential equations and partial differential equations. In the Black-Scholes model this leads to the Black-Scholes partial differential equations.

Example 7.2.4. In the standard Black-Scholes model in Example 5.4.1 the share prices evolves according to

$$dS(t) = \mu S(t)\,dt + \sigma S(t)\,dW(t).$$

Clearly, the solution is $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$-adapted and thus, the model is complete.

Example 7.2.5. Consider the volatility model in Example 5.4.2:

$$dS(t) = \mu S(t)\,dt + \Phi(t)S(t)\,dW_1(t) \qquad \text{for all t} \in [0,\mathrm{T}], \tag{7.2.6}$$
$$d\Phi(t) = f(t,\Phi(t))\,dt + g(t,\Phi(t))\,dW_2(t) \qquad \text{for all t} \in [0,\mathrm{T}], \tag{7.2.7}$$

where W_1 and W_2 are two independent Brownian motions. Clearly, Φ is adapted with respect to the filtration $\{\mathfrak{F}_t^{W_2}\}_{t\geqslant 0}$ since the Brownian motion W_2 is the only random source in equation (7.2.7). However, since Φ is $\{\mathfrak{F}_t^{W_2}\}_{t\geqslant 0}$-adapted and W_1 is $\{\mathfrak{F}_t^{W_1}\}_{t\geqslant 0}$-adapted, we have to consider the stochastic differential equation in (7.2.6) with respect to the filtration $\{\mathscr{F}_t^{W_1,W_2}\}_{t\geqslant 0}$ defined by

$$\mathscr{F}_t^{W_1,W_2} := \sigma\Big(\big(W_1(s),W_2(s)\big)^{-1}([a_1,b_1]\times[a_2,b_2]) : \, s\in[0,t],\, -\infty < a_i \leqslant b_i < \infty,\, i=1,2\Big).$$

In particular, an option C with maturity T on the share S is $\mathscr{F}_T^{W_1,W_2}$-measurable, and thus we can not apply Theorem 7.2.3 to conclude the existence of a hedging strategy for C.

7.3. Exercise

1. In each of the following cases find an $\{\mathfrak{F}_t^W\}_{t\in[0,T]}$-adapted stochastic process $(\Phi(t) : t\in[0,T])\in\mathscr{H}$ such that

$$X = E[X] + \int_0^T \Phi(s)\,dW(s),$$

where

(a) $X = W(T)$;

(b) $X = \int_0^T W(s)\,ds$;

(c) $X = W^2(T)$;

(d) $X = \int_0^T W^2(s)\,ds$;

(e) $X = W^3(T)$.

2. Let Z be an $\mathfrak{F}_T^W$-measurable random variable with $E[|Z|^2] < \infty$ and define

$$M(t) := E[Z|\mathfrak{F}_t^W] \qquad \text{for all } t \in [0,T].$$

Recall from Exercise 2.5.2 that $(M(t) : t \in [0,T])$ is a martingale w.r.t. $\{\mathfrak{F}_t^W\}_{t\in[0,T]}$.

(a) Show that $E[|M(T)|^2] < \infty$.

(b) Find in each of the following cases an $\{\mathfrak{F}_t^W\}_{t\in[0,T]}$-adapted stochastic process $(\Phi(t) : t \in [0,T]) \in \mathscr{H}$ such that

$$M(t) = E[M(0)] + \int_0^T \Phi(s)\, dW(s) \qquad \text{for all } t \in [0,T]$$

where

(i) $Z = W^2(T)$.

(ii) $Z = W^3(T)$.

3. The following result is a conclusion of a special martingale representation form which is often applied in financial mathematics. We can assume an arbitrary filtration $\{\mathscr{F}_t\}_{t\geqslant 0}$ such that (i) and (ii) in the introductory part of this chapter are satisfied.

Let $(X(t) : t \in [0,T])$ be an Itô process of the form

$$X(t) = X(0) + \int_0^t \Upsilon(s)\, ds + \int_0^t \Phi(s)\, dW(s) \qquad \text{for all t} \in \text{[0, T]},$$

where $(\Upsilon(t) : t \in [0,T])$ and $(\Phi(t) : t \in [0,T])$ are adapted stochastic processes satisfying

$$E\left[\int_0^T |\Upsilon(s)|\; ds\right] + E\left[\int_0^T |\Phi(s)|^2\; ds\right] < \infty.$$

(a) Show that if X is a martingale w.r.t. $\{\mathscr{F}_t\}_{t\in[0,T]}$, then it follows ($\natural$)

$$E\big[\Upsilon(t)|\mathscr{F}_s\big] = 0 \quad \text{for a.a. } t \geqslant s.$$

(b) Show that if X is a martingale w.r.t. $\{\mathscr{F}_t\}_{t\in[0,T]}$, then it follows ($\natural$)

$$\Upsilon(t)(\omega) = 0 \quad \text{for a.a. } (t,\omega) \in [0,T] \times \Omega.$$

Part (b) can be deduced from (a) by using the following result:

If $X : \Omega \to \mathbb{R}$ is a random variable with $E[|X|] < \infty$ and $\{\mathscr{C}_k\}_{k\in\mathbb{N}}$ is a family of increasing σ-algebras $\mathscr{C}_k \subseteq \mathscr{A}$ then

$$\lim_{k\to\infty} E[X|\mathscr{C}_k] = E[X|\mathscr{C}] \qquad P\text{-a.s. and in } L_P^1(\Omega),$$

where $\mathscr{C} := \sigma(\cup_{k=1}^\infty \mathscr{C}_k)$.

A

Solutions

A.1. Solution Chapter 1

1. (a) For every $t_0 \in I$ one has

$$\{X(t) = Y(t) \quad \text{for all } t \in I\} \subseteq \{X(t_0) = Y(t_0)\},$$

which shows the claim by

$$1 = P\left(X(t) = Y(t) \quad \text{for all } t \in I\right) \leqslant P\left(X(t_0) = Y(t_0)\right).$$

(b) The continuity[1] of the paths imply

$$C := \{X(t) = Y(t) \text{ for all } t \in [0,\infty)\} = \{X(t) = Y(t) \text{ for all } t \in [0,\infty) \cap \mathbb{Q}\}.$$

Consequently, the set C is in $\mathscr{A}$ and we have

$$\begin{aligned}
P(C) &= P\big(X(t) = Y(t) \text{ for all } t \in [0,\infty) \cap \mathbb{Q}\,\big) \\
&= 1 - P\big(X(t) \neq Y(t) \text{ for some } t \in [0,\infty) \cap \mathbb{Q}\,\big) \\
&= 1 - P\left(\bigcup_{t \in [0,\infty)\cap\mathbb{Q}} \{X(t) \neq Y(t)\}\right) \\
&\geqslant 1 - \sum_{t \in [0,\infty)\cap\mathbb{Q}} P\big(X(t) \neq Y(t)\big) \\
&= 1.
\end{aligned}$$

[1] for simplicity we assume that all paths are continuous, not only almost all

2. For fixed $t \geqslant 0$ it follows that

$$P\big(X(t) = Y(t)\big) = P\big(0 = Y(t)\big) = P\big(T \neq t\big) = 1.$$

The last equality follows from the assumption that the distribution of T has a density, say $f : \mathbb{R} \to \mathbb{R}_+$. Then

$$P(T \in B) = \int_B f(u)\, du \qquad \text{for all } B \in \mathfrak{B}(\mathbb{R}).$$

It follows for a set B of the form $B = \{a\}$ for a constant $a \in \mathbb{R}$ that $P(T \in \{a\}) = 0$. However,

$$\begin{aligned}\{\omega \in \Omega : X(t)(\omega) = Y(t)(\omega) \text{ for all } t \geqslant 0\} &= \{\omega \in \Omega : 0 = Y(t)(\omega) \text{ for all } t \geqslant 0\} \\ &= \emptyset,\end{aligned}$$

which shows that X and Y are not indistinguishable.

3. Since X and Y are modifications from each other the set $N_t := \{X(t) \neq Y(t)\}$ obeys $P(N_t) = 0$ for each $t \geqslant 0$. Consequently, for every $t_1, \dots, t_n \geqslant 0$ we have $P(N_{t_1} \cup \dots \cup N_{t_n}) = 0$, and thus

$$\begin{aligned}&P\big(X(t_1) \in B_1, \dots, X(t_n) \in B_n\big) \\ &\qquad = P\big((X(t_1) \in B_1, \dots, X(t_n) \in B_n) \backslash (N_{t_1} \cup \dots \cup N_{t_n})\big) \\ &\qquad = P\big((Y(t_1) \in B_1, \dots, Y(t_n) \in B_n) \backslash (N_{t_1} \cup \dots \cup N_{t_n})\big) \\ &\qquad = P\big(Y(t_1) \in B_1, \dots, Y(t_n) \in B_n\big).\end{aligned}$$

4. (a) $\mathscr{C} = \{\emptyset, \Omega, \{1,2,3\}, \{3,4,5\}, \{3\}, \{1,2,4,5\}, \{1,2\}, \{4,5\}\}$.

   (b) The random variable X is measurable w.r.t. $\mathscr{C}$ since we have for each $A \in \mathfrak{B}(\mathbb{R})$:

$$\begin{aligned}&\text{if } 0 \in A, 1, 10 \notin A: \quad X^{-1}(A) = \{1,2\} \in \mathscr{C}; \\ &\text{if } 1 \in A, 0, 10 \notin A: \quad X^{-1}(A) = \{4,5\} \in \mathscr{C}; \\ &\text{if } 10 \in A, 0, 1 \notin A: \quad X^{-1}(A) = \{3\} \in \mathscr{C}; \\ &\text{if } 0, 1, 10 \notin A: \quad X^{-1}(A) = \emptyset \in \mathscr{C}; \\ &\text{if } 0, 1, 10 \in A: \quad X^{-1}(A) = \Omega \in \mathscr{C};\end{aligned}$$

All other cases can be reduced to these, e.g. if $0, 1, \in A$ but $10 \notin A$ then:

$$X^{-1}(A) = X^{-1}(\{0\}) \cup X^{-1}(\{1\}) = \{1,2\} \cup \{4,5\} = \{1,2,4,5\} \in \mathscr{C}.$$

   (c) $\mathscr{D} = \sigma(Y) = \{\Omega, \emptyset, \{1\}, \{2,3,4,5\}\}$.

5. For every $n \in \mathbb{N}$ and $C \in \mathfrak{B}(\mathbb{R})$ the definition of X_n yields

$$\begin{aligned}(X_n(C))^{-1} &= \begin{cases} \frac{1}{2}C \cap [0, 1 - \frac{1}{n}], & \text{if } 0 \notin C, \\ \left(\frac{1}{2}C \cap [0, 1 - \frac{1}{n}]\right) \bigcup (1 - \frac{1}{n}, 1], & \text{if } 0 \in C, \end{cases} \\ &= \begin{cases} \frac{1}{2}C \cap (0, 1 - \frac{1}{n}], & \text{if } 0 \notin C, \\ \left(\frac{1}{2}C \cap (0, 1 - \frac{1}{n}]\right) \bigcup (1 - \frac{1}{n}, 1] \bigcup \{0\}, & \text{if } 0 \in C, \end{cases}\end{aligned}$$

where we use the shorthand notation $\frac{1}{2}C := \{\frac{1}{2}y : y \in C\}$. Thus, the σ-field $\mathscr{F}_n^X$ must contain at least

$$\mathfrak{C} := \left\{\tfrac{1}{2}C \cap (0, 1 - \tfrac{1}{n}], \tfrac{1}{2}C \cap [0, 1 - \tfrac{1}{n}] \cup (1 - \tfrac{1}{n}, 1] \cup \{0\} \quad \text{for all } C \in \mathfrak{B}(\mathbb{R})\right\}.$$

Since $\{\frac{1}{2}C \cap (0, 1 - \frac{1}{n}] : C \in \mathfrak{B}(\mathbb{R})\} = \mathfrak{B}\left((0, 1 - \frac{1}{n}]\right)$ this can be written in the form

$$\mathfrak{C} = \left\{A \cup B : A \in \mathfrak{B}\left((0, 1 - \tfrac{1}{n}]\right), B = \emptyset \text{ or } B = (1 - \tfrac{1}{n}, 1] \cup \{0\}\right\}.$$

Since this is already a σ-field (check!) and $\mathscr{F}_n^X$ is the smallest σ-field which contains $\mathfrak{C}$ it follows $\mathscr{F}_n^X = \mathfrak{C}$.

6. For each $t \geqslant 0$ it follows from the very definition of stopping times that

$$\begin{aligned}\{\sigma \wedge \tau \leqslant t\} &= \{\sigma \leqslant t\} \cup \{\tau \leqslant t\} \in \mathscr{F}_t, \\ \{\sigma \vee \tau \leqslant t\} &= \{\sigma \leqslant t\} \cap \{\tau \leqslant t\} \in \mathscr{F}_t,\end{aligned}$$

which shows both part (b) and (c).

7. (a) The empty set $\emptyset$ is in $\mathscr{F}_\tau$ since $\emptyset \cap \{\tau \leqslant t\} = \emptyset \in \mathscr{F}_t$ for all $t \geqslant 0$. If $B \in \mathscr{F}_\tau$ it follows that

$$B^c \cap \{\tau \leqslant t\} = \{\tau \leqslant t\} \backslash (B \cap \{\tau \leqslant t\}) \in \mathscr{F}_t \qquad \text{for all } t \geqslant 0,$$

since both $\{\tau \leqslant t\}$ and $B \cap \{\tau \leqslant t\}$ are in $\mathscr{F}_t$ for each $t \geqslant 0$. Thus, $B^c \in \mathscr{F}_\tau$. For $B_1, B_2, \ldots \in \mathscr{F}_\tau$ we obtain

$$\left(\bigcup_{k=1}^{\infty} B_k\right) \cap \{\tau \leqslant t\} = \bigcup_{k=1}^{\infty} (B_k \cap \{\tau \leqslant t\}) \in \mathscr{F}_t \qquad \text{for all } t \geqslant 0.$$

Thus, $\bigcup_{k=1}^{\infty} B_k \in \mathscr{F}_\tau$.

(b) Since $\sigma \leqslant \tau$ we have that $\{\tau \leqslant t\} \subseteq \{\sigma \leqslant t\}$ for all $t \geqslant 0$ (this is NOT a typo; think!). It follows for $A \in \mathscr{F}_\sigma$ that

$$A \cap \{\tau \leqslant t\} = \underbrace{A \cap \{\sigma \leqslant t\}}_{\in \mathscr{F}_t \text{ since } A \in \mathscr{F}_\sigma} \cap \underbrace{\{\tau \leqslant t\}}_{\in \mathscr{F}_t \text{ since } \tau \text{ stopping time}} \in \mathscr{F}_t \qquad \text{for all } t \geqslant 0.$$

Thus, we obtain $A \in \mathscr{F}_\tau$.

8. In order to establish (a)-(e) the same notations as in 1.1.2 are used. We use several times the fact that if X and Y are two independent, non-negative random variables and Y has the density f_Y then

$$P\big(X + Y > s,\, Y \leqslant s\big) = \int_0^s P\big(X > s - u\big) f_y(u)\, du \qquad \text{for all } s \geqslant 0. \tag{1.1.1}$$

(a) First we show for each $n \in \mathbb{N}$ that

$$P\left(S_n > t\right) = e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} \qquad \text{for all } t \geqslant 0, \tag{1.1.2}$$

which can be done by induction: for $n = 1$ the definition of X_1 implies

$$P(S_1 > t) = P(X_1 > t) = e^{-\lambda t} \qquad \text{for all } t \geqslant 0.$$

Assume that (1.1.2) is true for $n \in \mathbb{N}$. Then we obtain for every $t \geqslant 0$

$$\begin{aligned} P\big(S_{n+1} > t\big) &= P\big(S_n + X_{n+1} > t\big) \\ &= P\big(S_n + X_{n+1} > t, X_{n+1} > t\big) + P\big(S_n + X_{n+1} > t, X_{n+1} \leqslant t\big) \end{aligned} \tag{1.1.3}$$

Due to the equality $\{S_n + X_{n+1} > t,\, X_{n+1} > t\} = \{X_{n+1} > t\}$ we obtain

$$P(S_n + X_{n+1} > t, X_{n+1} > t) = P(X_{n+1} > t) = e^{-\lambda t}. \tag{1.1.4}$$

Since X_{n+1} is exponentially distributed with parameter λ, the density f_X of its probability distribution is given by $f_X(s) := \lambda \exp(-\lambda s)$ for all $s \geqslant 0$. By using (1.1.2) for n and applying formula (1.1.1) (S_n and X_{n+1} are independent) implies

$$\begin{aligned} P\big(S_n + X_{n+1} > t, X_{n+1} \leqslant t\big) &= \int_0^t P\big(S_n > t - s\big) f_X(s)\, ds \\ &= \int_0^t e^{-\lambda(t-s)} \sum_{k=0}^{n-1} \frac{(\lambda(t-s))^k}{k!} \lambda e^{-\lambda s}\, ds \\ &= e^{-\lambda t} \sum_{k=0}^{n-1} \frac{\lambda^{k+1}}{k!} \int_0^t (t-s)^k\, ds. \end{aligned} \tag{1.1.5}$$

By using (1.1.4) and (1.1.5) in (1.1.3) it follows (1.1.2) for $n + 1$.

In order to show (a) note that $\{N(t) < n\} = \{S_n > t\}$ for all $t \geqslant 0$ and $n \in \mathbb{N}_0$. Thus,

$$\begin{aligned} P\big(N(t) = n\big) &= P\big(N(t) < n + 1\big) - P\big(N(t) < n\big) \\ &= P\big(S_{n+1} > t\big) - P\big(S_n > t\big) \\ &= e^{-\lambda t} \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!} - e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} \\ &= e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \end{aligned}$$

(b) The definition of R_n^t yields for every $t \geqslant 0$ and $n \in \mathbb{N}$

$$R_n^t = S_{N(t)+1} - t + X_{N(t)+2} + \cdots + X_{N(t)+n} = S_{N(t)+n} - t.$$

Thus, we have for each $s > 0$ and $t \geqslant 0$

$$\begin{aligned} Q^t(s) &= \max\big\{k \in \{0,1,\dots,\} : \, S_{N(t)+k} - t \leqslant s\big\} \\ &= \max\big\{k \in \{0,1,\dots,\} : \, S_k \leqslant t+s\big\} - N(t) \\ &= N(t+s) - N(t). \end{aligned}$$

Clearly, $Q^t(0) = 0$ for all $t \geqslant 0$.

(c) We first show for every $s, t \geqslant 0$ and $n \in \mathbb{N}$:

$$P\big(Y_1^t > s | N(t) = n\big) = P\big(X_1 > s\big). \tag{1.1.6}$$

For this purpose, note (draw a picture!) the equalities of the sets

$$\{N(t) = n\} = \{X_{n+1} > t - S_n, \, t \geqslant S_n\} \tag{1.1.7}$$

$$\{Y_1^t > s, \, N(t) = n\} = \{X_{n+1} > s + t - S_n, \, t \geqslant S_n\}. \tag{1.1.8}$$

Since X_{k+1} is exponentially distributed its probability distribution obeys

$$P\big(X_{k+1} > x + y\big) = e^{-\lambda(x+y)} = P\big(X_{k+1} > x\big)P\big(\big(X_{k+1} > y\big) \qquad \text{for every } x, y \geqslant 0.$$

Let f_{S_n} denote the density of the probability distribution for S_n for each $n \in \mathbb{N}$. Since X_{n+1} and S_n are independent, equality (1.1.1) yields

$$\begin{aligned} P\big(Y_1^t > s, \, N(t) = n\big) &= P\big(X_{n+1} > s + t - S_n, \, S_n \leqslant t\big) \\ &= \int_0^t P\big(X_{n+1} > s + t - u\big) f_{S_n}(u)\, du \\ &= P\big(X_{n+1} > s\big) \int_0^t P\big(X_{n+1} > t - u\big) f_{S_n}(u)\, du \\ &= P\big(X_1 > s\big) P\big(X_{n+1} > t - S_n, \, S_n \leqslant t\big) \\ &= P\big(X_1 > s\big) P\big(N(t) = n\big), \end{aligned}$$

which shows (1.1.6). In order to consider the general case, not that equality (1.1.8) can be extended to

$$\begin{aligned} &\{Y_1^t > s_1, \, Y_2^t > s_2, \, \dots, \, Y_k^t > s_k, \, N(t) = n\} \\ &\quad = \{X_{n+1} > s_1 + t - S_n, \, t \geqslant S_n\} \cap \{X_{n+2} > s_2\} \cap \cdots \cap \{X_{n+k} > s_k\} \end{aligned}$$

for all $s_1, \dots, s_k, \, t \geqslant 0$ and $n \in \mathbb{N}$. By using this equality together with (1.1.8) and our first result (1.1.6), we can conclude

$$\begin{aligned} &P\big(Y_1^t > s_1, \, Y_2^t > s_2, \, \dots, \, Y_k^t > s_k, \, N(t) = n\big) \\ &\quad = P\big(X_{n+1} > s_1 + t - S_n, \, t \geqslant S_n\big)\, P\big(X_{n+2} > s_2\big) \cdots P\big(X_{n+k} > s_k\big) \\ &\quad = P\big(Y_1^t > s_1, \, N(t) = n\big) P\big(X_2 > s_2\big) \dots P\big(X_k > s_k\big) \\ &\quad = P\big(Y_1^t > s_1 | N(t) = n\big) P\big(N(t) = n\big) P\big(X_2 > s_2\big) \dots P\big(X_k > s_k\big) \\ &\quad = P\big(X_1 > s_1\big) P\big(N(t) = n\big) P\big(X_2 > s_2\big) \dots P\big(X_k > s_k\big), \end{aligned} \tag{1.1.9}$$

which shows the claim by dividing both sides with $P(N(t)=n)$.

