Universitext

Paolo Baldi

# Stochastic Calculus

## An Introduction Through Theory and Exercises

Springer

Universitext

# Universitext

**Series editors**

Sheldon Axler
*San Francisco State University*

Carles Casacuberta
*Universitat de Barcelona*

Angus MacIntyre
*Queen Mary, University of London*

Kenneth Ribet
*University of California, Berkeley*

Claude Sabbah
*École polytechnique, CNRS, Université Paris-Saclay, Palaiseau*

Endre Süli
*University of Oxford*

Wojbor A. Woyczyński
*Case Western Reserve University*

*Universitext* is a series of textbooks that presents material from a wide variety of mathematical disciplines at master's level and beyond. The books, often well class-tested by their author, may have an informal, personal even experimental approach to their subject matter. Some of the most successful and established books in the series have evolved through several editions, always following the evolution of teaching curricula, to very polished texts.

Thus as research topics trickle down into graduate-level teaching, first textbooks written for new, cutting-edge courses may make their way into *Universitext.*

More information about this series at http://www.springer.com/series/223

Paolo Baldi

# Stochastic Calculus

## An Introduction Through Theory and Exercises

Springer

Paolo Baldi
Dipartimento di Matematica
Università di Roma "Tor Vergata"
Roma, Italy

ISSN 0172-5939 ISSN 2191-6675 (electronic)
Universitext
ISBN 978-3-319-62225-5 ISBN 978-3-319-62226-2 (eBook)
DOI 10.1007/978-3-319-62226-2

Library of Congress Control Number: 2017948389

Mathematics Subject Classification (2010): 60H10, 60H05, 60H30, 60G42, 60G44

© Springer International Publishing AG 2017
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*To the memory of Gianni Di Masi*

# Preface

Courses in Stochastic Calculus have in the last two decades changed their target audience. Once this advanced part of mathematics was of interest mainly to postgraduates intending to pursue an academic research career, but now many professionals cannot do without the ability to manipulate stochastic models.

The aim of this book is to provide a tool in this direction, starting from a basic probability background (with measure theory, however). The intended audience should, moreover, have serious mathematical bases.

The entire content of this book should provide material for a two-semester class. My experience is that Chaps. 2–9 provide the material for a course of 72 h, including the time devoted to the exercises.

To be able to manipulate these notions requires the reader to acquire not only the elements of the theory but also the ability to work with them and to understand their applications and their connections with other branches of mathematics.

The first of these objectives is taken care of (or at least I have tried to…) by the development of a large set of exercises which are provided together with extensive solutions. Exercises are hardwired with the theory and are intended to acquaint the reader with the full meaning of the theoretical results. This set of exercises with their solution is possibly the most original part of this work.

As for the applications, this book develops two kinds.

The first is given by modeling applications. Actually there are very many situations (in finance, telecommunications, control, …) where stochastic processes, and in particular diffusions, are a natural model. In Chap. 13 we develop financial applications, currently a rapidly growing area.

Stochastic processes are also connected with other fields in pure mathematics and in particular with PDEs. Knowledge of diffusion processes contributes to a better understanding of some aspects of PDE problems and, conversely, the solution of PDE problems can lead to the computation of quantities of interest related to diffusion processes. This two-way tight connection between processes and PDEs is developed in Chap. 10. Further interesting connections between diffusion processes and other branches of mathematics (algebraic structures, differential geometry, ...) are unfortunately not present in this text, primarily for reasons of space.

The first goal is to make the reader familiar with the basic elements of stochastic processes, such as Brownian motion, martingales, and Markov processes, so that it is not surprising that stochastic calculus proper begins almost in the middle of the book.

Chapters 2–3 introduce stochastic processes. After the description of the general setting of a continuous time stochastic process that is given in Chap. 2, Chap. 3 introduces the prototype of diffusion processes, that is Brownian motion, and investigates its, sometimes surprising, properties.

Chapters 4 and 5 provide the main elements on conditioning, martingales, and their applications in the investigation of stochastic processes. Chapter 6 is about Markov processes.

From Chap. 7 begins stochastic calculus proper. Chapters 7 and 8 are concerned with stochastic integrals and Ito's formula. Chapter 9 investigates stochastic differential equations, Chap. 10 is about the relationship with PDEs. After the detour on numerical issues related to diffusion processes of Chap. 11, further notions of stochastic calculus are investigated in Chap. 12 (Girsanov's theorem, representation theorems of martingales) and applications to finance are the object of the last chapter.

The book is organized in a linear way, almost every section being necessary for the understanding of the material that follows. The few sections and the single chapter that can be avoided are marked with an asterisk.

This book is based on courses that I gave first at the University of Pisa, then at Roma "Tor Vergata" and also at SMI (Scuola Matematica Interuniversitaria) in Perugia. It has taken advantage of the remarks and suggestions of many cohorts of students and of colleagues who tried the preliminary notes in other universities. The list of the people I am indebted to is a long one, starting with the many students that have suffered under the first versions of this book. G. Letta was very helpful in clarifying to me quite a few complicated situations. I am also indebted to C. Costantini, G. Nappo, M. Pratelli, B. Trivellato, and G. Di Masi for useful remarks on the earlier versions.

I am also grateful for the list of misprints, inconsistencies, and plain mistakes pointed out to me by M. Gregoratti and G. Guatteri at Milano Politecnico and B. Pacchiarotti at my University of Roma "Tor Vergata". And mainly I must mention L. Caramellino, whose class notes on mathematical finance were the main source of Chap. 13.

Roma, Italy Paolo Baldi

# Contents

# Common Notations

## Real, complex numbers, $\mathbb{R}^m$

| | |
|---|---|
| $x \vee y$ | $= \max(x, y)$ the largest of the real numbers $x$ and $y$ |
| $x \wedge y$ | $= \min(x, y)$ the smallest of the real numbers $x$ and $y$ |
| $\langle x, y\rangle$ | The scalar product of $x, y \in \mathbb{R}^m$ |
| $x^+, x^-$ | The positive and negative parts of $x \in \mathbb{R}$: $x^+ = \max(x, 0)$, $x^- = \max(-x, 0)$ |
| $\|x\|$ | According to the context the absolute value of the real number $x$, the modulus of the complex number $x$, or the euclidean norm of $x \in \mathbb{R}^m$ |
| $\Re z$, $\Im z$ | The real and imaginary parts of $z \in \mathbb{C}$ |
| $B_R(x)$ | $\{y \in \mathbb{R}^m, \|y - x\| < R\}$, the open ball centered at $x$ with radius $R$ |
| $A^*$, $\operatorname{tr} A$, $\det A$ | The transpose, trace, determinant of