(d) Taking expectation at both sides of (1.3.1) yields for all $s_i \geqslant 0$, $i=1,\dots,k$ and $k \in \mathbb{N}$:

$$P\big(Y_1^t > s_1, \dots, Y_k^t > s_k\big) = P(X_1 > s_1) \cdots P(X_k > s_k). \tag{1.1.10}$$

By choosing $s_i = 0$ for $i = 1,\dots,k$, $i \neq j$ and $s_j \geqslant 0$ we obtain

$$P\big(Y_j^t > s_j\big) = P\big(X_j > s_j\big). \tag{1.1.11}$$

Using this equality in (1.1.10) results for all $k \in \mathbb{N}$ in

$$P\big(Y_1^t > s_1, \dots, Y_k^t > s_k\big) = P(Y_1^t > s_1) \cdots P(Y_k^t > s_k) \qquad \text{for all } s_i \geqslant 0, i=1,\dots,k,$$

which shows that $Y_1^t, \dots, Y_k^t$ are independent for all $k \in \mathbb{N}$. Since the equality (1.1.11) implies that Y_i^t is exponentially distributed for all $i \in \mathbb{N}$ it follows that the stochastic process Q^t is defined as N in Example 1.1.2, i.e. Q^t is a Poisson process for each $t \geqslant 0$. Equation (1.1.11) implies that for each $t \geqslant 0$ we have $Q^t(s) \stackrel{\mathscr{D}}{=} N(s)$ for all $s \geqslant 0$.

By applying (1.1.11) to (1.1.9) we obtain for all $t \geqslant 0$, $s_i \geqslant 0$, $i = 1,\dots,k$ and $k \in \mathbb{N}$:

$$\begin{aligned} &P\big(Y_1^t > s_1,\, Y_2^t > s_2,\, \dots,\, Y_k^t > s_k,\, N(t)=n\big) \\ &\quad = P\big(Y_1^t > s_1\big)P\big(Y_2^t > s_2\big) \dots P\big(Y_k^t > s_k\big)P\big(N(t)=n\big), \end{aligned}$$

which shows that $N(t)$, $Y_1^t, \dots, Y_k^t$ are independent and thus, $N(t)$ and Q^t are independent.

(e) Since for each $u \geqslant 0$ we have $Q^u(v) \stackrel{\mathscr{D}}{=} N(v)$ for all $v \geqslant 0$ it follows for all $0 \leqslant s \leqslant t$ from (b):

$$N(t) - N(s) = N(t-s+s) - N(s) = Q^s(t-s) \stackrel{\mathscr{D}}{=} N(t-s).$$

The result from (a) implies (iii) in Definition 1.3.1.

The independence of the increments is shown by induction: if $n=2$ then the result in part (d) guarantees that $Q^t(s) \stackrel{\mathscr{D}}{=} N(t+s) - N(t)$ is independent of $N(t)$). For $0 \leqslant t_1 \leqslant \ldots \leqslant t_{n+1}$ we have

$$\begin{aligned} &\Big(N(t_2) - N(t_1),\, \dots, N(t_{n+1}) - N(t_n)\Big) \\ &\qquad = \Big(Q^{t_1}(t_2 - t_1), Q^{t_1}(t_3 - t_1) - Q^{t_1}(t_2 - t_1), \dots, Q^{t_1}(t_{n+1} - t_1) - Q^{t_1}(t_n - t_1)\Big). \end{aligned}$$

Since Q^{t_1} is a Poisson process according to part (d) and since these are n increments, the induction hypothesis guarantees that these random variables are independent. Moreover, part (d) guarantees that Q^{t_1} is independent of $N(t_1)$ which implies that

$$N(t_1), N(t_2) - N(t_1),\, \dots, N(t_{n+1}) - N(t_n)$$

are independent, completing the induction and showing Condition (ii) in Definition 1.3.1.

For proofing that Definition 1.3.1 implies the definition in Example 1.1.2, define recursively $\sigma_0 := 0$ and

$$\sigma_k := \inf\{t > \sigma_{k-1} : N(t) \geqslant k\} \qquad \text{for } k \in \mathbb{N}.$$

Independent increments of N (Condition (ii)) and their Poisson distribution (Condition (iii)) imply for every $0 \leqslant a_1 < b_1 \leqslant \cdots \leqslant a_n < b_n$:

$$\begin{aligned}
&P\left(\bigcap_{k=1}^{n}\{a_k < \sigma_k \leqslant b_k\}\right) \\
&\quad = P\Bigg(\bigcap_{k=1}^{n-1}\{N(a_k) - N(b_{k-1}) = 0,\ N(b_k) - N(a_k) = 1\} \\
&\qquad\qquad\qquad \cap \{N(a_n) - N(b_{n-1}) = 0, N(b_n) - N(a_n) \geqslant 1\}\Bigg) \\
&\quad = e^{-\lambda(a_n - b_{n-1})}\big(1 - e^{-\lambda(b_n - a_n)}\big)\prod_{k=1}^{n-1} e^{-\lambda(a_k - b_{k-1})}\lambda(b_k - a_k)e^{-\lambda(b_k - a_k)} \\
&\quad = \big(e^{-\lambda a_n} - e^{-\lambda b_n}\big)\lambda^{n-1}\prod_{k=1}^{n-1}(b_k - a_k) \\
&\quad = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \lambda^n e^{-\lambda y_n}\, dy_n \ldots dy_1 \\
&\quad = \int_{a_1}^{b_1}\int_{a_2 - x_1}^{b_2 - x_1} \cdots \int_{a_n - y_1 - \cdots - y_{n-1}}^{b_n - y_1 - \cdots - y_{n-1}} \lambda^n e^{-\lambda(y_1 + \cdots + y_n)}\, dy_n \ldots dy_1.
\end{aligned}$$

Thus, the random variables $Y_k := \sigma_k - \sigma_{k-1}$ for $k \in \mathbb{N}$ obey

$$\begin{aligned}
P\left(\bigcap_{k=1}^{n}\{Y_k \in (a_k, b_k]\}\right) &= P\left(\bigcap_{k=1}^{n}\{\sigma_k \in (a_1 + \cdots + a_k, b_1 + \cdots + b_k]\}\right) \\
&= \int_{a_1}^{b_1}\int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} \lambda^n e^{-\lambda(y_1 + \cdots + y_n)}\, dy_n \ldots dy_1 \\
&= \int_{a_1}^{b_1} \lambda e^{-\lambda y_1}\, dy_1 \cdots \int_{a_n}^{b_n} \lambda e^{-\lambda y_n}\, dy_n,
\end{aligned}$$

which implies that $Y_1, \ldots, Y_n$ are independent and exponentially distributed with parameter λ. Since the definition of σ_k, $k \in \mathbb{N}$, implies for $t > 0$

$$\begin{aligned}
N(t) &= \max\{k \in \{0, 1, 2, \ldots\} : \sigma_k \leqslant t\} \\
&= \max\{k \in \{0, 1, 2, \ldots\} : Y_1 + \cdots + Y_k \leqslant t\},
\end{aligned}$$

it follows that N satisfies the Definition 1.1.2.

9. Let $t = 0$ corresponds to the time 8pm and let $(N(t) : t \geqslant 0)$ denote the Poisson process with rate $\lambda = 4$. According to the Definition in Example 1.1.2 we can assume that there exist independent, exponentially distributed random variables $X_1, X_2, \dots$ with parameter λ such that

$$N(t) := \begin{cases} 0, & \text{if } t = 0, \\ \max\{k \in \{0, 1, 2, \dots\} : X_1 + \dots + X_k \leqslant t\}, & \text{if } t > 0, \end{cases}$$

(a) The expected waiting time for the next bus is modelled by X_1. A simple calculation to obtain the expectation of an exponentially distributed random variable shows that $E[X_1] = \frac{1}{\lambda} = 15\text{min}$.

(b) Let Y model the waiting time for the next bus given that John has already waited for 30min. Then it follows from the definition of conditional probability for each $t \geqslant 0$

$$\begin{aligned} P(Y > t) &= P(X_1 > 0.5 + t | X_1 > 0.5) \\ &= P(X_1 > 0.5 + t)P(X > 0.5) = e^{-4(0.5+t)}e^{-4(0.5)} = e^{-4t}. \end{aligned}$$

Thus, also Y is exponentially distributed with the same parameter. It means that John is not in a better situation than at 8pm; in particular the expected waiting time is again $E[Y] = \frac{1}{\lambda} = 15\text{min}$.

This result is quite surprising but it can be explained. It is known as inspection paradox of renewal processes.

10. It is easy to check the conditions in Definition 1.3.1.

A.2. Solution Chapter 2

Some of the solutions use results from Chapter 3, such as the formula in Proposition 3.2.1.(c).

1. It follows from Definition 1.3.1 that for each $t \geqslant 0$ the random variable $N(t)$ has a Poisson distribution with parameter λt and therefore

$$E[N(t)] = \lambda t, \qquad E[N^2(t)] = \lambda t + \lambda^2 t^2. \tag{1.2.12}$$

The stationary increments of the Poisson process imply $E[(N(t) - N(s))^k] = E[(N(t - s))^k]$ for each $0 \leqslant s \leqslant t$ and $k \in \mathbb{N}$.

(a) We check the conditions in Definition 2.1.1:

(i) for each $t \geqslant 0$ the random variable $N(t) - \lambda t$ is $\mathscr{F}_t^N$-measurable since the function $t \mapsto \lambda t$ is deterministic.

(ii) $E\big[\,|N(t) - \lambda t|\,\big] \leqslant E\big[\,|N(t)|\,\big] + \lambda t \leqslant 2\lambda t$ for all $t \geqslant 0$.

(iii) for every $0 \leqslant s \leqslant t$ the independent increments and (1.2.12) imply

$$\begin{aligned} E[N(t) - \lambda t | \mathscr{F}_s^N] &= E[N(t) - N(s) | \mathscr{F}_s^N] + E[N(s) | \mathscr{F}_s^N] - \lambda t \\ &= E[N(t) - N(s)] + N(s) - \lambda t \\ &= (t-s)\lambda + N(s) - \lambda t \\ &= N(s) - \lambda s. \end{aligned}$$

(b) We check the conditions in Definition 2.1.1:

(i) for each $t \geqslant 0$ the random variable $(N(t) - \lambda t)^2 - \lambda t$ is $\mathscr{F}_t^N$-measurable since it is the image of the continuous function

$$f \colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = (x - \lambda t)^2 - \lambda t$$

applied to the $\mathscr{F}_t^N$-measurable random variable $N(t)$.

(ii) for every $t \geqslant 0$ equation (1.2.12) implies

$$E\Big[\Big|(N(t) - \lambda t)^2 - \lambda t\Big|\Big] \leqslant E\big[(N(t) - \lambda t)^2\big] + \lambda t = \lambda t + \lambda t < \infty.$$

(iii) ffor every $0 \leqslant s \leqslant t$ the independent and stationary increments and (1.2.12) imply

$$\begin{aligned} &E\big[(N(t) - \lambda t)^2 - \lambda t | \mathscr{F}_s^N\big] \\ &\quad = E\big[(N(t) - N(s) + N(s) - \lambda t)^2 - \lambda t | \mathscr{F}_s^N\big] \\ &\quad = E\big[(N(t) - N(s))^2 | \mathscr{F}_s^N\big] + 2E[(N(t) - N(s))(N(s) - \lambda t) | \mathscr{F}_s^N] \\ &\qquad + E\big[(N(s) - \lambda t)^2 | \mathscr{F}_s^N\big] - \lambda t \\ &\quad = E\big[(N(t) - N(s))^2\big] + 2(N(s) - \lambda t)E[(N(t) - N(s))] \\ &\qquad + (N(s) - \lambda t)^2 - \lambda t \\ &\quad = \lambda(t-s) + \lambda^2(t-s)^2 + 2(N(s) - \lambda t)\lambda(t-s) + (N(s) - \lambda t)^2 - \lambda t \\ &\quad = (N(s) - \lambda s)^2 - \lambda s. \end{aligned}$$

(c) Follows directly from part (a) and (b) as

$$\big(N(t) - \lambda t\big)^2 - N(t) = \big(N(t) - \lambda t\big)^2 - \lambda t - \big(N(t) - \lambda t\big)$$

for all $t \geqslant 0$.

2. By the very definition of conditional expectation the stochastic process $(Y(t) : t \geqslant 0)$ is adapted to $\{\mathscr{F}_t\}_{t \geqslant 0}$. Moreover, properties of conditional expectation yield for every $t \geqslant 0$

$$E[|Y(t)|] = E\big[\,|E[X|\mathscr{F}_t]|\,\big] \leqslant E\big[E[|X|\,|\mathscr{F}_t]\big] = E[|X|] < \infty.$$

The tower property of conditional expectation implies for every $0 \leqslant s \leqslant t$

$$E\big[Y(t)|\mathscr{F}_s\big] = E\big[E[X|\mathscr{F}_t]|\mathscr{F}_s\big] = E\big[X|\mathscr{F}_s\big] = Y(s).$$

3. (a) Yes, see Theorem 3.2.3.

(b) No, since for $0 \leqslant s < t$ we obtain

$$E[W^2(t)] = t \neq s = E[W^2(s)],$$

which contradicts the fact that martingale has constant expectation.

(c) We check the conditions in Definition 2.1.1 (see also Corollary 3.2.4):

(i) for each $t \geqslant 0$ the random variable $\exp\left(cW(t) - \frac{c^2}{2}t\right)$ is $\mathscr{F}_t^W$-measurable since it is the image of the continuous function

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = \exp\left(cx - \tfrac{c^2}{2}t\right),$$

applied to the $\mathscr{F}_t^W$-measurable random variable $W(t)$.

(ii) for all $t \geqslant 0$ we have

$$E\big[\,|X(t)|\,\big] = E\big[\exp(cW(t))\big]\exp(-\tfrac{1}{2}c^2 t) = \exp(\tfrac{1}{2}c^2 t)\exp(-\tfrac{1}{2}c^2 t) = 1 < \infty.$$

(iii) for every $0 \leqslant s \leqslant t$ the independent increments and Proposition 3.2.1.(c). imply

$$\begin{aligned}
E\left[e^{cW(t)-\frac{c^2}{2}t}|\mathscr{F}_s^W\right] &= X(s)E\left[e^{c(W(t)-W(s))}|\mathscr{F}_s^W\right]e^{-\frac{c^2}{2}(t-s)} \\
&= X(s)E\left[e^{c(W(t)-W(s))}\right]e^{-\frac{c^2}{2}(t-s)} \\
&= X(s)e^{\frac{1}{2}c^2(t-s)}e^{-\frac{c^2}{2}(t-s)} \\
&= X(s).
\end{aligned}$$

(d) We check the conditions in 2.1.1:

(i) for each $t \geqslant 0$ the random variable $W^3(t) - 3tW(t)$ is $\mathscr{F}_t^W$-measurable since it is the image of the continuous function

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) = x^3 - 3tx,$$

applied to the $\mathscr{F}_t^W$-measurable random variable $W(t)$.

(ii) Since $W(t)$ is normally distributed we obtain by Lemma 3.3.4 for all $t \geqslant 0$

$$E\big[\,|X(t)|\,\big] \leqslant E\left[|W(t)|^3\right] + 3tE\left[|W(t)|\right] < \infty.$$

(iii) for every $0 \leqslant s \leqslant t$ we obtain by using (a)

$$\begin{aligned}
&E\left[W^3(t) - 3tW(t)|\mathscr{F}_s^W\right] \\
&= E\left[\big(W(t) - W(s) + W(s)\big)^3|\mathscr{F}_s^W\right] - E[3tW(t)|\mathscr{F}_s^W] \\
&= E\left[(W(t) - W(s))^3|\mathscr{F}_s^W\right] + E\left[3(W(t) - W(s))^2 W(s)|\mathscr{F}_s^W\right] \\
&\quad + E\left[3(W(t) - W(s))W^2(s)|\mathscr{F}_s^W\right] + E\left[W^3(s)|\mathscr{F}_s^W\right] - 3tW(s) \\
&= 0 + 3(t-s)W(s) + 0 + W^3(s) - 3tW(s) \\
&= X(s).
\end{aligned}$$

(e) We check the conditions in Definition 2.1.1:

(i) for each $t \geqslant 0$ the random variable $t^2 W(t) - 2\int_0^t sW(s)\,ds$ is $\mathscr{F}_t^W$-measurable since it is the image of the continuous function

$$f\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}, \qquad f(x,y) = t^2 x - 2y,$$

applied to the $\mathscr{F}_t^W$-measurable random variables $W(t)$ and $\int_0^t sW(s)\,ds$. The latter is $\mathscr{F}_t^W$ measurable since it is the limit of $\mathscr{F}_t^W$ measurable random variables in the Riemann sum:

$$\int_0^t sW(s)\,ds = \lim_{|\pi_n|\to 0} \sum_{i=0}^{m_n-1} \zeta_i^{(n)} W\left(\zeta_i^{(n)}\right)\left(t_{i+1}^{(n)} - t_i^{(n)}\right),$$

where $\pi_n = \{t_i^{(n)}\}_{i=0,\dots,m_n}$ is a partition of $[0,t]$ for each $n\in\mathbb{N}$ and $\zeta_i^{(n)} \in [t_i^{(n)}, t_{i+1}^{(n)}]$.

(ii) for each $t \geqslant 0$ we obtain by using Fubini's theorem for the second term:

$$\begin{aligned} E\left[\left|t^2 W(t) - 2\int_0^t sW(s)\,ds\right|\right] &\leqslant t^2 E[|W(t)|] + 2\int_0^t sE[|W(s)|]\,ds \\ &\leqslant t^2\left(E[W^2(t)]\right)^{1/2} + 2\int_0^t s\left(E[W^2(s)]\right)^{1/2}\,ds \\ &< \infty. \end{aligned}$$

(iii) for each $0 \leqslant s \leqslant t$ Theorem 3.2.3 and Fubini's theorem imply

$$\begin{aligned} &E\left[t^2 W(t) - 2\int_0^t uW(u)\,du|\mathscr{F}_s^W\right] \\ &\quad = t^2 E[W(t)|\mathscr{F}_s^W] - 2\int_0^t E[uW(u)|\mathscr{F}_s^W]\,du \\ &\quad = t^2 W(s) - 2\int_0^s uE[W(u)|\mathscr{F}_s^W]\,du - 2\int_s^t uE[W(u)|\mathscr{F}_s^W]\,du \\ &\quad = t^2 W(s) - 2\int_0^s uW(u)\,du - 2\int_s^t uW(s)\,du \\ &\quad = s^2 W(s) - 2\int_0^s uW(u)\,du. \end{aligned}$$

We use in the second line above that the order of the conditional expectation and the integral can be exchanged (Fubini's theorem).

(f) No, since for $0 \leqslant s < t$ we obtain

$$E[W^4(t) - 4t^2 W(t)] = 3t^2 \neq 3s^2 = E[W^4(s) - 4s^2 W(s)],$$

which contradicts the fact that martingale has constant expectation.

4. (a) For $A := (-\infty, -a] \cup [a, \infty)$ the random time τ obeys the equation

$$\tau = \inf\{t \geqslant 0 : W(t) \in A\}.$$

Since A is a closed set in $\mathbb{R}$ and W has continuous path, Proposition 1.2.7 implies that τ is a stopping time

(b) Recall that $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$ for all $x \in \mathbb{R}$. We check the conditions in Definition 2.1.1:

(i) For each $t \geqslant 0$ the random variable $M(t)$ is $\mathscr{F}_t^W$-measurable since it is the image of the continuous function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \exp(-\frac{c^2}{2}t)\cosh(cx)$, applied to $W(t)$.

(ii) for all $t \geqslant 0$ we obtain

$$E\left[\left|e^{-\frac{c^2}{2}t}\cosh(cW(t))\right|\right] = \tfrac{1}{2}e^{-\frac{c^2}{2}t}E\left[e^{cW(t)} + e^{-cW(t)}\right] = e^{-\frac{c^2}{2}t}e^{\frac{c^2}{2}t} = 1 < \infty.$$

(iii) for each $0 \leqslant s \leqslant t$ it follows from Exercise 2.5.3.c

$$\begin{aligned}
E\left[e^{-\frac{c^2}{2}t}\cosh(cW(t))|\mathscr{F}_s^W\right] &= \frac{1}{2}\left(E\left[e^{-\frac{c^2}{2}t+cW(t)}|\mathscr{F}_s^W\right] + E\left[e^{-\frac{c^2}{2}t-cW(t)}|\mathscr{F}_s^W\right]\right) \\
&= \frac{1}{2}\left(e^{-\frac{c^2}{2}s+cW(s)} + e^{-\frac{c^2}{2}s-cW(s)}\right) \\
&= e^{-\frac{c^2}{2}s}\cosh(cW(s)).
\end{aligned}$$

(c) Since $(M(t) : t \geqslant 0)$ is a martingale according to (b), the optional sampling Theorem 2.3.1 implies

$$1 = E[M(0)] = E[M(t \wedge \tau)] \qquad \text{for all } t \geqslant 0. \tag{1.2.13}$$

Moreover, we have $M(t\wedge\tau) \to M(\tau)$ P-a.s. as $t \to \infty$ and it follows from the monotonicity and symmetry of the function $x \mapsto \cosh(x)$ that

$$\begin{aligned}
&|M(t \wedge \tau)| \leqslant |\cosh(cW(t \wedge \tau))| \leqslant \cosh(ca) \\
&E[|\cosh(ca)|] = \cosh(ca) < \infty.
\end{aligned}$$

Consequently, Lebesgue's dominated convergence theorem implies

$$E[M(\tau)] = E[\lim_{t\to\infty} M(t \wedge \tau)] = \lim_{t\to\infty} E[M(t \wedge \tau)] \ = 1, \tag{1.2.14}$$

where we use equation (1.2.13). On the other hand, the symmetry of $x \mapsto \cosh(x)$ implies

$$\begin{aligned}
E[M(\tau)] &= E\left[e^{-\frac{c^2}{2}\tau}\cosh(cW(\tau))\mathbb{1}_{W(\tau)=a}\right] + E\left[e^{-\frac{c^2}{2}\tau}\cosh(cW(\tau))\mathbb{1}_{W(\tau)=-a}\right] \\
&= E\left[e^{-\frac{c^2}{2}\tau}\mathbb{1}_{W(\tau)=a}\right]\cosh(ca) + E\left[e^{-\frac{c^2}{2}\tau}\mathbb{1}_{W(\tau)=-a}\right]\cosh(-ca) \\
&= E\left[e^{-\frac{c^2}{2}\tau}\right]\cosh(ca).
\end{aligned}$$

Together with equation (1.2.14) this results in

$$E\left[e^{-\lambda\tau}\right] = \Big(\cosh(\sqrt{2\lambda}\, a)\Big)^{-1}$$

for each $\lambda > 0$.

5. Define $c := b + \sqrt{b^2 + 2\lambda}$ and $M(t) := \exp(cW(t) - \frac{c^2}{2}t)$ for each $t \geqslant 0$. Since $(M(t) : t \geqslant 0)$ is a martingale according to Exercise 2.5.3.c, the optional sampling Theorem 2.3.1 implies

$$1 = E[M(0)] = E[M(t \wedge \tau)] \qquad \text{for all } t \geqslant 0. \tag{1.2.15}$$

Moreover, we have $M(t\wedge\tau) \to M(\tau)$ P-a.s. as $t \to \infty$ and

$$\begin{aligned} |M(t\wedge\tau)| &\leqslant \exp\big(c(a + b(t\wedge\tau)) - \tfrac{1}{2}c^2(t\wedge\tau)\big) \\ &= \exp(ca)\exp\big((cb - \tfrac{1}{2}c^2)(t\wedge\tau)\big) \\ &\leqslant \exp(ca), \end{aligned}$$

because $cb - \frac{1}{2}c^2 = -\lambda < 0$. Since

$$E[|\exp(ca)|] = \exp(ca) < \infty,$$

Lebesgue's theorem of dominated convergence implies

$$E[M(\tau)] = E[\lim_{t\to\infty} M(t\wedge\tau)] = \lim_{t\to\infty} E[M(t\wedge\tau)] = 1, \tag{1.2.16}$$

where we use Equation (1.2.15). On the other hand, a straightforward calculation yields

$$E[M(\tau)] = E\left[\exp\big(c(a + b\tau) - \tfrac{1}{2}c^2\tau\big)\right] = E\left[\exp\big((cb - \tfrac{1}{2}c^2)\tau\big)\right]\exp(ca).$$

Together with equation (1.2.16) this results in

$$E\left[e^{-\lambda\tau}\right] = \exp\Big(-a(b + \sqrt{b^2 + 2\lambda}\,)\Big)$$

for each $\lambda > 0$.

Can somebody solve this exercise for $b < 0$?

6. To show existence define $A_0 := 0$ and

$$A_k := \sum_{i=1}^{k} E[X_i - X_{i-1}|\mathscr{F}_{i-1}] \qquad \text{for all } k \in \mathbb{N}\,.$$

Since X is a submartingale, the sequence $(A_k : k \in \mathbb{N}_0)$ is increasing and by definition, each A_k is $\mathscr{F}_{k-1}$-measurable. Define $M_k := X_k - A_k$ for each $k \in \mathbb{N}_0$, which is adapted and satisfies $E[|M_k|] < \infty$. By using the definition of A_k we obtain for ever $k \in \mathbb{N}$

$$\begin{aligned} E[M_k|\mathscr{F}_{k-1}] &= E[X_k|\mathscr{F}_{k-1}] - A_k \\ &= E[X_k|\mathscr{F}_{k-1}] - E[X_k - X_{k-1}|\mathscr{F}_{k-1}] - A_{k-1} \\ &= X_{k-1} - A_{k-1} = M_{k-1}. \end{aligned}$$

The power property shows that $(M_k)_{k\in\mathbb{N}_0}$ is a martingale. To show uniqueness assume that there are M and A with the same properties satisfying

$$X_k = M_k + A_k = M_k' + A_k' \qquad \text{for all } k \in \mathbb{N}_0\,.$$

It follows that $M_k - M_k' = A_k - A_k'$ is an $\mathscr{F}_{k-1}$-measurable random variable which implies

$$M_k - M_k' = E[M_k - M_k' | \mathscr{F}_{k-1}] = M_{k-1} - M_{k-1}'.$$

It follows recursively

$$M_k - M_k' = M_{k-1} - M_{k-1}' = \cdots = M_0 - M_0' = A_0 - A_0' = 0,$$

which shows uniqueness.

7. This exercise is an example of the use of the following result from probability theory:

   Lemma Let $\mathscr{B}$ be a sub-σ-algebra of $\mathscr{A}$. Let X be a random variable which is $\mathscr{B}$-measurable and let Y be a random variable which is independent of $\mathscr{B}$. If $f : \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is a function with $E[|f(X,Y)|] < \infty$ then

$$E[f(X,Y)|\mathscr{B}] = h(X) \qquad P\text{-a.s.},$$

   where $h : \mathbb{R} \to \mathbb{R}$ is a function defined by

$$h(x) := E[f(x,Y)].$$

   The payoff of the European call option is giving by $g(S(T))$ where $g(x) = (x-K)^+$ for all $x \in \mathbb{R}$. The given representation of S yields that

$$S(t) = S(u)\exp\left(\left(r - \tfrac{1}{2}\sigma^2\right)(t-u) + \sigma(W(t) - W(u))\right)$$

   for all $0 \leqslant u \leqslant t$. Applying this representation for $t = T$ and $u = t$ we obtain by the lemma that

$$\begin{aligned} V(t) &= E\left[e^{-r(T-t)} g\left(\underbrace{S(t)}_{\mathscr{F}_t\text{-meas.}}\ \exp\left(\left(r-\tfrac{1}{2}\sigma^2\right)(T-t) + \sigma\underbrace{(W(T)-W(t))}_{\text{ind. of } \mathscr{F}_t}\right)\right) |\mathscr{F}_t\right] \\ &= h(t, S(t)) \end{aligned}$$

   for a function $h : [0,T]\times\mathbb{R}\to\mathbb{R}$ where

$$\begin{aligned} h(t,x) &= E\left[e^{-r(T-t)} g\left(x\exp\left(\left(r-\tfrac{1}{2}\sigma^2\right)(T-t) + \sigma(W(T)-W(t))\right)\right)\right] \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-r(T-t)} g\left(x\exp\left(\sigma\sqrt{T-t}\,y + \left(r-\tfrac{1}{2}\sigma^2\right)\right)(T-t)\right) e^{-\frac{1}{2}y^2}\,dy. \end{aligned}$$

Using the definition $g(x) = (x-K)^+$ we can continue and evaluate the integral to obtain:

$$
\begin{aligned}
&h(t,x)\\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-r(T-t)} \left(x \exp\left(\left(r - \tfrac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}\, y \right) - K \right)^+ e^{-\frac{1}{2}y^2}\, dy\\
&\vdots \quad \vdots\\
&= xF_N(d_+(T-t,x)) - Ke^{-r(T-t)} F_N(d_-(T-t,x)),
\end{aligned}
$$

where F_N denotes the distribution function of the standard normal distribution and the constants d_+ and d_- are defined by

$$
d_\pm(\tau,x) := \frac{1}{\sigma\sqrt{\tau}} \left(\ln \frac{x}{K} + (r \pm \tfrac{\sigma^2}{2})\tau \right).
$$

Remark for students in financial mathematics: in risk-neutral valuation you do this calculation under the risk-neutral measure Q, i.e. the expectation here is with respect to Q.

A.3. Solution Chapter 3

1. (a) Define $\widetilde{\mathrm{W}}(t) := -W(t)$ for all $t \geqslant 0$. We show the conditions in Definition 3.0.1:

   (i) $\widetilde{\mathrm{W}}(0) = -W(0) = 0$ P-a.s.

   (ii) Note, that if X and Y are independent random variables then $f(X)$ and $g(Y)$ are also independent for any measurable functions $f, g : \mathbb{R} \to \mathbb{R}$. By applying this together with the independent increments of W it follows by induction that for every $0 \leqslant t_1 \leqslant \ldots \leqslant t_n$ and $n \in \mathbb{N}$ the random variables

   $$
   -\big(W(t_2) - W(t_1)\big), \ldots, -\big(W(t_n) - W(t_{n-1})\big)
   $$

   are independent. Since

   $$
   \begin{aligned}
   &\widetilde{\mathrm{W}}(t_2) - \widetilde{\mathrm{W}}(t_1), \ldots, \widetilde{\mathrm{W}}(t_n) - \widetilde{\mathrm{W}}(t_{n-1})\\
   &\quad = -\big(W(t_2) - W(t_1)\big), \ldots, -\big(W(t_n) - W(t_{n-1})\big)
   \end{aligned}
   $$

   it follows that $\widetilde{\mathrm{W}}$ also has independent increments.

   (iii) For every $0 \leqslant s \leqslant t$ the symmetry of the standard normal distribution implies

   $$
   \widetilde{\mathrm{W}}(t) - \widetilde{\mathrm{W}}(s) = -(W(t) - W(s)) \stackrel{\mathscr{D}}{=} N(0, t-s).
   $$

   (iv) Since the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := -x$ is continuous, continuity of W implies that the mapping $t \mapsto f(W(t))$ is continuous, too.

(b) Define $\widetilde{W}(t) := W(t+t_0) - W(t_0)$ for all $t \geqslant 0$. We show the conditions in Definition 3.0.1:

(i) $\widetilde{W}(0) = W(t_0) - W(t_0) = 0$ P-a.s.

(ii) For every $0 \leqslant t_1 \leqslant \ldots \leqslant t_n$ and $n \in \mathbb{N}$ we obtain

$$\begin{aligned}&\widetilde{W}(t_2) - \widetilde{W}(t_1), \ldots, \widetilde{W}(t_n) - \widetilde{W}(t_{n-1})\\ &\quad = W(t_2+t_0) - W(t_1+t_0), \ldots, W(t_n+t_0) - W(t_{n-1}+t_0)\end{aligned}$$

are independent, since W has independent increments and $t_1 + t_0 \leqslant t_2 + t_0 \leqslant \ldots \leqslant t_n + t_0$.

(iii) For every $0 \leqslant s \leqslant t$ we obtain

$$\widetilde{W}(t) - \widetilde{W}(s) = W(t+t_0) - W(s+t_0)) \stackrel{\mathscr{D}}{=} N(0, t+t_0-(s+t_0)) = N(0, t-s).$$

(iv) Clearly, the stochastic process $(W(t+t_0) : t \geqslant 0)$ has continuous trajectories. Since $\widetilde{W}(t) = W(t+t_0) - W(t_0)$ and $W(t_0)$ does not depend on t, also $\widetilde{W}$ has continuous trajectories.

(c) Define $\widetilde{W}(t) := cW(\frac{t}{c^2})$ for all $t \geqslant 0$. We show the conditions in Definition 3.0.1:

(i) $\widetilde{W}(0) = W(0) = 0$ P-a.s.

(ii) For every $0 \leqslant t_1 \leqslant \ldots \leqslant t_n$ and $n \in \mathbb{N}$ the independent increments of W guarantee that

$$\left(W\left(\frac{t_2}{c^2}\right) - W\left(\frac{t_1}{c^2}\right)\right), \ldots, \left(W\left(\frac{t_n}{c^2}\right) - W\left(\frac{t_{n-1}}{c^2}\right)\right)$$

are independent, since $\frac{t_1}{c^2} \leqslant \ldots \leqslant \frac{t_n}{c^2}$. As multiplication by a constant preserves independence (see part (a)), we obtain

$$\begin{aligned}&\widetilde{W}(t_2) - \widetilde{W}(t_1), \ldots, \widetilde{W}(t_n) - \widetilde{W}(t_{n-1})\\ &\quad = c\left(W\left(\frac{t_2}{c^2}\right) - W\left(\frac{t_1}{c^2}\right)\right), \ldots, c\left(W\left(\frac{t_n}{c^2}\right) - W\left(\frac{t_{n-1}}{c^2}\right)\right)\end{aligned}$$

are also independent.

(iii) For every $0 \leqslant s \leqslant t$ we obtain

$$\widetilde{W}(t) - \widetilde{W}(s) = c\left(W\left(\frac{t}{c^2}\right) - W\left(\frac{s}{c^2}\right)\right) \stackrel{\mathscr{D}}{=} N\left(0, c^2\left(\frac{t}{c^2} - \frac{s}{c^2}\right)\right) = N(0, t-s).$$

(iv) Clearly, the stochastic process $(W(\frac{t}{c^2}) : t \geqslant 0)$ has continuous trajectories. Since $\widetilde{W}(t) = cW(\frac{t}{c^2})$, also $\widetilde{W}$ has continuous trajectories.

2. Let X be normally distributed with expectation 0 and variance $\sigma^2 > 0$. For $k \in \mathbb{N}$ we obtain by partial integration

$$\begin{aligned}
E[X^k] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^k e^{-\frac{1}{2\sigma^2}x^2}\, dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \left(\left[\frac{1}{k+1} x^{k+1} e^{-\frac{1}{2\sigma^2}x^2} \right]_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{1}{k+1} x^{k+1} \frac{-x}{\sigma^2} e^{-\frac{1}{2\sigma^2}x^2}\, dx \right) \\
&= \frac{1}{(k+1)\sigma^2} E[X^{k+2}]. \qquad (1.3.17)
\end{aligned}$$

We show by induction that

$$E[X^k] = \begin{cases} 0, & \text{if } k \text{ is odd,} \\ \frac{k!}{2^{k/2}(k/2)!}\sigma^k, & \text{if } k \text{ is even.} \end{cases} \qquad (1.3.18)$$

Proof: If k is odd it follows from (1.3.17) that

$$0 = E[X] = E[X^3] = \cdots = E[X^k].$$

Let k be even. For $k = 2$ we know that $E[X^2] = \sigma^2$ and thus, the claim is satisfied. For the induction, we assume that the claim is true for k and show that this implies the claim for $k + 2$:

$$E[X^{k+2}] = (k+1)\sigma^2 E[X^k] = (k+1)\sigma^2 \frac{k!}{2^{k/2}(k/2)!}\sigma^k = \frac{(k+2)!}{2^{(k+2)/2}((k+2)/2)!}\sigma^{k+2}.$$

Thus, the proof of formula (1.3.18) is completed.

3. Fix $0 \leqslant t_1 \leqslant \cdots \leqslant t_n$ and note that the n-dimensional random vector

$$\big(W(t_1), W(t_2) - W(t_1), \dots, W(t_n) - W(t_{n-1})\big)$$

is normally distributed as the entries of this random vector are independent and normally distributed. Since

$$\begin{pmatrix} W(t_1) \\ W(t_2) \\ \vdots \\ W(t_n) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ 1 & 1 & 1 & \dots & 1 & 1 \end{pmatrix} \begin{pmatrix} W(t_1) \\ W(t_2) - W(t_1) \\ \vdots \\ W(t_n) - W(t_{n-1}) \end{pmatrix},$$

it follows that $(W(t_1), \dots, W(t_n))$ is normally distributed. Thus, the stochastic process $(W(t) : t \geqslant 0)$ is Gaussian.

4. By properties of the Brownian motion in Proposition 3.2.1 we obtain for each $0 \leqslant s \leqslant t$:

$$\begin{aligned}
E[B(t)] &= E[W(t)] - tE[W(1)] = 0, \\
\operatorname{Cov}(B(s)B(t)) &= E\big[\big(W(s) - sW(1)\big)\big(W(t) - tW(1)\big)\big] \\
&= E\big[W(s)W(t)\big] - tE\big[W(s)W(1)\big] - sE\big[W(1)W(t)\big] + stE\big[W^2(1)\big] \\
&= s - st.
\end{aligned}$$

Fix $0 \leqslant t_1 \leqslant \dots \leqslant t_n$ and note that the $(n+1)$-dimensional random vector

$$\big(W(t_1), W(t_2), \dots, W(t_n), W(1)\big)$$

is normally distributed by Exercise 3.4.3. Since

$$\begin{pmatrix} W(t_1) - t_1 W(1) \\ W(t_2) - t_2 W(1) \\ \vdots \\ W(t_n) - t_n W(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & -t_1 \\ 0 & 1 & 0 & \dots & 0 & -t_2 \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -t_{n-1} \\ 0 & 0 & 0 & \dots & 1 & -t_n \end{pmatrix} \begin{pmatrix} W(t_1) \\ W(t_2) \\ \vdots \\ W(t_n) \\ W(1) \end{pmatrix},$$

it follows that $(W(t_1) - t_1 W(1), \dots, W(t_n) - t_n W(1))$ is normally distributed. Thus, the Brownian bridge $(B(t) : t \geqslant 0)$ is a Gaussian process.

5. In this exercise we apply Fubini's theorem, that is the interchange of the order of integration (including expectation), without showing that the necessary conditions are satisfied. A rigorous mathematical proof would require that, of course!

   (a) For every $t \geqslant 0$ we obtain

$$E[Y(t)] = \int_0^t E[W(s)f(s)]\,ds = \int_0^t E[W(s)]\,f(s)\,ds = 0.$$

Using the formula for the covariance of Browian motion we obtain for each $0 \leqslant s \leqslant t$ that

$$\begin{aligned} \operatorname{Cov}(Y(s)Y(t)) &= E[Y(s)Y(t)] \\ &= E\left[\int_0^s \int_0^t W(r)f(r)W(u)f(u)\,du\,dr\right] \\ &= \int_0^s \int_0^t E[W(r)W(u)]f(r)f(u)\,du\,dr \\ &= \int_0^s \left(\int_0^r E[W(r)W(u)]f(r)f(u)\,du + \int_r^t E[W(r)W(u)]f(r)f(u)\,du\right) dr \\ &= \int_0^s \left(\int_0^r uf(r)f(u)\,du + \int_r^t rf(r)f(u)\,du\right) dr \\ &= \int_0^s \left(\int_0^r uf(r)f(u)\,du\right) dr + \int_0^s \left(\int_r^t rf(r)f(u)\,du\right) dr. \end{aligned}$$

Changing the order of integration in the first integral yields

$$\begin{aligned} &= \int_0^s \left(\int_u^s uf(r)f(u)\,dr\right) du + \int_0^s \left(\int_r^t rf(r)f(u)\,du\right) dr \\ &= \int_0^s uf(u) \left(\int_u^s f(r)\,dr\right) du + \int_0^s rf(r) \left(\int_r^t f(u)\,du\right) dr. \end{aligned}$$

Redefine $u \to r$ and $r \to u$ in the first integral to obtain

$$= \int_0^s r f(r) \left(\int_r^s f(u)\, du \right) dr + \int_0^s r f(r) \left(\int_r^t f(u)\, du \right) dr$$
$$= \int_0^s r f(r) \left(2 \int_r^s f(u)\, du + \int_s^t f(u)\, du \right) dr.$$

(b) By the definition of Riemann integrals we have for every $\omega \in \Omega$ that

$$\begin{aligned} Y(t)(\omega) &= \int_0^t W(s)(\omega) f(s)\, ds \\ &= \lim_{|\pi_n| \to 0} \sum_{i=0}^{m_n - 1} W\left(\zeta_i^{(n)}\right)(\omega)\, f\left(\zeta_i^{(n)}\right) \left(t_{i+1}^{(n)} - t_i^{(n)}\right), \end{aligned} \tag{1.3.19}$$

where $\pi_n = \{t_i^{(n)}\}_{i=0,\dots,m_n}$ is a partition of $[0,t]$ for each $n \in \mathbb{N}$ and $\zeta_i^{(n)} \in [t_i^{(n)}, t_{i+1}^{(n)}]$. Since $(W(t) : t \geqslant 0)$ is a Gaussian process according to Exercise 3.4.3, it follows that the sum

$$\sum_{i=0}^{m_n - 1} W\left(\zeta_i^{(n)}\right) f\left(\zeta_i^{(n)}\right) \left(t_{i+1}^{(n)} - t_i^{(n)}\right)$$

is a normally distributed random variable for every partition π_n. Since the P-a.s. limit of Gaussian random variables is again normally distributed[2], it follows that the left hand side in (1.3.19) is normally distributed.

6. We assume[3] that the probability distribution of X has a density f.

(a) Since X is assumed to be a non-negative random variable, the density f has only support in $\mathbb{R}_+$ and we obtain by Fubini's theorem that

$$E[X] = \int_{\mathbb{R}} y f(y)\, dy = \int_0^\infty f(y) \int_0^y dx\, dy = \int_0^\infty \int_x^\infty f(y)\, dy\, dx = \int_0^\infty P(X \geqslant x)\, dx.$$

(b) Assume for a contradiction that there is $\delta > 0$ such that $P(|X| = \infty) = \delta$. It follows for every constant $K > 0$ that

$$E[|X|] \geqslant E[|X|\, \mathbb{1}_{|X|>K}] \geqslant K E[\mathbb{1}_{|X|>K}] = K P(|X| > K) \geqslant K\delta.$$

Since K is arbitrary we can take $K \to \infty$ and it follows that $E[|X|] = \infty$ which is a contradiction.

Measure theoretical proof: define $B := \{X = \infty\}$ and assume that $P(B) > 0$ for contradiction. Then

$$E[|X|] = \int_\Omega |X(\omega)|\; P(d\omega) \geqslant \int_B |X(\omega)|\; P(d\omega) = \infty \cdot P(B) = \infty,$$

which is a contradiction.

[2] this is not given in the lecture notes, but it can be read out of the characteristic function

[3] if you know measure theory, then replace $f(y)\, dy$ by $P(dy)$ in the following

7. (a) Since $(W^H(t) : t \geqslant 1)$ is a Gaussian process, it follows for every $s, t \geqslant 0$ that the random vector $(W^H(s), W^H(t))$ is normally distributed, and, by its definition, with expectation 0 and covariance matrix

$$\begin{pmatrix} r(s,s) & r(s,t) \\ r(t,s) & r(t,t) \end{pmatrix}.$$

It follows for $0 \leqslant s \leqslant t$ that $W^H(t) - W^H(s)$ is normally distributed with expectation 0 and

$$\begin{aligned} \mathrm{Var}[W^H(t) - W^H(s)] &= E[(W^H(t) - W^H(s))^2] \\ &= E[(W^H(t))^2] + E[(W^H(s))^2] - 2E[W^H(t)W^H(s)] \\ &= r(t,t) + r(s,s) - 2r(t,s) \\ &= |t-s|^{2H} \\ &= r(t-s, t-s). \end{aligned}$$

Consequently, the random variables $W^H(t) - W^H(s)$ and $W^H(t-s)$ are both normally distributed with the same expectations and variances, which implies in the case of a Gaussian distribution that they are distributed according to the same distribution.

(b) For $0 \leqslant s \leqslant t$ and $n \in \mathbb{N}$ it follows from part (a) and Lemma 3.3.4 that

$$E\left[|W^H(t) - W^H(s)|^{2n}\right] = |t-s|^{2Hn} E[X^{2n}] = \frac{(2k)!}{2^k k!} |t-s|^{2Hn},$$

where $X \stackrel{\mathscr{D}}{=} N(0,1)$. Consequently, for each $n \in \mathbb{N}$ we can choose $\alpha = 2n$ and $\beta = 2Hn - 1$ in Kolmogorov's Theorem 3.3.3 which implies that W^H is Hölder continuous of each order smaller than $H - \frac{1}{2n}$. Since n is arbitrary the claim follows.

8. (a) Let $\{t_k\}_{k=0,\dots n}$ be a partition of $[a,b]$. Since f is Hölder continuous of order 1 there exists a constant $c > 0$ such that

$$|f(x) - f(y)| \leqslant c\,|x-y| \qquad \text{for all } x, y \in [a,b].$$

Consequently, we obtain that

$$\sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)| \leqslant \sum_{k=1}^{n} c\,|t_{k+1} - t_k| = c(b-a),$$

and therefore, f is of finite variation.

(b) Let $n \in \mathbb{N}$ be even and define

$$\begin{aligned} &t_0 = 0, \qquad t_n = 1, \\ &t_k = \left((n-1-k)\pi + \tfrac{\pi}{2}\right)^{-1} \qquad \text{for } k = 1, \dots, n-1. \end{aligned}$$

Consequently, $\{t_k\}_{k=0,\dots,n}$ is a partition of $[0,1]$ and

$$f(t_k) = \begin{cases} -t_k & \text{if } k \text{ even}, \\ t_k & \text{if } k \text{ odd}, \end{cases} \qquad \text{for } k = 0, \dots, n-1.$$

It follows that

$$\begin{aligned} \sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)| &\geqslant \sum_{k=0}^{n-2} |t_{k+1} + t_k| \\ &= 2 \sum_{k=1}^{n-2} t_k + t_{n-1} + t_0 \\ &\geqslant 2 \sum_{k=1}^{n-2} (k\pi + \tfrac{\pi}{2})^{-1} \\ &\to \infty \qquad \text{as } n \to \infty. \end{aligned}$$

Consequently, f is not of finite variation. But f is continuous in $x = 0$, since

$$-x \leqslant f(x) \leqslant x \qquad \text{for all } x \in (0,1].$$

A.4. Solution Chapter 4

1. (a) $(\Phi(t) : t \in [0,4])$ is adapted, since

for $t \in [0,1]$: $\Phi(t) = 2$ and thus $\Phi(t)$ is $\mathscr{F}_t^W$-measurable;
for $t \in (1,2]$: $\Phi(t) = Y_1$ and thus $\Phi(t)$ is $\mathscr{F}_t^W$-measurable as Y_1 is $\mathscr{F}_1^W$-measurable;
for $t \in (2,3]$: $\Phi(t) = Y_2$ and thus $\Phi(t)$ is $\mathscr{F}_t^W$-measurable as Y_2 is $\mathscr{F}_2^W$-measurable;
for $t \in (3,4]$: $\Phi(t) = 0$ and thus $\Phi(t)$ is $\mathscr{F}_t^W$-measurable.

(b) The linearity of the stochastic integral yields

$$\begin{aligned} \int_0^4 \Phi(s)\, dW(s) &= \int_0^1 \Phi(s)\, dW(s) + \int_1^2 \Phi(s)\, dW(s) + \int_2^3 \Phi(s)\, dW(s) + \int_3^4 \Phi(s)\, dW(s) \\ &= \int_0^1 2\, dW(s) + \int_1^2 Y_1\, dW(s) + \int_2^3 Y_2\, dW(s) + \int_3^4 0\, dW(s) \\ &= 2W(1) + Y_1\big(W(2) - W(1)\big) + Y_2\big(W(3) - W(2)\big). \end{aligned}$$

(c) The process $(\Phi(t) : t \in [0,4])$ is in $\mathscr{H}$ since

$$\begin{aligned}
&\int_0^4 E\left[|\Phi(s)|^2\right] ds \\
&\quad = \int_0^1 E\left[|\Phi(s)|^2\right] ds + \int_1^2 E\left[|\Phi(s)|^2\right] ds + \int_2^3 E\left[|\Phi(s)|^2\right] ds + \int_3^4 E\left[|\Phi(s)|^2\right] ds \\
&\quad = 4 + E\left[|Y_1|^2\right] + E\left[|Y_2|^2\right] \\
&\quad = 4 + 2 + 4 = 10 < \infty.
\end{aligned}$$

Thus, Theorem 4.3.1 implies that the mean and expectation exist and that they are given by

$$\begin{aligned}
&E\left[\int_0^T \Phi(s)\, dW(s)\right] = 0, \\
&\operatorname{Var}\left[\int_0^T \Phi(s)\, dW(s)\right] = E\left[\left|\int_0^T \Phi(s)\, dW(s)\right|^2\right] = \int_0^T E\left[|\Phi(s)|^2\right] ds = 10.
\end{aligned}$$

2. (a) Let $T \geqslant b$. The stochastic process $(|W(t)| : t \in [0,T])$ is in $\mathscr{H}$ since

$$\int_0^T E\left[|W(t)|^2\right] dt = \int_0^T t\, dt < \infty.$$

It follows from Theorem 4.3.1 that the mean and variance exists. For the expectation we compute

$$E\left[\int_a^b |W(t)|\, dW(t)\right] = E\left[\int_0^b |W(t)|\, dW(t)\right] - E\left[\int_0^a |W(t)|\, dW(t)\right] = 0.$$

For the variance Itô's isometry (4.3.11) implies

$$\begin{aligned}
\operatorname{Var}\left[\int_a^b |W(t)|\, dW(t)\right] &= E\left[\left|\int_a^b |W(t)|\, dW(t)\right|^2\right] \\
&= \int_a^b E[|W(t)|^2]\, dt = \int_a^b t\, dt = \frac{1}{2}(b^2 - a^2).
\end{aligned}$$

(b) Let $T \geqslant b$. The stochastic process $(\sqrt{t}\, \exp(W(t)) : t \in [0,T])$ is in $\mathscr{H}$ since Proposition 3.2.1.(c) enables us to conclude

$$\int_0^T E\left[\left|\sqrt{t}\, \exp(W(t))\right|^2\right] dt = \int_0^T t e^{2t}\, dt < \infty.$$

It follows from Theorem 4.3.1 that the mean and variance exists. For the expectation we compute

$$E\left[\int_a^b \sqrt{t}\ \exp(W(t))\,dW(t)\right] = 0.$$

For the variance Itô's isometry (4.3.11) implies

$$\begin{aligned}\operatorname{Var}\left[\int_a^b \sqrt{t}\ \exp(W(t))\,dW(t)\right] &= E\left[\left|\int_a^b \sqrt{t}\ \exp(W(t))\,dW(t)\right|^2\right] \\ &= \int_a^b E\left[\left|\sqrt{t}\ \exp(W(t))\right|^2\right]\,dt = \int_a^b te^{2t}\,dt = \tfrac{1}{2}(b-\tfrac{1}{2})e^{2b} - \tfrac{1}{2}(a-\tfrac{1}{2})e^{2a}.\end{aligned}$$

(c) Let $T \geqslant b$. The stochastic process $(\operatorname{sgn}(W(t)) : t \in [0,T])$ is in $\mathscr{H}$ since

$$\int_0^T E\left[|\operatorname{sgn}(W(t))|^2\right]\,dt = \int_0^T 1\,dt = T < \infty.$$

It follows from Theorem 4.3.1 that that the mean and variance exists. For the expectation we compute as before

$$E\left[\int_a^b \operatorname{sgn}(W(t))\,dW(t)\right] = 0$$

For the variance Itô's isometry (4.3.11) implies

$$\begin{aligned}\operatorname{Var}\left[\int_a^b \operatorname{sgn}(W(t))\,dW(t)\right] &= E\left[\left|\int_a^b \operatorname{sgn}(W(t))\,dW(t)\right|^2\right] \\ &= \int_a^b E[|\operatorname{sgn}(W(t))|^2]\,dt = \int_a^b 1\,dt = b-a.\end{aligned}$$

3. For a first way see Exercise 2c in Section 2.5.3. For a second way define the function

$$f\colon [0,T]\times\mathbb{R}\to\mathbb{R}, \qquad f(t,x) := e^{x-\frac{1}{2}t}.$$

The function f is in $C^{1,2}$ and we obtain

$$f_t(t,x) = -\tfrac{1}{2}e^{x-\frac{1}{2}t}, \quad f_x(t,x) = e^{x-\frac{1}{2}t}, \quad f_{xx}(t,x) = e^{x-\frac{1}{2}t}.$$

Itô's formula in Theorem 4.5.6 implies

$$\begin{aligned}df(t,W(t)) &= \left(f_t(t,W(t)) + \tfrac{1}{2}f_{xx}(t,W(t))\right)\,dt + f_x(t,W(t))\,dW(t) \\ &= \left(-\tfrac{1}{2}e^{W(t)-\frac{1}{2}t} + \tfrac{1}{2}e^{W(t)-\frac{1}{2}t}\right)\,dt + e^{W(t)-\frac{1}{2}t}\,dW(t) \\ &= e^{W(t)-\frac{1}{2}t}\,dW(t).\end{aligned}$$

Since $f(0, W(0)) = 1$ this is

$$e^{W(t)-\frac{1}{2}t} = 1 + \int_0^t e^{W(s)-\frac{1}{2}s}\, dW(s) \qquad \text{for all } t \in [0,T]. \tag{1.4.20}$$

The stochastic process $\big(\exp(W(t) - \frac{1}{2}t) : t \in [0,T]\big)$ is in $\mathscr{H}$ since Proposition 3.2.1.(c) enables us to conclude

$$\int_0^T E\left[\left|\exp(W(t) - \tfrac{1}{2}t)\right|^2\right] dt = \int_0^T e^{-t}e^{2t}\, dt = \int_0^T e^t\, dt < \infty.$$

Consequently, Theorem 4.3.1 guarantees that $\big(\int_0^t \exp(W(u) - \frac{1}{2}u)\, dW(u) : t \in [0,T]\big)$ is a martingale which completes the exercise by equation (1.4.20).

4. (a) It follows directly from the definition of the sums $R(\pi)$ and $L(\pi)$ that

$$R(\pi) - L(\pi) = \sum_{k=0}^{m-1} \left(f(t_{k+1}) - f(t_k)\right)^2,$$
$$R(\pi) + L(\pi) = \sum_{k=0}^{m-1} \left(f(t_{k+1})^2 - f(t_k)^2\right) = f^2(T) - f^2(0).$$

By solving for $R(\pi)$ and $L(\pi)$ we obtain

$$L(\pi) = \frac{1}{2}\left(f^2(T) - f^2(0) - \sum_{k=0}^{m-1} \left(f(t_{k+1}) - f(t_k)\right)^2\right),$$
$$R(\pi) = \frac{1}{2}\left(f^2(T) - f^2(0) + \sum_{k=0}^{m-1} \left(f(t_{k+1}) - f(t_k)\right)^2\right).$$

(b) It follows from the Definition of the Riemann-Stieltjes integral that a necessary requirement for the existence of the Riemann-Stieltjes integral $\int_0^T f(s)\, df(s)$ is

$$\lim_{|\pi_n|\to 0} R(\pi_n) = \lim_{|\pi_n|\to 0} L(\pi_n)$$

for every sequence $\{\pi_n\}_{n\in\mathbb{N}}$ of partitions $\pi_n = \{t_k^{(n)}\}_{k=0,\dots,m_n}$ in $P[0,T]$. By part (a) this is satisfied if and only if

$$\lim_{|\pi_n|\to 0} \sum_{k=0}^{m_n-1} \left|f(t_{k+1}^{(n)}) - f(t_k^{(n)})\right|^2 = 0.$$

5. Define the function

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) := x^m.$$

The function f is in C^2 and we obtain

$$f_x(x) = mx^{m-1}, \quad f_{xx}(x) = m(m-1)x^{m-2}.$$

Itô's formula in Theorem 4.5.1 implies

$$\begin{aligned}(W(t))^m &= f(W(t)) \\ &= f(0) + \int_0^t f_x(W(s))\,dW(s) + \frac{1}{2}\int_0^t f_{xx}(W(s))\,ds \\ &= m\int_0^t W^{m-1}(s)\,dW(s) + \frac{1}{2}m(m-1)\int_0^t W^{m-2}(s)\,ds.\end{aligned}$$

6. Let $(H(u) : u \in [0,T])$ be a simple stochastic process in $\mathscr{H}_0$ of the form

$$H(u)(\omega) = \sum_{k=0}^{m-1} \mathbb{1}_{(u_k,u_{k+1}]}(u)X_k(\omega) \qquad \text{for all } u \in [0,T],\ \omega \in \Omega,$$

where $X_k : \Omega \to \mathbb{R}$ are $\mathscr{F}_{u_k}$-measurable random variables with $E[|X_k|^2] < \infty$ and $0 = u_0 < u_1 < \cdots < u_m = T$. For the given $0 \leqslant s \leqslant t$ we can assume without loss of generality that there exist $k_0, m_0 \in \{0,\dots,m\}$ such that $u_{k_0} = s$ and $u_{m_0} = t$.

(a) The linearity of the stochastic integral yields

$$\int_s^t H(u)\,dW(u) = \int_0^t H(u)\,dW(u) - \int_0^s H(u)\,dW(u) = \sum_{k=k_0}^{m_0-1} X_k\big(W(u_{k+1}) - W(u_k)\big).$$

By using properties of conditional expectation and the fact that the Brownian motion W is a martingale, we obtain

$$\begin{aligned}E\left[\int_s^t H(u)\,dW(u)\middle|\mathscr{F}_s\right] &= \sum_{k=k_0}^{m_0-1} E\Big[X_k(W(u_{k+1}) - W(u_k))\Big|\mathscr{F}_s\Big] \\ &= \sum_{k=k_0}^{m-1} E\Big[X_k E[W(u_{k+1}) - W(u_k)|\mathscr{F}_{u_k}]\Big|\mathscr{F}_s\Big] = 0.\end{aligned}$$

(b) For every $0 \leqslant i < j \leqslant m$ properties of conditional expectation and the fact that the Brownian motion W yield

$$\begin{aligned}&E\left[X_iX_j\big(W(u_{i+1}) - W(u_i)\big)\big(W(u_{j+1}) - W(u_j)\big)\big|\mathscr{F}_s\right] \\ &\quad = E\left[X_iX_j\big(W(u_{i+1}) - W(u_i)\big)E\big[(W(u_{j+1}) - W(u_j))|\mathscr{F}_{u_j}\big]\big|\mathscr{F}_s\right] = 0.\end{aligned}$$

It follows that the mixed terms vanish in the following calculation:

$$\begin{aligned}
&E\left[\left(\int_s^t H(u)\,dW(u)\right)^2\middle|\mathscr{F}_s\right]\\
&\quad=\sum_{k=k_0}^{m_0-1} E\left[(X_k(W(u_{k+1})-W(u_k)))^2\,\middle|\mathscr{F}_s\right]\\
&\qquad+2\sum_{0\leqslant i<j\leqslant m_0-1} E\left[X_iX_j\big(W(u_{i+1})-W(u_i)\big)\big(W(u_{j+1})-W(u_j)\big)\big|\mathscr{F}_s\right]\\
&\quad=\sum_{k=k_0}^{m_0-1} E\left[X_k^2 E\left[\big(W(u_{k+1})-W(u_k)\big)^2\big|\mathscr{F}_{u_k}\right]\middle|\mathscr{F}_s\right]\\
&\quad=\sum_{k=k_0}^{m_0-1} E\left[X_k^2 E\left[\big(W(u_{k+1})-W(u_k)\big)^2\right]\middle|\mathscr{F}_s\right]\\
&\quad=\sum_{k=k_0}^{m_0-1} E\left[X_k^2(u_{k+1}-u_k)\big|\mathscr{F}_s\right]\\
&\quad=\sum_{k=k_0}^{m_0-1} \int_{u_k}^{u_{k+1}} E\left[X_k^2\big|\mathscr{F}_s\right]du\\
&\quad=\int_s^t E[H^2(u)\big|\mathscr{F}_s]\,du.
\end{aligned}$$

(c) For each $X\in\mathscr{H}$ there exists a sequence $\{H_n\}_{n\in\mathbb{N}}$ of simple stochastic processes in $\mathscr{H}_0$ such that $\|H_n-X\|_{\mathscr{H}}\to 0$ and $E[|I(H_n)-I(X)|^2]\to 0$ as $n\to\infty$. By properties of conditional expectation[4] it follows by part (a) that

$$E\left[\int_s^t X(u)\,dW(u)\middle|\mathscr{F}_s\right]=\lim_{n\to\infty}E\left[\int_s^t H_n(u)\,dW(u)\middle|\mathscr{F}_s\right]=0,$$

where the limit takes place in $L^1_P(\Omega)$. Analogously, we conclude from part (a), again for limits in $L^1_P(\Omega)$, that

$$\begin{aligned}
E\left[\left(\int_s^t X(u)\,dW(u)\right)^2\middle|\mathscr{F}_s\right]&=\lim_{n\to\infty}E\left[\left(\int_s^t H_n(u)\,dW(u)\right)^2\middle|\mathscr{F}_s\right]\\
&=\lim_{n\to\infty}E\left[\int_s^t H_n^2(u)\,du\middle|\mathscr{F}_s\right]=E\left[\int_s^t X^2(u)\,du|\mathscr{F}_s\right],
\end{aligned}$$

where the last equality follows from $\|H_n-X\|_{\mathscr{H}}\to 0$.

[4] If $E[|Y_n\to Y|]\to 0$ then $E[|Y_n-Y|\,|\mathscr{C}]\to 0$ in mean for each sub-σ-algebra $\mathscr{C}$.

7. (i) The stochastic process $(W^2(t) : t \in [0,T])$ is in $\mathscr{H}$ since Lemma 3.3.4 guarantees

$$\int_0^T E\left[\left|W^2(s)\right|^2\right]\, ds = \int_0^T 3s^2\, ds < \infty.$$

(ii) Define a simple stochastic process $(H_n(t) : t \in [0,T])$ in $\mathscr{H}_0$ by

$$H_n(t) := \sum_{k=0}^{n-1} \mathbb{1}_{(t_k, t_{k+1}]}(t) W^2(t_k),$$

where $t_k := \frac{k}{n}T$ for $k = 0, \dots, n$. For each $0 \leqslant t \leqslant s$, the independent increments of W implies

$$\begin{aligned}
&E[W^2(t)W^2(s)] \\
&\quad = E\left[W^2(t)\big(W(s) - W(t) + W(t)\big)^2\right] \\
&\quad = E\left[W^2(t)\big(W(s) - W(t)\big)^2\right] + 2E\left[W^3(t)\big(W(s) - W(t)\big)\right] + E\left[W^4(t)\right] \\
&\quad = E\left[W^2(t)\right] E\left[\big(W(s) - W(t)\big)^2\right] + 2E\left[W^3(t)\right] E\left[\big(W(s) - W(t)\big)\right] + E\left[W^4(t)\right] \\
&\quad = t(s-t) + 2\,0 + 3t^3 = st + 2t^2.
\end{aligned}$$

Thus, we obtain for every $n \in \mathbb{N}$

$$\begin{aligned}
\left\|H_n - W^2\right\|_{\mathscr{H}}^2 &= \int_0^T E\left[\left|H_n(s) - W^2(s)\right|^2\right]\, ds \\
&= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} E\left[\left|W^2(t_k) - W^2(s)\right|^2\right]\, ds \\
&= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \big(E[W^4(t_k)] + E[W^4(s)] - 2E[W^2(t_k)W^2(s)]\big)\, ds \\
&= \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} \big(3t_k^2 + 3s^2 - 2(t_k s + 2t_k^2)\big)\, ds.
\end{aligned}$$

Using the inequality $-t_k^2 + 3s^2 - 2t_k s \leqslant -t_k^2 + 3s^2 - 2t_k^2 = 3(s^2 - t_k^2)$ for $t_k \leqslant s$ we obtain

$$\begin{aligned}\big\|H_n - W^2\big\|_{\mathscr{H}}^2 &\leqslant \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} 3(s^2 - t_k^2)\, ds \\ &\leqslant \sum_{k=0}^{n-1} 3(t_{k+1}^2 - t_k^2)(t_{k+1} - t_k) \\ &= 3\frac{T}{n} \sum_{k=0}^{n-1} t_{k+1}^2 - t_k^2 \\ &= 3\frac{T}{n} T^2 \\ &\to 0 \qquad \text{as } n \to \infty.\end{aligned}$$

(iii) Since $H_n \to W^2$ in $\mathscr{H}$ by (ii), we obtain according to Definition 4.2.7

$$\int_0^T W^2(s)\, dW(s) = \lim_{n\to\infty} \sum_{k=0}^{n-1} W^2(t_k)\Big(W(t_{k+1}) - W(t_k)\Big)$$

where the limit takes place in $L_P^2(\Omega)$. By using the algebraic identity

$$(b-a)^3 = b^3 - a^3 - 3ab^2 + 3a^2 b = b^3 - a^3 - 3a^2(b-a) - 3a(b-a)^2 \qquad \text{for all } a, b \in \mathbb{R},$$

we obtain for every $n \in \mathbb{N}$

$$\begin{aligned}&\sum_{k=0}^{n-1} W^2(t_k)\Big(W(t_{k+1}) - W(t_k)\Big) \\ &\quad = \frac{1}{3} W^3(T) - \frac{1}{3}\sum_{k=0}^{n-1}\big(W(t_{k+1}) - W(t_k)\big)^3 - \sum_{k=0}^{n-1} W(t_k)\big(W(t_{k+1}) - W(t_k)\big)^2.\end{aligned}$$

Independent increments of W and Lemma 3.3.4 yield for the first sum

$$\begin{aligned}E\left[\left|\sum_{k=0}^{n-1}\big(W(t_{k+1}) - W(t_k)\big)^3\right|^2\right] &= \sum_{k=0}^{n-1} E\left[\left|\big(W(t_{k+1}) - W(t_k)\big)^3\right|^2\right] \\ &= 15 \sum_{k=0}^{n-1} (t_{k+1} - t_k)^3 \\ &\leqslant 15 \sup_{k\in\{0,\dots,n-1\}} |t_{k+1} - t_k|^2 \sum_{k=0}^{n-1} t_{k+1} - t_k \\ &= 15 \frac{T^3}{n^2} \to 0 \quad \text{as } n \to \infty.\end{aligned}$$

For the second sum we show

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} W(t_k)\big(W(t_{k+1}) - W(t_k)\big)^2 = \int_0^T W(s)\,ds \qquad \text{in } L^2_P(\Omega). \tag{1.4.21}$$

The integral on the right hand side is a classical Riemann integral, thus we have by (4.1.1)

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} W(t_k)(t_{k+1} - t_k) = \int_0^T W(s)\,ds \qquad P\text{-a.s.}$$

Independent increments of W imply for every $n \in \mathbb{N}$

$$\begin{aligned}
E&\left[\left|\sum_{k=0}^{n-1} W(t_k)\big(W(t_{k+1}) - W(t_k)\big)^2 - \sum_{k=0}^{n-1} W(t_k)(t_{k+1} - t_k)\right|^2\right] \\
&= \sum_{k=0}^{n-1} E\left[W^2(t_k)\Big(\big(W(t_{k+1}) - W(t_k)\big)^2 - \big(t_{k+1} - t_k\big)\Big)^2\right] + 0 \\
&= \sum_{k=0}^{n-1} E\left[W^2(t_k)\right] E\left[\Big(\big(W(t_{k+1}) - W(t_k)\big)^2 - \big(t_{k+1} - t_k\big)\Big)^2\right] \\
&= \sum_{k=0}^{n-1} t_k \left(E\left[\big(W(t_{k+1}) - W(t_k)\big)^4\right] - \left(E\left[\big(W(t_{k+1}) - W(t_k)\big)^2\right]\right)^2\right) \\
&= \sum_{k=0}^{n-1} t_k (t_{k+1} - t_k)^2 \big(3(t_{k+1} - t_k) - 1\big)^2 \\
&\leqslant \frac{T^3}{n}\left(3\frac{T}{n} - 1\right)^2 \to 0,
\end{aligned}$$

which establishes (1.4.21). The last two estimates show that

$$\begin{aligned}
\lim_{n\to\infty} &\Big(\frac{1}{3}W^3(T) - \frac{1}{3}\sum_{k=0}^{n-1} \big(W(t_{k+1}) - W(t_k)\big)^3 - \sum_{k=0}^{n-1} W(t_k)\big(W(t_{k+1}) - W(t_k)\big)^2 \\
&= \frac{1}{3}W^3(T) - 0 - \int_0^T W(s)\,ds,
\end{aligned}$$

which completes part (a).

(b) Define the function

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) := x^3.$$

The function f is in C^2 and we obtain

$$f_x(x) = 3x^2, \quad f_{xx}(x) = 6x.$$

Itô's formula in Theorem 4.5.1 implies

$$\begin{aligned}(W(t))^3 = f(W(t)) &= f(0) + \int_0^t f_x(W(s))\,dW(s) + \frac{1}{2}\int_0^t f_{xx}(W(s))\,ds \\ &= 3\int_0^t W^2(s)\,dW(s) + 3\int_0^t W(s)\,ds.\end{aligned}$$

8. Define the function

$$f : [0,T] \times \mathbb{R} \to \mathbb{R} \qquad f(t,x) := tx.$$

The function f is in $C^{1,2}$ and we obtain that

$$f_t(t,x) = x, \quad f_x(t,x) = t, \quad f_{xx}(t,x) = 0.$$

Itô's formula in Theorem 4.5.6 implies

$$\begin{aligned}df(t,W(t)) &= \big(f_t(t,W(t)) + \tfrac{1}{2}f_{xx}(t,W(t))\big)\,dt + f_x(t,W(t))\,dW(t) \\ &= (W(t)+0)\,dt + t\,dW(t) \\ &= W(t)\,dt + t\,dW(t).\end{aligned}$$

Since $f(0,W(0)) = 0$ this means

$$tW(t) = \int_0^t W(s)\,ds + \int_0^t s\,dW(s).$$

(b) Define the function

$$f : [0,T] \times \mathbb{R} \to \mathbb{R} \qquad f(t,x) := e^{\frac{1}{2}t}\cos x.$$

The function f is in $C^{1,2}$ and we obtain that

$$f_t(t,x) = \tfrac{1}{2}e^{\frac{1}{2}t}\cos x, \quad f_x(t,x) = -e^{\frac{1}{2}t}\sin x, \quad f_{xx}(t,x) = -e^{\frac{1}{2}t}\cos x.$$

Itô's formula in Theorem 4.5.6 implies

$$\begin{aligned}df(t,W(t)) &= \big(f_t(t,W(t)) + \tfrac{1}{2}f_{xx}(t,W(t))\big)\,dt + f_x(t,W(t))\,dW(t) \\ &= (\tfrac{1}{2}e^{\frac{1}{2}t}\cos W(t) - \tfrac{1}{2}e^{\frac{1}{2}t}\cos W(t))\,dt - e^{\frac{1}{2}t}\sin W(t)\,dW(t) \\ &= -e^{\frac{1}{2}t}\sin W(t)\,dW(t).\end{aligned}$$

Since $f(0,W(0)) = 1$ this means

$$e^{\frac{1}{2}t}\cos W(t) = 1 - \int_0^t e^{\frac{1}{2}s}\sin W(s)\,dW(s).$$

(c) Define the function

$$f : [0,T] \times \mathbb{R} \to \mathbb{R} \qquad f(t,x) := \tfrac{1}{3}x^3 - tx.$$

The function f is in $C^{1,2}$ and we obtain

$$f_t(t,x) = -x, \quad f_x(t,x) = x^2 - t, \quad f_{xx}(t,x) = 2x.$$

Itô's formula in Theorem 4.5.6 implies

$$\begin{aligned} df(t,W(t)) &= \big(f_t(t,W(t)) + \tfrac{1}{2} f_{xx}(t,W(t))\big)\, dt + f_x(t,W(t))\, dW(t) \\ &= (-W(t) + W(t))\, dt + (W^2(t) - t)\, dW(t) \\ &= (W^2(t) - t)\, dW(t). \end{aligned}$$

Since $f(0,W(0)) = 0$ this means

$$\tfrac{1}{3}W^3(t) - tW(t) = \int_0^t (W^2(s) - s)\, dW(s). \tag{1.4.22}$$

The stochastic process $(W^2(t) - t : t \in [0,T])$ is in $\mathscr{H}$ since

$$\int_0^T E\left[\left|W^2(s) - s\right|^2\right] ds \leqslant 2 \int_0^T E[|W(s)|^4] + s^2\, ds = 8 \int_0^T s^2\, ds < \infty.$$

Consequently, $(\int_0^t (W^2(s) - s)\, dW(s) : t \in [0,T])$ is a martingale which completes the exercise due to the equality (1.4.22).

9. Follows later ...

10. (a) Brownian motion $((W_1(t), \dots, W_n(t)) : t \geqslant 0)$ in $\mathbb{R}^n$ is an Itô process of the form (4.7.21) with $\Upsilon = 0$ and $\Phi(s) = \mathrm{Id}_n$ for all $s \geqslant 0$ since

$$\begin{pmatrix} W_1(t) \\ W_2(t) \\ \vdots \\ W_n(t) \end{pmatrix} = 0 + \int_0^t \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \dots & \dots & 1 \end{pmatrix} dW(s) \qquad \text{for all } t \geqslant 0.$$

Since $f \in C^2(\mathbb{R}^n)$ Theorem 4.7.2 implies for all $t \in [0,T]$

$$\begin{aligned} f(W(t)) &= f(W(0)) + \frac{1}{2} \sum_{i=1}^n \int_0^t \Phi_{ii}(s)\Phi_{ii}(s) \frac{\partial^2 f}{\partial x_i \partial x_i}(W(s))\, ds \\ &\qquad + \sum_{i=1}^n \int_0^t \Phi_{ii}(s) \frac{\partial f}{\partial x_i}(W(s))\, dW_i(s) \\ &= f(W(0)) + \frac{1}{2} \int_0^t \Delta f(W(s))\, ds + \int_0^t \nabla f(s)\, dW(s). \end{aligned}$$

(b) Note that for all $i \in \{1, \dots, n\}$ Theorem 4.3.3 implies

$$W_i(t \wedge \tau) = \int_0^{t \wedge \tau} dW_i(s) = \int_0^t \mathbb{1}_{[0,\tau]}(s)\, dW_i(s) \qquad \text{for all } t \geqslant 0.$$

Thus, stopped Brownian motion $((W_1(t \wedge \tau), \dots, W_n(t \wedge \tau)) : t \geqslant 0)$ in $\mathbb{R}^n$ is an Itô process of the form (4.7.21) with $\Upsilon = 0$ and $\Phi(s) = \mathbb{1}_{[0,\tau]}(s)\,\mathrm{Id}_n$ for all $s \geqslant 0$ since

$$\begin{pmatrix} W_1(t \wedge \tau) \\ W_2(t \wedge \tau) \\ \vdots \\ W_n(t \wedge \tau) \end{pmatrix} = 0 + \int_0^t \begin{pmatrix} \mathbb{1}_{[0,\tau]}(s) & 0 & \dots & \dots & 0 \\ 0 & \mathbb{1}_{[0,\tau]}(s) & \dots & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \dots & \dots & \mathbb{1}_{[0,\tau]}(s) \end{pmatrix} dW(s) \quad \text{for all } t \geqslant 0.$$

Since f is in $C^2(\mathbb{R}^n)$ and it is harmonic on D Theorem 4.7.2 implies for all $t \in [0, T]$

$$\begin{aligned} f(W(t \wedge \tau)) &= f(W(0)) + \frac{1}{2} \sum_{i=1}^n \int_0^t \Phi_{ii}(s)\Phi_{ii}(s) \frac{\partial^2 f}{\partial x_i \partial x_i}\big(W(s \wedge \tau)\big)\, ds \\ &\qquad + \sum_{i=1}^n \int_0^t \Phi_{ii}(s) \frac{\partial f}{\partial x_i}\big(W(s \wedge \tau)\big)\, dW_i(s) \\ &= f(W(0)) + \int_0^t \mathbb{1}_{[0,\tau]}(s) \nabla f(W(s \wedge \tau))\, dW(s). \end{aligned}$$

Consequently, $(f(W(t \wedge \tau)) : t \in [0, T])$ is a local martingale.

A.5. Solution Chapter 5

1. Define the function

$$f \colon [0, T] \times \mathbb{R} \to \mathbb{R} \qquad f(t, x) := x^m.$$

The function f is in $C^{1,2}$ and we obtain

$$f_t(t, x) = 0, \quad f_x(t, x) = m x^{m-1}, \quad f_{xx}(t, x) = m(m-1)x^{m-2}.$$

Since S is the solution of a stochastic differential equation it is an Itô process. Consequently, Itô's formula in Theorem 4.6.5 implies

$$\begin{aligned} dZ(t) = df(t, S(t)) &= \Big(f_t(t, S(t)) + f_x(t, S(t))\Upsilon(t) + \tfrac{1}{2} f_{xx}(t, S(t))\Phi^2(t)\Big)\, dt \\ &\qquad + f_x\big(t, S(t)\big)\Phi(t)\, dW(t) \\ &= \Big(m S^{m-1}(t) \alpha S(t) + \tfrac{1}{2} m(m-1) S^{m-2}(t) \sigma^2 S^2(t)\Big)\, dt \\ &\qquad + m S^{m-1}(t) \sigma S(t)\, dW(t) \\ &= \big(m\alpha + \tfrac{1}{2} m(m-1)\sigma^2\big) Z(t)\, dt + m\sigma Z(t)\, dW(t). \end{aligned}$$

Since the application of Itô's formula above implies all other conditions in Definition 5.1.1 the stochastic process Z is the solution of the stochastic differential equation

$$\begin{aligned} dZ(t) &= \big(m\alpha + \tfrac{1}{2}m(m-1)\sigma^2\big)Z(t)\,dt + m\sigma Z(t)\,dW(t)\\ Z(0) &= S^m(0). \end{aligned}$$

2. (a) Define the function

$$f\colon \mathbb{R}\to\mathbb{R} \qquad f(x) := \begin{cases} (x-a)^3, & \text{if } x \geqslant a,\\ 0, & \text{else.}\end{cases}$$

By considering the function f on $(-\infty, a)$ and $[a,\infty)$ and realising that the derivate of f from the left and right in a coincide, it follows that f is in C^2. The derivatives are given by

$$f_x(x) = \begin{cases} 3(x-a)^2, & \text{if } x \geqslant a,\\ 0, & \text{else,}\end{cases} \qquad f_{xx}(x) = \begin{cases} 6(x-a), & \text{if } x \geqslant a,\\ 0, & \text{else.}\end{cases}$$

Since $f_x(x) \geqslant 0$ and $f_{xx}(x) \geqslant 0$ for all $x\in\mathbb{R}$ the roots in the following are well defined, and the derivatives can be written as

$$f_x(x) = 3\big(f(x)\big)^{2/3}, \quad f_{xx}(t,x) = 6\big(f(x)\big)^{1/3}.$$

Itô's formula in Theorem 4.5.1 implies

$$\begin{aligned} dX(t) = df(W(t)) &= f_x(W(t))\,dW(t) + \tfrac{1}{2}f_{xx}(W(t))\,dt\\ &= 3\big(f(W(t)\big)^{2/3}\,dW(t) + 3\big(f(W(t))\big)^{1/3}\,dt\\ &= 3\,(X(t))^{2/3}\,dW(t) + 3\,(X(t))^{1/3}\,dt. \end{aligned}$$

Since the application of Itô's formula above implies all other conditions in Definition 5.1.1 and since we have $X(0) = 0$ by the definition of X, it follows that X is a solution of the given stochastic differential equation.

(b) In part (a) we show that the stochastic process

$$X(t) = \begin{cases} (W(t)-a)^3, & \text{if } W(t) \geqslant a,\\ 0, & \text{else}\end{cases}$$

is a solution of the stochastic differential equation (5.7.21) for all $a > 0$. Consequently, there exists infinitely many solutions of (5.7.21).

The coefficients of this stochastic differential equation are

$$f\colon \mathbb{R}_+\to\mathbb{R},\quad f(x) = 3x^{1/3} \quad\text{and}\quad g\colon \mathbb{R}_+\to\mathbb{R},\quad g(x) = 3x^{2/3}.$$

Both functions are not Lipschitz continuous on any interval $[0,h]$ for every $h > 0$ since their derivatives tend to infinity for $x \searrow 0$.

I am a bit sloppy here since the functions f and g are only defined on $\mathbb{R}_+$ whereas Theorem 5.1.5 considers functions defined on the whole line $\mathbb{R}$. This can be easily fixed by taking the modulus in the considered stochastic differential equation.

3. The exercise assumes that X is a solution which makes only sense if $X(t) > 0$ P-a.s. for all $t \in [0,T]$. Part (d) finally confirms that the representation of X – derived under this assumption – solves the stochastic differential equation.

   (a) Define the function

$$f\colon [0,T]\times(0,\infty)\to\mathbb{R} \qquad f(t,x):=\ln(x).$$

   The function f is in $C^{1,2}([0,T]\times(0,\infty)$ and we obtain

$$f_t(t,x)=0, \quad f_x(t,x)=\frac{1}{x}, \quad f_{xx}(t,x)=-\frac{1}{x^2}.$$

   The stochastic process $(X(t): t\in[0,T])$ defined by

$$X(t):=X(0)+\int_0^t \underbrace{b(a-\ln(X(s)))X(s)}_{=:\Upsilon(s)}\,ds+\int_0^t \underbrace{\sigma X(s)}_{=:\Phi(s)}\,dW(s)$$

   is an Itô process. Since the stochastic process X satisfies $X(t)\in(0,\infty)$ P-a.s. for all $t\in[0,T]$ we can apply the usual Itô's formula although the function f is not defined on the complete space $\mathbb{R}$, see Remark 4.7.3. Consequently, Itô's formula in Theorem 4.6.5 implies

$$\begin{aligned} dY(t)=df(t,X(t)) &= \big(f_t(t,X(t))+f_x(t,X(t))\Upsilon(t)+\tfrac12 f_{xx}(t,X(t))\Phi^2(t)\big)\,dt \\ &\qquad + f_x(t,X(t))\Phi(t)\,dW(t) \\ &= \left(\frac{1}{X(t)}b(a-\ln(X(t)))X(t)+\frac{-1}{2X^2(t)}\sigma^2X^2(t)\right)dt \\ &\qquad +\frac{1}{X(t)}\sigma X(t)\,dW(t) \\ &= \left(ab-\tfrac{\sigma^2}{2}-bY(t)\right)dt+\sigma\,dW(t). \end{aligned}$$

   Since the application of Itô's formula above implies all other conditions in Definition 5.1.1 and since we have $Y(0)=\ln(X(0))=\ln(x_0)$ by the definition of Y, it follows that Y is a solution of the stochastic differential equation

$$\begin{aligned} dY(t) &= \left(ab-\tfrac{\sigma^2}{2}-bY(t)\right)dt+\sigma\,dW(t) \qquad \text{for all t} \in [0,\text{T}], \\ Y(0) &= \ln(x_0). \end{aligned}$$

   (b) Define the function

$$g\colon [0,T]\times\mathbb{R}\to\mathbb{R} \qquad g(t,x):=e^{bt}x.$$

   The function g is in $C^{1,2}([0,T]\times\mathbb{R})$ and we obtain

$$g_t(t,x)=be^{bt}x, \quad g_x(t,x)=e^{bt}, \quad g_{xx}(t,x)=0.$$

The stochastic process $(Y(t) : t \in [0, T])$ defined by

$$Y(t) := Y(0) + \int_0^t \underbrace{\Big(ab - \tfrac{\sigma^2}{2} - bY(s)\Big)}_{=:\Upsilon(s)} ds + \int_0^t \underbrace{\sigma}_{=:\Phi(s)} dW(s),$$

is an Itô process. Consequently, for $Z(t) := Y(t)\exp(bt)$, Itô's formula in Theorem 4.6.5 implies

$$\begin{aligned} dZ(t) = dg(t, Y(t)) &= \big(g_t(t, Y(t)) + g_x(t, Y(t))\Upsilon(t) + \tfrac{1}{2} g_{xx}(t, Y(t))\Phi^2(t)\big)\, dt \\ &\qquad + g_x(t, Y(t))\Phi(t)\, dW(t) \\ &= \Big(be^{bt}Y(t) + e^{bt}\Big(ab - \tfrac{\sigma^2}{2} - bY(t)\Big)\Big)\, dt + e^{bt}\sigma\, dW(t) \\ &= \Big(ab - \tfrac{\sigma^2}{2}\Big) e^{bt}\, dt + \sigma e^{bt}\, dW(t). \end{aligned}$$

This means that for all $t \in [0, T]$ we have

$$\begin{aligned} Z(t) &= Z(0) + \int_0^t \Big(ab - \tfrac{\sigma^2}{2}\Big) e^{bs}\, ds + \int_0^t \sigma e^{bs}\, dW(s) \\ &= \ln(x_0) + \frac{1}{b}\Big(ab - \tfrac{\sigma^2}{2}\Big)\big(e^{bt} - 1\big) + \sigma \int_0^t e^{bs}\, dW(s). \end{aligned} \tag{1.5.23}$$

(c) It follows from (b) for all $t \in [0, T]$

$$\begin{aligned} Y(t) &= e^{-bt} Z(t) \\ &= e^{-bt}\ln(x_0) + \frac{1}{b}\Big(ab - \tfrac{\sigma^2}{2}\Big)\big(1 - e^{-bt}\big) + \sigma \int_0^t e^{-b(t-s)}\, dW(s). \end{aligned}$$

By the definition of Y we obtain for all $t \in [0, T]$

$$\begin{aligned} X(t) &= \exp(Y(t)) \\ &= \exp\Big(e^{-bt}\ln(x_0) + \Big(ab - \tfrac{\sigma^2}{2}\Big)\big(1 - e^{-bt}\big) + \sigma \int_0^t e^{-b(t-s)}\, dW(s)\Big). \end{aligned} \tag{1.5.24}$$

(d) Define the function

$$f \colon [0, T] \times \mathbb{R} \to \mathbb{R} \qquad f(t, x) := \exp\Big(\exp(-bt)\, x\Big).$$

The function f is in $C^{1,2}$ and we obtain

$$f_t(t, x) = -bxe^{-bt} f(t, x), \quad f_x(t, x) = e^{-bt} f(t, x), \quad f_{xx}(t, x) = e^{-2bt} f(t, x).$$

Since Z is an Itô process as it can be seen in (1.5.23), Itô's formula in Theorem 4.6.5 implies

$$\begin{aligned} df(t,Z(t)) &= \big(f_t(t,Z(t)) + f_x(t,Z(t))\Upsilon(t) + \tfrac{1}{2} f_{xx}(t,Z(t))\Phi^2(t)\big)\, dt \\ &\qquad + f_x(t,Z(t))\Phi(t)\, dW(t) \\ &= f(t,Z(t))\Big(-bZ(t)e^{-bt} + e^{-bt}\big(ab - \tfrac{\sigma^2}{2}\big)e^{bt} + \tfrac{1}{2}e^{-2bt}\sigma^2 e^{2bt}\Big) \\ &\qquad + f(t,Z(t))e^{-bt}\sigma e^{bt}\, dW(t) \\ &= b\, f(t,Z(t))\big(a - e^{-bt}Z(t)\big)\, dt + \sigma f(t,Z(t))\, dW(t). \end{aligned}$$

Note that according to the representation (1.5.24), we have $X(t) = f(t,Z(t))$ for all $t \in [0,T]$ and thus, $\ln(X(t)) = Z(t)\exp(-bt)$. Since the application of Itô's formula above implies all other conditions in Definition 5.1.1 and since we have $X(0) = f(0,Z(0)) = x_0$ by the definition of Z, it follows that X is a solution of the stochastic differential equation

$$\begin{aligned} dX(t) &= bX(t)\big(a - \ln(X(t)\big)\, dt + \sigma X(t)\, dW(t) \qquad \text{for all } t \in [0,T], \\ X(0) &= x_0. \end{aligned}$$

4. Define the function

$$f : [0,T] \times (0,\infty) \to \mathbb{R} \qquad f(t,x) := \frac{1}{x}.$$

The function f is in $C^{1,2}([0,T] \times (0,\infty)$ and we obtain

$$f_t(t,x) = 0, \quad f_x(t,x) = -x^{-2}, \quad f_{xx}(t,x) = 2x^{-3}.$$

The stochastic process $(X(t) : t \in [0,T])$ defined by

$$X(t) := X(0) + \int_0^t \underbrace{\alpha X(s)}_{=:\Upsilon(s)}\, ds + \int_0^t \underbrace{\sigma X(s)}_{=:\Phi(s)}\, dW(s),$$

is an Itô process. According to Proposition 5.3.2 the stochastic process X satisfies $X(t) \in (0,\infty)$ P-a.s. for all $t \in [0,T]$. Thus, we can apply the usual Itô's formula although the function f is not defined on the complete space $\mathbb{R}$, see Remark 4.7.3. Consequently, Itô's formula in Theorem 4.6.5 implies

$$\begin{aligned} dX^{-1}(t) = df(t,X(t)) &= \big(f_t(t,X(t)) + f_x(t,X(t))\Upsilon(t) + \tfrac{1}{2} f_{xx}(t,X(t))\Phi^2(t)\big)\, dt \\ &\qquad + f_x(t,X(t))\Phi(t)\, dW(t) \\ &= (-X^{-2}(t)\alpha X(t) + \tfrac{1}{2} 2X^{-3}(t)\sigma^2 X^2(t))\, dt \\ &\qquad - X^{-2}(t)\sigma X(t)\, dW(t) \\ &= (-\alpha + \sigma^2)X^{-1}(t)\, dt - \sigma X^{-1}(t)\, dW(t). \end{aligned}$$

Since the application of Itô's formula above implies all other conditions in Definition 5.1.1, it follows that X^{-1} is a solution of the stochastic differential equation

$$\begin{aligned} dZ(t) &= (-\alpha + \sigma^2) Z(t)\,dt - \sigma Z(t)\,dW(t), \qquad \text{for all t} \in [0, \mathrm{T}], \\ Z(0) &= \frac{1}{x_0}. \end{aligned}$$

5. Define the function

$$f\colon [0,T] \times \mathbb{R}^2 \to \mathbb{R} \qquad f(t,x,y) := xy.$$

The function f is in $C^{1,2,2}$ and we obtain

$$f_x(t,x,y) = y, \quad f_y(t,x,y) = x, \quad f_{xy}(t,x,y) = f_{yx}(t,x,y) = 1,$$

and all other derivatives vanish.

The two-dimensional stochastic process $((X(t), Y(t)) : t \in [0,T])$ is of form

$$\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} X(0) \\ Y(0) \end{pmatrix} + \int_0^t \begin{pmatrix} \Upsilon_1(s) \\ \Upsilon_2(s) \end{pmatrix} ds + \int_0^t \begin{pmatrix} \Phi_1(s) \\ \Phi_2(s) \end{pmatrix} dW(s).$$

Since X and Y are assumed to be real-valued Itô processes it follows that $((X(t), Y(t)) : t \in [0,T])$ is a two-dimensional Itô processes of the form (4.7.21) with $n = 2$ and $d = 1$. Itô's formula in Theorem 4.7.2 implies

$$\begin{aligned} &d(X(t)Y(t)) \\ &\quad = df(t, X(t), Y(t)) \\ &\quad = \Big(\Upsilon_1(t) f_x(t, X(t), Y(t)) + \Upsilon_2(t) f_y(t, X(t), Y(t)) \Big)\,dt \\ &\qquad\quad + \tfrac{1}{2} \Big(\Phi_1(t)\Phi_2(t) f_{x,y}(t, X(t), Y(t)) + \Phi_2(t)\Phi_1(t) f_{y,x}(t, X(t), Y(t)) \Big)\,dt \\ &\qquad\quad + \Phi_1(t) f_x(t, X(t), Y(t))\,dW(t) + \Phi_2(t) f_y(t, X(t), Y(t))\,dW(t) \\ &\quad = \Big(\Upsilon_1(t) Y(t) + \Upsilon_2(t) X(t) \Big)\,dt + \Phi_1(t)\Phi_2(t)\,dt + \Phi_1(t) Y(t)\,dW(t) + \Phi_2(t) X(t)\,dW(t) \\ &\quad = Y(t)\,dX(t) + X(t)\,dY(t) + \Phi_1(t)\Phi_2(t)\,dt. \end{aligned}$$

6. (a) The coefficients are given by the functions

$$f(t,z) = \begin{pmatrix} \alpha x \\ \alpha y \end{pmatrix}, \qquad g(t,z) = \begin{pmatrix} -y \\ x \end{pmatrix}$$

for $t \in [0,T]$ and $z = (x,y) \in \mathbb{R}^2$. For each $t \in [0,T]$ and $z_i = (x_i, y_i) \in \mathbb{R}^2$, $i = 1, 2$, we obtain

$$\begin{aligned}
\|f(t,z_1)-f(t,z_2)\|_2 &= |\alpha|\left(|x_1-x_2|^2+|y_1-y_2|^2\right)^{1/2} = |\alpha|\,\|z_1-z_2\|_2\,,\\
\|f(t,z_1)\|_2^2 &= \alpha^2\,\|z_1\|_2^2 \leqslant \alpha^2(1+\|z_1\|_2^2)\\
\|g(t,z_1)-g(t,z_2)\|_{\mathrm{HS}} &= \left(|-y_1+y_2|^2+|x_1-x_2|^2\right)^{1/2} = \|z_1-z_2\|_2\,,\\
\|g(t,z_1)\|_{\mathrm{HS}}^2 &= \left(|-y_1|^2+|x_1|^2\right) = \|z_1\|_2^2 \leqslant 1+\|z_1\|_2^2\,.
\end{aligned}$$

Thus, the functions f and g satisfy the Lipschitz and linear growth conditions in Theorem 5.5.2 and it follows that there exists a unique solution of the stochastic differential equation.

(b) Define the function

$$f\colon [0,T]\times\mathbb{R}\to\mathbb{R} \qquad f(t,x):=x^2.$$

The function f is in $C^{1,2}$ and we obtain

$$f_t(t,x)=0,\quad f_x(t,x)=2x,\quad f_{xx}(t,x)=2.$$

Let $(Z(t):t\in[0,T])$ be the solution of the given stochastic differential equation and denote $Z(t)=\big(X(t),Y(t)\big)$ for all $t\in[0,T]$. Then the stochastic process $(X(t):t\in[0,T])$ obeys

$$X(t):=X(0)+\int_0^t \underbrace{\alpha X(s)}_{=:\Upsilon(s)}\,ds+\int_0^t \underbrace{-Y(s)}_{=:\Phi(s)}\,dW(s),$$

and thus, X is an Itô process. Consequently, Itô's formula in Theorem 4.6.5 implies

$$\begin{aligned}
dX^2(t)=df(t,X(t)) &= \big(f_t(t,X(t))+f_x(t,X(t))\Upsilon(t)+\tfrac12 f_{xx}(t,X(t))\Phi^2(t)\big)\,dt\\
&\qquad + f_x(t,X(t))\Phi(t)\,dW(t)\\
&= (2X(t)\alpha X(t)+\tfrac12 2Y^2(t))\,dt+2X(t)Y(t)\,dW(t)\\
&= \big(2\alpha X^2(t)+Y^2(t)\big)\,dt-2X(t)Y(t)\,dW(t).
\end{aligned}$$

Analogously, we obtain that

$$dY^2(t)=\big(2\alpha Y^2(t)+X^2(t)\big)\,dt+2X(t)Y(t)\,dW(t).$$

Applying the representations above of X and Y we obtain for $R(t):=X^2(t)+Y^2(t)$, $t\in[0,T]$:

$$\begin{aligned}
R(t) &= R(0)+\int_0^t \big(2\alpha X^2(s)+Y^2(s)+2\alpha Y^2(t)+X^2(t)\big)\,ds\\
&\qquad +\int_0^t (-2X(s)Y(s)+2X(s)Y(s))\,dW(s)\\
&= x_0^2+y_0^2+\int_0^t (1+2\alpha)R(s)\,ds.
\end{aligned}$$

Since the integral on the right hand side is a standard integral it follows from the fundamental theorem of calculus that for each $\omega \in \Omega$ the function $t \mapsto R(t)(\omega)$ is differentiable, which results in

$$R'(t) = (1+2\alpha)R(t) \qquad \text{for all } t \in [0,T].$$

Thus, for each $\omega \in \Omega$ the function $g_\omega(\cdot) := R(\cdot)(\omega)$ solves the ordinary differential equation

$$\begin{aligned} h'(t) &= (1+2\alpha)h(t) \qquad \text{for all } t \in [0,T] \\ h(0) &= x_0^2 + y_0^2. \end{aligned}$$

This ordinary differential equation has a unique solution h given by $h(t) = h(0)\exp((1+2\alpha)t)$ for $t \in [0,T]$. Since this solution is unique it follows

$$R(t)(\omega) = h(t) \quad \text{for all } t \in [0,T] \text{ and } \omega \in \Omega,$$

which shows that R does not depend on ω.

(c) Theorem 5.5.2 implies that there exists a constant $\alpha > 0$ such that

$$E\left[\|(X(t),Y(t))\|_2^2\right] \leqslant \alpha(1 + \|(x_0,y_0)\|_2^2) \qquad \text{for all } t \in [0,T].$$

Thus, the moments $E[X(t)]$, $E[Y(t)]$ and $E[X(t)Y(t)]$ exist and we also have

$$\begin{aligned} \int_0^T E\left[|X(s)|^2\right]\, ds &\leqslant \alpha(1 + \|(x_0,y_0)\|_2^2)T, \\ \int_0^T E\left[|Y(s)|^2\right]\, ds &\leqslant \alpha(1 + \|(x_0,y_0)\|_2^2)T. \end{aligned}$$

It follows that the stochastic processes X, Y are in $\mathscr{H}$. Taking expectation of the stochastic differential equation solved by X we obtain

$$\begin{aligned} E[X(t)] &= E\left[x_0 + \alpha\int_0^t X(s)\, ds - \int_0^t Y(s)\, dW(s)\right] \\ &= x_0 + \alpha\int_0^t E\left[X(s)\right]\, ds. \end{aligned}$$

Thus, the function $g(t) := E[X(t)]$ satisfies

$$g'(t) = \alpha g(t) \qquad \text{for all } t \in [0,T]$$

and therefore we obtain that $E[X(t)] = x_0\exp(\alpha t)$ for all $t \in [0,T]$. Analogously, we derive $E[Y(t)] = y_0\exp(\alpha t)$ for all $t \in [0,T]$.

Since X and Y are real-valued Itô processes it follows from Exercise 5 that

$$\begin{aligned} d(X(t)Y(t)) &= Y(t)\, dX(t) + X(t)\, dY(t) + \Phi_1(t)\Phi_2(t)\, dt \\ &= Y(t)\alpha X(t)\, dt + Y^2(t)\, dW(t) + X(t)\alpha Y(t)\, dt - X^2(t)\, dW(t) - Y(t)X(t)\, dt \\ &= \Big(2\alpha - 1\Big)X(t)Y(t)\, dt + \Big(Y^2(t) - X^2(t)\Big)\, dW(t). \end{aligned}$$

Taking expectation from this stochastic differential equation results in

$$E[X(t)Y(t)] = x_0y_0(2\alpha - 1)\int_0^t E[X(s)Y(s)]\,ds \qquad \text{for all t} \in [0, \mathrm{T}]\,.$$

As above we obtain that $E[X(t)Y(t)] = x_0y_0\exp((2\alpha-1)t)$ for all $t \in [0,T]$, which enables us to calculate

$$\operatorname{Cov}(X(t),Y(t)) = E[X(t)Y(t)] - E[X(t)]E[Y(t)] = x_0y_0e^{2\alpha t}\left(e^{-t}-1\right).$$

7. (a) Define the function

$$f:[0,T]\times\mathbb{R}\to\mathbb{R} \qquad f(t,x) := x\exp\left(-\frac{1}{2}\int_0^t \alpha(u)\,du\right).$$

The function f is in $C^{1,2}$ and we obtain

$$f_t(t,x) = -\tfrac{1}{2}\alpha(t)f(t,x), \quad f_x(t,x) = \exp\left(-\frac{1}{2}\int_0^t \alpha(u)\,du\right), \quad f_{xx}(t,x) = 0.$$

Since the functions α and σ are assumed to be continuous, the function

$$\Phi\colon [0,T]\to\mathbb{R}, \qquad \Phi(t) := \tfrac{1}{2}\sigma(t)\exp\left(\tfrac{1}{2}\int_0^t \alpha(u)\,du\right)$$

is also continuous, and thus satisfies

$$\int_0^T E\left[|\Phi(s)|^2\right]\,ds = \int_0^T |\Phi(s)|^2\,ds \leqslant T\sup_{t\in[0,T]}|\Phi(s)|^2 < \infty.$$

Consequently, we can define a stochastic process $(Y_i(t) : t\in[0,T])$ for $i=1,2$ by

$$Y_i(t) := x_i + \int_0^t \underbrace{0}_{=:\Upsilon(s)}\,ds + \int_0^t \underbrace{\left(\tfrac{1}{2}\sigma(s)\exp\left(\tfrac{1}{2}\int_0^s \alpha(u)\,du\right)\right)}_{=\Phi(s)}\,dW_i(s),$$

and Y_i is seen to be an Itô process. Itô's formula in Theorem 4.6.5 implies

$$\begin{aligned} df(t,Y_i(t)) &= \left(f_t(t,Y_i(t)) + f_x(t,Y_i(t))\Upsilon(t) + \tfrac{1}{2}f_{xx}(t,Y_i(t))\Phi^2(t)\right)\,dt \\ &\qquad + f_x(t,Y_i(t))\Phi(t)\,dW_i(t) \\ &= \left(-\tfrac{1}{2}\alpha(t)f(t,Y_i(t))\right)dt + \tfrac{1}{2}\sigma(t)\,dW_i(t). \end{aligned}$$

Since $X_i(t) = f(t,Y_i(t))$ for all $t\in[0,T]$ by definition of X_i, this shows that X_i is the solution of the stochastic differential equation

$$\begin{aligned} dX_i(t) &= -\tfrac{1}{2}\alpha(t)X_i(t)\,dt + \tfrac{1}{2}\sigma(t)\,dW_i(t) \qquad \text{for all } t\in[0,T], \\ X_i(0) &= x_i. \end{aligned}$$

(b) Lemma 5.2.3 implies that for each $t \in [0,T]$ the random variable $Y_i(t)$ is normally distributed with

$$Y_i(t) \stackrel{\mathscr{D}}{=} N\left(x_i, \int_0^t \Phi^2(s)\,ds\right).$$

Since $X_i(t) = f(t, Y_i(t))$ and $f(t,x)$ is for fixed $t \in [0,T]$ just the multiplication of x by a constant, it follows that $X_i(t)$ is also normally distributed with

$$\begin{aligned} E\left[X_i(t)\right] &= x_i \exp\left(-\tfrac{1}{2}\int_0^t \alpha(u)\,du\right) \\ \operatorname{Var}\left[X_i(t)\right] &= \exp\left(-\int_0^t \alpha(u)\,du\right)\int_0^t \Phi^2(s)\,ds. \end{aligned}$$

c) Define the function

$$f : [0,T] \times \mathbb{R}^2 \to \mathbb{R} \qquad f(t,x,y) := x^2 + y^2.$$

The function f is in $C^{1,2,2}$ and we obtain

$$f_x(t,x,y) = 2x, \quad f_y(t,x,y) = 2y, \quad f_{xx}(t,x,y) = f_{yy}(t,x,y) = 2,$$

and all other derivatives vanish. The 2-dimensional process $((X_1(t), X_2(t)) : t \in [0,T])$ can be written in the form as in (4.7.21):

$$\begin{aligned} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \int_0^t \begin{pmatrix} \Upsilon_1(s) \\ \Upsilon_2(s) \end{pmatrix} ds + \int_0^t \begin{pmatrix} \Phi_{1,1}(s) & \Phi_{1,2}(s) \\ \Phi_{2,1}(s) & \Phi_{2,2}(s) \end{pmatrix} d\begin{pmatrix} W_1(s) \\ W_2(s) \end{pmatrix} \\ &:= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \int_0^t \begin{pmatrix} -\frac{1}{2}\alpha(s)X_1(s) \\ -\frac{1}{2}\alpha(s)X_2(s) \end{pmatrix} ds + \int_0^t \begin{pmatrix} \frac{1}{2}\sigma(s) & 0 \\ 0 & \frac{1}{2}\sigma(s) \end{pmatrix} d\begin{pmatrix} W_1(s) \\ W_2(s) \end{pmatrix}. \end{aligned}$$

Since both X_1 and X_2 are real-valued Itô processes, the two-dimensional stochastic process $((X_1(t), X_2(t)) : t \in [0,T])$ is an Itô process. Itô's formula in Theorem 4.7.2 implies

$$\begin{aligned} &d(R(t)) \\ &\quad = df(t, X_1(t), X_2(t)) \\ &\quad = \Big(\Upsilon_1(t) f_x(t, X_1(t), X_2(t)) + \Upsilon_2(t) f_y(t, X_1(t), X_2(t)\Big)\,dt \\ &\qquad + \tfrac{1}{2}\Big(\Phi_{1,1}(t)\Phi_{1,1}(t) f_{x,x}(t, X_1(t), X_2(t)) + \Phi_{2,2}(t)\Phi_{2,2}(t) f_{y,y}(t, X_1(t), X_2(t))\Big)\,dt \\ &\qquad + \Phi_{1,1}(t) f_x(t, X_1(t), X_2(t))\,dW_1(t) + \Phi_{2,2}(t) f_y(t, X_1(t), X_2(t))\,dW_2(t) \\ &\quad = \Big(-\frac{1}{2}\alpha(t)X_1(t)2X_1(t) - \tfrac{1}{2}\alpha(t)X_2(t)2X_2(t)\Big)\,dt + \tfrac{1}{2}\Big(\tfrac{1}{4}\sigma^2(t)2 + \tfrac{1}{4}\sigma^2(t)2\Big)dt \\ &\qquad + \tfrac{1}{2}\sigma(t)2X_1(t)\,dW_1(t) + \tfrac{1}{2}\sigma(t)2X_2(t)\,dW_2(t) \\ &\quad = \Big(\tfrac{1}{2}\sigma^2(t) - \alpha(t)R(t)\Big)\,dt + \sigma(t)X_1(t)\,dW_1(t) + \sigma(t)X_2(t)\,dW_2(t). \end{aligned}$$

8. Since the stochastic differential equation looks very similar to the one which defines the Ornstein-Uhlenbeck process in Definition 5.2.1 we try the same approach as in the proof of Proposition 5.2.2. Define the function

$$f : [0,T] \times \mathbb{R} \to \mathbb{R} \qquad f(t,x) := e^{-at}x.$$

The function f is in $C^{1,2}$ and we obtain

$$f_t(t,x) = -ae^{-at}x, \quad f_x(t,x) = e^{-at}, \quad f_{xx}(t,x) = 0.$$

The stochastic process $(Y(t) : t \in [0,T])$ defined by

$$Y(t) := \underbrace{r_0 - b}_{=:Y(0)} + \int_0^t \underbrace{0}_{=:\Upsilon(s)} \, ds + \int_0^t \underbrace{\sigma e^{as}}_{=:\Phi(s)} \, dW(s),$$

is an Itô process. Consequently, Itô's formula in Theorem 4.6.5 implies

$$\begin{aligned} df(t,Y(t)) &= \big(f_t(t,Y(t)) + f_x(t,Y(t))\Upsilon(t) + \tfrac{1}{2} f_{xx}(t,Y(t))\Phi^2(t)\big)\, dt \\ &\qquad + f_x(t,Y(t))\Phi(t)\, dW(t) \\ &= -ae^{-at}Y(t)\, dt + e^{-at}\sigma e^{at}\sigma\, dW(t). \end{aligned}$$

This means that

$$e^{-at}Y(t) = r_0 - b + \int_0^t -ae^{-as}Y(s)\, ds + \sigma W(t) \qquad \text{for all t} \in [0,\mathrm{T}]\,.$$

If we define $R(t) := e^{-at}Y(t) + b$ for all $t \in [0,T]$, a simple rewriting shows

$$\begin{aligned} R(t) &= r_0 + \int_0^t -ae^{-as}Y(s)\, ds + \sigma W(t) \\ &= r_0 + \int_0^t -a(e^{-as}Y(s) + b - b)\, ds + \sigma W(t) \\ &= r_0 + \int_0^t -a(R(s) - b)\, ds + \sigma W(t). \end{aligned}$$

Thus, the stochastic process R solves the stochastic differential equation.

A.6. Solution Chapter 6

1. (a) The random variable $Y := W(T)\exp(-\int_0^T s^2\, dW(s))$ can be written as

$$\begin{aligned} Y &= W(T)\exp\left(\tfrac{1}{2}\int_0^T s^4\, ds\right)\exp\left(-\int_0^T s^2\, dW(s) - \tfrac{1}{2}\int_0^T s^4\, ds\right) \\ &= W(T)\exp\left(\tfrac{T^5}{10}\right)\exp\left(-\int_0^T s^2\, dW(s) - \tfrac{1}{2}\int_0^T s^4\, ds\right). \end{aligned} \tag{1.6.25}$$

Define the stochastic process $L := (L(t) : t \in [0,T])$ by

$$L(t) := \exp\left(-\int_0^t s^2\, dW(s) - \tfrac{1}{2}\int_0^t s^4\, ds\right).$$

Due to

$$P\left(\int_0^T s^4(s)\, ds < \infty\right) = P\left(\tfrac{T^5}{5} < \infty\right) = 1,$$
$$E\left[\exp\left(\tfrac{1}{2}\int_0^T s^4\, ds\right)\right] = \exp\left(\tfrac{T^5}{10}\right) < \infty,$$

Novikov's Theorem 6.1.8 implies that L is a martingale, and thus we can conclude from Girsanov's Theorem 6.1.3 that

$$Q : \mathscr{A} \to [0,\infty], \qquad Q(A) := \int_A L(T)(\omega)\, P(d\omega) = E_P[\mathbb{1}_A L(T)]$$

is a probability measure and

$$\widetilde{W}(t) := W(t) + \int_0^t X(s)\, ds = W(t) + \int_0^t s^2\, ds = W(t) + \tfrac{t^3}{3} \tag{1.6.26}$$

defines a Brownian motion $(\widetilde{W}(t) : t \in [0,T])$ under the probability measure Q. By Lemma 6.1.5, equality (1.6.25) yields

$$E_P[Y] = \exp\left(\tfrac{T^5}{10}\right) E_P[W(T)L(T)] = \exp\left(\tfrac{T^5}{10}\right) E_Q[W(T)].$$

By solving for W in (1.6.26) we obtain

$$\exp\left(\tfrac{T^5}{10}\right) E_Q[W(T)] = \exp\left(\tfrac{T^5}{10}\right) E_Q\left[\widetilde{W}(T) - \tfrac{T^3}{3}\right], = -\exp\left(\tfrac{T^5}{10}\right)\tfrac{T^3}{3},$$

and thus $E[Y] = -\exp\left(\tfrac{T^5}{10}\right)\tfrac{T^3}{3}$.

(b) Just replace s^2 by $-s^2$ in (a) and observe that $(-s^2)^2 = s^4$. Then one obtains that

$$E\left[W(T)\exp\left(\int_0^T s^2\, dW(s)\right)\right] = \exp\left(\tfrac{T^5}{10}\right)\tfrac{T^3}{3}.$$

2. (a) We first find the stochastic differential equation which is satisfied by X. For that purpose, define the function

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) := e^x.$$

The function f is in C^2 and we obtain

$$f_x(t,x) = e^x, \quad f_{xx}(x) = e^x.$$

Itô's formula in Theorem 4.5.1 implies

$$dX(t) = df(W(t)) = f_x(W(t))\,dW(t) + \frac{1}{2} f_{xx}(W(t))\,dt$$
$$= X(t)\,dW(t) + \frac{1}{2}X(t)\,dt. \tag{1.6.27}$$

Assume that there exists an adapted stochastic process $(Y(t) : t \in [0,T])$ such that

$$P\left(\int_0^T Y^2(s)\,ds < \infty\right) = 1,$$

and that

$$L(t) := \exp\left(-\int_0^t Y(s)\,dW(s) - \tfrac{1}{2}\int_0^t Y^2(s)\,ds\right)$$

defines a martingale $(L(t) : t \in [0,T])$. Then Girsanov's Theorem 6.1.3 implies

$$Q : \mathscr{A} \to [0,\infty], \qquad Q(A) := \int_A L(T)(\omega)\,P(d\omega) = E_P[\mathbb{1}_A L(T)]$$

is a probability measure and

$$\widetilde{\mathrm{W}}(t) := W(t) + \int_0^t Y(s)\,ds$$

defines a Brownian motion $(\widetilde{\mathrm{W}}(t) : t \in [0,T])$ under the measure Q. It follows from equation (1.6.27) that

$$dX(t) = \frac{1}{2}X(t)\,dt + X(t)\,dW(t)$$
$$= \left(\frac{1}{2}X(t) - X(t)Y(t)\right)dt + X(t)\,d\widetilde{\mathrm{W}}(t).$$

Since $\widetilde{\mathrm{W}}$ is a Brownian motion under Q the stochastic process X is a martingale under Q if the dt-terms vanish and if X is in $\mathscr{H}$, where expectation is taken with respect to the probability measure Q. The dt-terms vanishes if

$$Y(t) = \frac{1}{2} \qquad \text{for all } t \in [0,T].$$

For this choice of the stochastic process $(Y(t) : t \in [0,T])$ we have $\widetilde{\mathrm{W}}(t) = W(t) + \frac{1}{2}t$ for all $t \in [0,T]$. Solving for W we obtain

$$\int_0^T E_Q\left[|X(s)|^2\right]ds = \int_0^T E_Q\big[\exp(2\,\widetilde{\mathrm{W}}(s) - s)\big]\,ds$$
$$= \int_0^T e^{-s} E_Q\big[\exp(2\,\widetilde{\mathrm{W}}(s))\big]\,ds$$
$$= \int_0^T e^{-s}e^{2s}\,ds = e^T - 1 < \infty.$$

Consequently, we obtain $X \in \mathscr{H}$ under the probability measure Q. Since

$$dX(t) = X(t)\, d\widetilde{W}(t) \qquad \text{for all } t \in [0,T],$$

and since $\widetilde{W}$ is a Brownian motion under Q, it follows from Theorem 4.3.1 that the stochastic process X is a martingale under Q.

Finally, we have to check that the stochastic process $(Y(t) : t \in [0,T])$ for $Y(t) = \frac{1}{2}$ satisfies the conditions assumed above:

$$\begin{aligned} &\int_0^T Y^2(s)\, ds = \int_0^T \tfrac{1}{4}\, ds < \infty \quad P\text{-a.s.,} \\ &E_P\left[\exp\left(\tfrac{1}{2}\int_0^T Y^2(s)\, ds\right)\right] = \exp\left(\tfrac{1}{8}T\right) < \infty. \qquad \text{(Novikov's condition)} \end{aligned}$$

Thus, our assumptions in the beginning are verified and we can conclude that X is a martingale under the measure Q.

(b) Since $\widetilde{W}$ is a Brownian motion under Q it follows from Lemma 3.3.4 that

$$E_Q\left[(W(t) + \tfrac{1}{2}t)^6\right] = E_Q\left[(\widetilde{W}(t))^6\right] = 15t^3 \qquad \text{for all } t \in [0,T].$$

3. For each $t \in [0,1]$ define

$$X(t) := \begin{cases} 0, & \text{if } t \in [0, \frac{1}{2}), \\ W(\frac{1}{2}), & \text{if } t \in [\frac{1}{2}, 1]. \end{cases}$$

It follows that $(X(t) : t \in [0,1])$ is an adapted, stochastic process satisfying

$$P\left(\int_0^1 X^2(s)\, ds < \infty\right) = P\Big(\tfrac{1}{2}W^2(\tfrac{1}{2}) < \infty\Big) = 1,$$

$$E\left[\exp\left(\tfrac{1}{2}\int_0^1 X^2(s)\, ds\right)\right] = E\left[\exp\left(\tfrac{1}{4}W^2(\tfrac{1}{2})\right)\right] = \tfrac{1}{\sqrt{\pi}}\int_{\mathbb{R}} e^{\frac{1}{4}u^2} e^{-u^2}\, du = \sqrt{\tfrac{4}{6}} < \infty.$$

Novikov's condition in Theorem 6.1.8 implies that

$$\begin{aligned} L(t) :&= \exp\left(-\int_0^t X(s)\, dW(s) - \tfrac{1}{2}\int_0^t X^2(s)\, ds\right) \\ &= \begin{cases} 1, & \text{if } t \in [0, \frac{1}{2}), \\ \exp\left(-W\left(\frac{1}{2}\right)\left(W(t) - W(\frac{1}{2})\right) - \frac{1}{2}\left(t - \frac{1}{2}\right)W^2(\frac{1}{2})\right), & \text{if } t \in (\frac{1}{2}, t], \end{cases} \\ &= \begin{cases} 1, & \text{if } t \in [0, \frac{1}{2}), \\ \exp\left(-W\left(\frac{1}{2}\right)W(t) - \frac{1}{2}\left(t - \frac{5}{2}\right)W^2(\frac{1}{2})\right), & \text{if } t \in (\frac{1}{2}, t], \end{cases} \end{aligned}$$

defines a martingale $(L(t) : t \in [0,1])$ under the measure P. From Girsanov's Theorem 6.1.3 we can conclude that

$$Q : \mathscr{A} \to [0,\infty], \qquad Q(A) := E_P[\mathbb{1}_A L(1)]$$

is a probability measure and that

$$\widetilde{W}(t) := W(t) + \int_0^t X(s)\,ds = \begin{cases} W(t), & \text{if } t \in [0, \frac{1}{2}), \\ W(t) + (t - \frac{1}{2})W(\frac{1}{2}), & \text{if } t \in [\frac{1}{2}, 1], \end{cases}$$

defines a Brownian motion $(\widetilde{W}(t) : t \in [0,T])$ under the measure Q.

If you wander how to find X, see the solution to Exercise 6.3.2.

4. Let $(U(t) : t \in [0,T])$ denote the solution of the given stochastic differential equation.

   (a) For each $t \in [0,T]$ define

$$X(t) := \tfrac{a-r}{b}.$$

Then $(X(t) : t \in [0,T])$ is an adapted stochastic process satisfying

$$P\left(\int_0^T X^2(s)\,ds < \infty\right) = P\Big(T\left(\tfrac{a-r}{b}\right)^2 < \infty\Big) = 1.$$

For each $t \in [0,T]$ define

$$\begin{aligned} L(t) :&= \exp\left(-\int_0^t X(s)\,dW(s) - \tfrac{1}{2}\int_0^t (X(s))^2\,ds\right) \\ &= \exp\left(-\tfrac{a-r}{b}W(t) - \tfrac{1}{2}\left(\tfrac{a-r}{b}\right)^2 t\right). \end{aligned}$$

Corollary 3.2.4 guarantees that $(L(t) : t \in [0,T])$ is a martingale. Consequently, Girsanov's Theorem 6.1.3 implies that

$$Q : \mathscr{A} \to [0,\infty], \qquad Q(A) := E_P[\mathbb{1}_A L(T)]$$

is a probability measure and

$$\widetilde{W}(t) := W(t) + \int_0^t X(s)\,ds = W(t) + \tfrac{a-r}{b}t \tag{1.6.28}$$

defines a Brownian motion $(\widetilde{W}(t) : t \in [0,T])$ under the measure Q. According to Proposition 5.3.2, the stochastic process $(U(t) : t \in [0,T])$ is of the form

$$U(t) = u_0 \exp\Big((a - \tfrac{1}{2}b^2)t + bW(t)\Big) \qquad \text{for all } t \in [0,T].$$

By applying (1.6.28) we can represent U as

$$\begin{aligned}U(t) &= u_0 \exp\big((a-\tfrac{1}{2}b^2)t + bW(t)\big)\\ &= u_0 \exp\Big((a-\tfrac{1}{2}b^2)t + b\,\widetilde{\mathrm{W}}(t) - (a-r)t\Big)\\ &= u_0 \exp\Big((r-\tfrac{1}{2}b^2)t + b\,\widetilde{\mathrm{W}}(t)\Big) \qquad \text{for all } t\in[0,T].\end{aligned}$$

(b) For each $t \geqslant 0$ the random variable $\hat{U}(t)$ is $\mathscr{F}_t^W$-measurable since it is the image of the continuous function

$$f\colon\ \mathbb{R}\to\mathbb{R}, \qquad f(x) = e^{-rt}u_0 \exp\big((a-\tfrac{1}{2}b^2)t + bx\big)$$

applied to the $\mathscr{F}_t^W$-measurable random variable $W(t)$. For each $t\in[0,T]$ we have

$$\begin{aligned}E_Q\Big[\big|\hat{U}(t)\big|\Big] &= e^{-rt}E_Q\Big[|u_0)|\exp\Big((r-\tfrac{1}{2}b^2)t + b\,\widetilde{\mathrm{W}}(t)\Big)\Big]\\ &= e^{-rt}u_0\exp\big((r-\tfrac{1}{2}b^2)t\big)\,E_Q\Big[\exp\Big(b\,\widetilde{\mathrm{W}}(t)\Big)\Big]\\ &= e^{-rt}u_0\exp\big((r-\tfrac{1}{2}b^2)t\big)\exp\big(\tfrac{1}{2}b^2t\big)\\ &< \infty.\end{aligned}$$

For each $0\leqslant s\leqslant t$ we obtain

$$\begin{aligned}E_Q[\hat{U}(t)|\mathscr{F}_s] &= e^{-rt}E_Q\Big[u_0\exp\Big((r-\tfrac{1}{2}b^2)t + b\,\widetilde{\mathrm{W}}(t)\Big)\,|\mathscr{F}_s\Big]\\ &= e^{-\frac{1}{2}b^2t}u_0E_Q\Big[\exp\Big(b(\widetilde{\mathrm{W}}(t)-\widetilde{\mathrm{W}}(s)+\widetilde{\mathrm{W}}(s))\Big)\,|\mathscr{F}_s\Big]\\ &= e^{-\frac{1}{2}b^2t}u_0\exp\Big(b\,\widetilde{\mathrm{W}}(s)\Big)\,E_Q\Big[\exp\Big(b(\widetilde{\mathrm{W}}(t)-\widetilde{\mathrm{W}}(s))\Big)\Big]\\ &= e^{-\frac{1}{2}b^2t}u_0\exp\Big(b\,\widetilde{\mathrm{W}}(s)\Big)\exp\big(\tfrac{1}{2}b^2(t-s)\big)\\ &= u_0\exp\Big(b\,\widetilde{\mathrm{W}}(s) + (r-\tfrac{1}{2}b^2)s\Big)\,e^{-rs}\\ &= \hat{U}(s).\end{aligned}$$

Easier but less educational is to cite Corollary 3.2.4.

(c) By using the representation of $U(T)$ derived in (a) and the fact that $\widetilde{\mathrm{W}}$ is a Brownian motion under Q, we obtain

$$\begin{aligned}
E_Q[C] &= E_Q\left[\mathbb{1}_{\left\{u_0 \exp\left((r-\frac{1}{2}b^2)T+b\widetilde{\mathrm{W}}(T)\right)\right\}\geqslant K}\right] \\
&= Q\left(u_0 \exp\left((r-\tfrac{1}{2}b^2)T+b\,\widetilde{\mathrm{W}}(T)\right) > K\right) \\
&= Q\left(\widetilde{\mathrm{W}}(T) \geqslant \frac{1}{b}\left(\ln\frac{K}{u_0}-(r-\tfrac{1}{2}b^2)T\right)\right) \\
&= Q\left(\widetilde{\mathrm{W}}(1) \geqslant \frac{1}{b\sqrt{T}}\left(\ln\frac{K}{u_0}-(r-\tfrac{1}{2}b^2)T\right)\right) \\
&= F_N\left(\frac{1}{b\sqrt{T}}\left(\ln\frac{u_0}{K}+(r-\tfrac{1}{2}b^2)T\right)\right),
\end{aligned}$$

where F_N denotes the probability distribution function of the standard normal distribution.

5. By defining for each $t\in[0,T]$

$$X(t) := \exp(W(t))\mathbb{1}_{\{|W(t)|\leqslant 1\}}$$

we obtain an adapted stochastic process $(X(t): t\in[0,T])$. Since $|X(t)|\leqslant \exp(1)$ P-a.s. for all $t\in[0,T]$ we conclude

$$\begin{aligned}
&P\left(\int_0^T X^2(s)\,ds<\infty\right)=1,\\
&E\left[\exp\left(\tfrac{1}{2}\int_0^T X^2(s)\,ds\right)\right]\leqslant \exp\left(\tfrac{1}{2}e^2T\right)<\infty.
\end{aligned}$$

Novikov's condition in Theorem 6.1.8 implies that

$$\begin{aligned}
L(t):&=\exp\left(-\int_0^t X(s)\,dW(s)-\tfrac{1}{2}\int_0^t X^2(s)\,ds\right)\\
&=\exp\left(-\int_0^t \exp(W(s))\mathbb{1}_{\{|W(s)|\leqslant 1\}}\,ds-\tfrac{1}{2}\int_0^t \exp(2W(s))\mathbb{1}_{\{|W(s)|\leqslant 1\}}\,ds\right)
\end{aligned}$$

defines a martingale $(L(t): t\in[0,T])$. Girsanov's Theorem 6.1.3 yields that

$$Q:\mathscr{A}\to[0,\infty],\qquad Q(A):=E_P[\mathbb{1}_A L(T)]$$

is a probability measure and

$$\widetilde{\mathrm{W}}(t):=W(t)+\int_0^t X(s)\,ds=W(t)+\int_0^t e^{W(s)}\mathbb{1}_{\{|W(s)|\leqslant 1\}}\,ds$$

defines a Brownian motion $(\widetilde{\mathrm{W}}(t): t\in[0,T])$ under the measure Q.

(b) Since $\widetilde{\mathrm{W}}$ is a Brownian motion under Q it follows from Lemma 3.3.4 that

$$E_Q\left[\left|W(t)+\int_0^t e^{W(s)}\mathbb{1}_{\{|W(s)|\leqslant 1\}}\,ds\right|^6\right]=E_Q\left[(\widetilde{\mathrm{W}}(t))^6\right]=15t^3.$$

(c) Proposition 3.3.1 enables us to conclude

$$E_P[W(\tau)]=-1\,P(W(\tau)=-1)+1\,P(W(\tau)=1)=-1\frac{1}{2}+1\frac{1}{2}=0.$$

Thus, we obtain

$$\begin{aligned}E_P[\widetilde{\mathrm{W}}(\tau)]&=E_P\left[W(\tau)+\int_0^\tau e^{W(s)}\mathbb{1}_{\{|W(s)|\leqslant 1\}}\,ds\right]\\&=E_P\left[\int_0^\tau e^{W(s)}\mathbb{1}_{\{|W(s)|\leqslant 1\}}\,ds\right]\\&=E_P\left[\int_0^\tau e^{W(s)}\,ds\right].\end{aligned}\tag{1.6.29}$$

On the other hand Itô's formula in Theorem 4.5.1 yields

$$e^{W(t)}=1+\int_0^t e^{W(s)}\,dW(s)+\tfrac{1}{2}\int_0^t e^{W(s)}\,ds,\tag{1.6.30}$$

which can be written as

$$\int_0^\tau e^{W(s)}\,ds=2e^{W(\tau)}-2-2\int_0^\tau e^{W(s)}\,dW(s).\tag{1.6.31}$$

Define $M(t):=\int_0^t e^{W(s)}\,dW(s)$ for all $t\geqslant 0$. Since Proposition 3.2.1 yields

$$\int_0^T E_P\left[e^{2W(s)}\right]\,ds=\int_0^T e^{2s}\,ds<\infty,$$

it follows that $(e^{W(t)}:t\in[0,T])\in\mathscr{H}$ (for each $T>0$) and thus, M is a martingale according to Theorem 4.3.1. The optional sampling Theorem 2.3.1 implies

$$0=E_P[M(0)]=E_P[M(t\wedge\tau)]\qquad\text{for all } t\geqslant 0.$$

Moreover, we have $M(t\wedge\tau)\to M(\tau)$ P-a.s. as $t\to\infty$ and by using (1.6.30) we obtain

$$\begin{aligned}|M(t\wedge\tau)|&=\left|e^{W(t\wedge\tau)}-1-\tfrac{1}{2}\int_0^{t\wedge\tau}e^{W(s)}\,ds\right|\\&\leqslant\left|e^{W(t\wedge\tau)}\right|+1+\tfrac{1}{2}\left|\int_0^{t\wedge\tau}e^{W(s)}\,ds\right|\leqslant 1+1+\tfrac{1}{2}(t\wedge\tau)\leqslant 2+\tfrac{1}{2}\tau.\end{aligned}$$

Since Proposition 3.3.1 guarantees $E_P[\tau] < \infty$, Lebesgue's dominated convergence theorem implies

$$E_P[M(\tau)] = E_P[\lim_{t\to\infty} M(t\wedge\tau)] = \lim_{t\to\infty} E_P[M(t\wedge\tau)] \ = 0.$$

Thus, we have $\int_0^\tau e^{W(s)}\, dW(s) = 0$, and taking expectation in (1.6.31) results in

$$E_P\left[\int_0^\tau e^{W(s)}\, ds\right] = 2\tfrac{1}{2}\left(e^{+1} + e^{-1}\right) - 2,$$

which together with (1.6.29) implies

$$E_P[\widetilde{\mathrm{W}}(\tau)] = \left(e^{+1} + e^{-1}\right) - 2.$$

(d) Corollary 3.2.4 guarantees that

$$M(t) := \exp\left(\widetilde{\mathrm{W}}(t) - \tfrac{1}{2}t\right)$$

defines a martingale $(M(t) : t \geqslant 0)$ under Q. Thus, the optional sampling Theorem 2.3.1 implies

$$1 = E_Q[M(0)] = E_Q[M(t\wedge\sigma)] \qquad \text{for all } t \geqslant 0. \tag{1.6.32}$$

Moreover, we have $M(t\wedge\sigma) \to M(\sigma)$ Q-a.s. as $t\to\infty$ and it follows that

$$\begin{aligned}&|M(t\wedge\sigma)| \leqslant \exp\left(\widetilde{\mathrm{W}}(t\wedge\sigma) - \tfrac{1}{2}t\wedge\sigma\right) \leqslant \exp\left(\widetilde{\mathrm{W}}(t\wedge\sigma)\right) \leqslant \exp(1),\\ &E_Q[\exp(1)] = \exp(1) < \infty.\end{aligned}$$

By using (1.6.32), Lebesgue's theorem of dominated convergence implies

$$E_Q[M(\sigma)] = E_Q[\lim_{t\to\infty} M(t\wedge\sigma)] = \lim_{t\to\infty} E_Q[M(t\wedge\sigma)] = 1. \tag{1.6.33}$$

On the other hand, since $\widetilde{\mathrm{W}}(\sigma) = 1$ we have

$$E_Q[M(\sigma)] = E_Q\left[e^{1-\frac{1}{2}\sigma}\right],$$

which together with equation (1.6.33) result in

$$E_Q\left[e^{-\sigma/2}\right] = \frac{1}{e}.$$

A.7. Solution Chapter 7

1. (a) If we define $\Phi(s) = 1$ for all $s \in [0, T]$ it follows

$$W(T) = \int_0^T dW(s) = E[X] + \int_0^T \Phi(s)\, dW(s).$$

The stochastic process $(\Phi(s) : s \in [0, T])$ is $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$-adapted since it is deterministic and in $\mathscr{H}$ since

$$\int_0^T E[|\Phi(s)|^2]\, ds = \int_0^T ds = T < \infty.$$

(b) By partial integration, see Lemma 5.2.4, we obtain

$$\int_0^T s\, dW(s) = TW(T) - \int_0^T W(s)\, ds.$$

Consequently, if we define $\Phi(s) = T - s$ for all $s \in [0, T]$ we obtain

$$\int_0^T W(s)\, ds = TW(T) - \int_0^T s\, dW(s) = \int_0^T (T - s)\, dW(s) = E[X] + \int_0^T \Phi(s)\, dW(s).$$

The stochastic process $(\Phi(s) : s \in [0, T])$ is $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$-adapted since it is deterministic and in $\mathscr{H}$ since

$$\int_0^T E[|\Phi(s)|^2]\, ds = \int_0^T (T - s)^2\, ds = \tfrac{1}{3}T^3 < \infty.$$

(c) Define the function

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) := x^2.$$

The function f is in C^2 and we obtain that

$$f_x(x) = 2x, \quad f_{xx}(x) = 2.$$

Itô's formula in Theorem 4.5.1 implies

$$d(W^2(t)) = f_x(W(t))\, dW(t) + \frac{1}{2} f_{xx}(W(t))\, dt = 2W(t)\, dW(t) + \frac{1}{2} 2\, dt.$$

Consequently, if we define $\Phi(s) = 2W(s)$ for all $s \in [0, T]$ we obtain

$$W^2(T) = T + \int_0^T 2W(s)\, dW(s) = E[X] + \int_0^T \Phi(s)\, dW(s).$$

The stochastic process $(\Phi(s) : s \in [0,T])$ is $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$-adapted since $\Phi(s) = f(W(s))$ for every $s \in [0,T]$ for the continuous function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2x$, and it is in $\mathscr{H}$ since

$$\int_0^T E[|\Phi(s)|^2]\,ds = \int_0^T E[|2W(s)|^2]\,ds = 4\int_0^T s\,ds = 2T^2 < \infty.$$

(d) Define the function

$$f : [0,T] \times \mathbb{R} \to \mathbb{R} \qquad f(t,x) := tx^2.$$

The function f is in $C^{1,2}$ and we obtain

$$f_t(t,x) = x^2, \quad f_x(t,x) = 2tx, \quad f_{xx}(t,x) = 2t.$$

Itô's formula in Theorem 4.5.6 implies

$$\begin{aligned} df(t,W(t)) &= \big(f_t(t,W(t)) + \tfrac{1}{2} f_{xx}(t,W(t))\big)\,dt + f_x(t,W(t))\,dW(t) \\ &= (W^2(t) + t)\,dt + 2tW(t)\,dW(t). \end{aligned}$$

Consequently, if we define $\Phi(s) = 2(T-s)W(s)$ for all $s \in [0,T]$, part (c) implies

$$\begin{aligned} \int_0^T W^2(s)\,ds &= TW^2(T) - \int_0^T s\,ds - \int_0^T 2sW(s)\,dW(s) \\ &= T\left(T + \int_0^T 2W(s)\,dW(s)\right) - \tfrac{1}{2}T^2 - \int_0^T 2sW(s)\,dW(s) \\ &= \tfrac{1}{2}T^2 + \int_0^T 2(T-s)W(s)\,dW(s) \\ &= E[X] + \int_0^T \Phi(s)\,dW(s). \end{aligned}$$

The stochastic process $(\Phi(s) : s \in [0,T])$ is $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$-adapted since $\Phi(s) = f(W(s))$ for every $s \in [0,T]$ for the continuous function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2(T-s)x$, and it is in $\mathscr{H}$ since

$$\int_0^T E[|\Phi(s)|^2]\,ds = \int_0^T 4(T-s)E[|W(s)|^2]\,ds = \int_0^T 4(T-s)^2 s\,ds < \infty.$$

(e) Define the function

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) := x^3.$$

The function f is in C^2 and we obtain

$$f_x(x) = 3x^2, \quad f_{xx}(x) = 6x.$$

Itô's formula in Theorem 4.5.1 implies

$$\begin{aligned} d(W^3(t)) &= f_x(W(t))\,dW(t) + \frac{1}{2} f_{xx}(W(t))\,dt \\ &= 3W^2(t)\,dW(t) + 3W(t)\,dt. \end{aligned}$$

Consequently, if we define $\Phi(s) = 3(T-s) + 3W^2(s)$ for all $s \in [0,T]$ part (b) implies

$$\begin{aligned} W(T)^3 &= \int_0^T 3W(s)\,ds + \int_0^T 3W^2(s)\,dW(s) \\ &= \int_0^T 3\Big((T-s) + W^2(s)\Big)\,dWs) \\ &= E[W^3(T)] + \int_0^T \Phi(s)\,dW(s). \end{aligned}$$

The stochastic process $(\Phi(s) : s \in [0,T])$ is $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$-adapted since $\Phi(s) = f(W(s))$ for every $s \in [0,T]$ for the continuous function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3(T-s) + 3x^2$, and it is in $\mathscr{H}$ since

$$\begin{aligned} \int_0^T E[|\Phi(s)|^2]\,ds &= \int_0^T E\big[9(T-s)^2 + 18(T-s)W^2(s) + 9W^4(s)\big]\,ds \\ &= \int_0^T \big(9(T-s)^2 + 18(T-s)s + 27s^2\big)\,ds < \infty. \end{aligned}$$

2. (a) By Jensen's inequality for conditional expectation one obtains

$$E[|M(T)|^2] = E\big[\,|E[Z|\mathscr{F}_T]|^2\,\big] \leqslant E\big[E[|Z|^2\,|\mathscr{F}_T]\big] = E[|Z|^2] < \infty.$$

(b) (i) Since $(W^2(t) - t : t \in [0,T])$ is a martingale according to Corollary 3.2.4, we obtain for every $0 \leqslant t \leqslant T$:

$$M(t) = E\left[W^2(T) - T + T|\mathscr{F}_t\right] = E\left[W^2(T) - T|\mathscr{F}_t\right] + T = W^2(t) - t + T.$$

Define $\Phi(s) = 2W(s)$ for each $s \in [0,T]$. Part (c) of Exercise 7.3.1 yields for each $t \in [0,T]$

$$M(t) = W^2(t) - t + T = t + \int_0^t 2W(s)\,dW(s) - t + T = E[M(0)] + \int_0^t \Phi(s)\,dW(s).$$

The stochastic process $(\Phi(s) : s \in [0,T])$ is $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$-adapted since $\Phi(s) = f(W(s))$ for every $s \in [0,T]$ for the continuous function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2x$, and it is in $\mathscr{H}$ since

$$\int_0^T E[|\Phi(s)|^2]\,ds = \int_0^T 4s\,ds = 2T^2 < \infty.$$

(ii) Part (d) in Exercise 2.5.3 guarantees that $(W(t)^3 - 3tW(t) : t \in [0,T])$ is a martingale. Together with the fact that W is also a martingale this implies

$$M(t) = E\left[W^3(T) - 3TW(T) + 3TW(T)|\mathscr{F}_t\right] = W^3(t) - 3tW(t) + 3TW(t).$$

We define $\Phi(s) = 3(T-s) + 3W^2(s)$ for all $s \in [0,T]$. Part (e) of Exercise 7.3.1 yields for each $t \in [0,T]$

$$\begin{aligned}
M(t) &= W^3(t) + 3(T-t)W(t) \\
&= \int_0^t 3\Big((t-s) + W^2(s)\Big)\, dW(s) + 3(T-t)W(t) \\
&= \int_0^t 3\Big((T-s) + W^2(s)\Big)\, dW(s) \\
&= E[M(0)] + \int_0^t \Phi(s)\, dW(s).
\end{aligned}$$

In Part (e) of Exercise 7.3.1 it is shown that the stochastic process $(\Phi(s) : s \in [0,T])$ is $\{\mathfrak{F}_t^W\}_{t\geqslant 0}$-adapted and in $\mathscr{H}$.

3. a) Theorem 4.3.1 implies for every $0 \leqslant s \leqslant t$ that

$$\begin{aligned}
E[X(t)|\mathscr{F}_s] &= X(0) + E\left[\int_0^t \Upsilon(u)\, du \,\middle|\, \mathscr{F}_s\right] + E\left[\int_0^t \Phi(u)\, dW(u) \,\middle|\, \mathscr{F}_s\right] \\
&= X(0) + \int_0^s \Upsilon(u)\, du + \int_0^s \Phi(u)\, dW(u) + E\left[\int_s^t \Upsilon(u)\, du \,\middle|\, \mathscr{F}_s\right] \\
&= X(s) + E\left[\int_s^t \Upsilon(u)\, du \,\middle|\, \mathscr{F}_s\right].
\end{aligned}$$

Since X is a martingale we obtain

$$E\left[\int_s^t \Upsilon(u)\, du | \mathscr{F}_s^W\right] = 0 \qquad \text{for all } 0 \leqslant s \leqslant t.$$

Interchanging the conditional expectation and integration results in

$$\int_s^t E\left[\Upsilon(u)|\mathscr{F}_s^W\right]\, du = 0 \qquad \text{for all } 0 \leqslant s \leqslant t.$$

By differentiating with respect to t we derive that

$$E[\Upsilon(t)|\mathscr{F}_s^W] = 0 \qquad \text{for all } 0 \leqslant s \leqslant t.$$

(b) For t fixed and $s_n \nearrow t$ it follows from (a) and the hint that

$$\Upsilon(t) = E\left[\Upsilon(t)|\mathscr{F}_t^W\right] = \lim_{s_n \uparrow t} E\left[\Upsilon(t)|\mathscr{F}_{s_n}^W\right] = 0.$$

B

FAQ

Symbols

$\mathbb{N}$	$\{1, 2, 3, \dots\}$
$\mathbb{N}_0$	$\{0, 1, 2, 3, \dots\}$
$\mathbb{R}$	real numbers
$\mathbb{R}_+$	$[0, \infty)$
$N(0, s^2)$	normal distribution with expectation 0 and variance s^2
$E[X]$	expectation of the random variable X (with respect to the standard measure P)
$\mathrm{Var}[X]$	variance of the random variable X (with respect to the standard measure P)
$\mathrm{Cov}(X, Y)$	covariance of two random variables X and Y
$N(a, V)$	normal distribution with expectation $a \in \mathbb{R}^n$ and covariance matrix $V \in \mathbb{R}^{n\times n}$
$a \wedge b$	$\min\{a, b\}$
$a \vee b$	$\max\{a, b\}$
$\frac{\partial}{\partial x_i}$	differential operator with respect to the i-th argument
$(a)^+$	$\max\{a, 0\}$
a^T	transpose of a vector $a \in \mathbb{R}^n$
$(c_{ij})_{i,j=1}^n$	a matrix in $\mathbb{R}^{n\times n}$ with entries c_{ij} (i-th row, j-th column)
$\mathscr{P}(\Omega)$	$\{A \subseteq \Omega\}$ = Powerset
$\mathfrak{B}(\mathbb{R}^d)$	Borel σ-algebra in $\mathbb{R}^d$
Id_d	identity matrix in $\mathbb{R}^{d\times d}$

Notations

something P-a.s.	P(something) $= 1$ (here P is a probability measure)
X random variable (r.v.)	$X : \Omega \to \mathbb{R}$ is $\mathscr{A}$-measurable
$f : [0, T] \to \mathbb{R}$ a deterministic function	$f(t)$ is a fixed number in $\mathbb{R}$ for each $t \in [0, T]$
$(Y(t) : t \in [0, T])$ a stochastic process	$Y(t) : \Omega \to \mathbb{R}$ is a random variable for each $t \in [0, T]$
A:=B	A is defined by B

$X \stackrel{\mathscr{D}}{=} Y$	random variables X and Y have the same distribution

Conventions

X,Y,S,Υ, Φ	random variables or stochastic processes (usually capital letters)
g,f,h	deterministic functions (usually small letters)
α, β, μ,σ	sometimes constants, sometimes deterministic functions

Bibliography

[1] P. Billingsley. *Probability and measure.* New Yorks:John Wiley & Sons Inc., 1995.

[2] T. Björk. *Arbitrage theory in continuous time.* Oxford: Oxford University Press, 2004.

[3] Z. Brzeźniak and T. Zastawniak. *Basic stochastic processes. A course through exercises.* Berlin: Springer, 1999.

[4] C. Dellacherie and P.-A. Meyer. *Probabilities and potential.* Amsterdam: North-Holland Publishing Co., 1978.

[5] C. Dellacherie and P.-A. Meyer. *Probabilities and potential. B.* Amsterdam: North-Holland Publishing Co., 1982.

[6] R. Durrett. *Stochastic calculus. A practical introduction.* Boca Raton, FL: CRC Press, 1996.

[7] G. R. Grimmett and D. R. Stirzaker. *Probability and random processes.* New York: The Clarendon Press Oxford University Press, 1992.

[8] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.* Amsterdam: North-Holland Publishing Co., 1989.

[9] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus.* New York: Springer, 1991.

[10] F. C. Klebaner. *Introduction to stochastic calculus with applications.* London: Imperial College Press, 2005.

[11] P. Kloeden and E. Platen. *Numerical solution of stochastic differential equations.* Berlin: Springer, 1992.

[12] H.-H. Kuo. *Introduction to stochastic integration.* New York, NJ: Springer , 2006.

[13] T. Mikosch. *Elementary stochastic calculus with finance in view.* Singapore: World Scientific, 1998.

[14] P. Mörters. *Stochastic analysis.* Lecture Notes, 2000.

[15] P. Mörters and Y. Peres. *Brownian motion. With an appendix by Oded Schramm and Wendelin Werner.* Cambridge: Cambridge University Press , 2010.

[16] B. Oksendal. *Stochastic differential equations. An introduction with applications.* Berlin: Springer, 2010.

[17] P. E. Protter. *Stochastic integration and differential equations.* Berlin: Springer, 2004.

[18] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Berlin: Springer, 1999.

[19] L.C.G. Rogers and D. Williams. *Diffusions, Markov processes and martingales. Vol. 1: Foundations.* . Cambridge: Cambridge University Press, 2000.

[20] L.C.G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2: Itô calculus.* Cambridge: Cambridge University Press , 2000.

[21] T. Seppäläinen. *Basics of Stochastic Analysis.* lecture notes, 2012.

[22] J. M. Steele. *Stochastic calculus and financial applications.* New York, NY: Springer, 2001.

[23] R. van Handel. *Stochastic Calculus, Filtering, and Stochastic Control.* lecture notes, 2007.

[24] H. von Weizsäcker and G. Winkler. *Stochastic integrals. An introduction.* Friedr. Vieweg & Sohn, Braunschweig, 1990.

Index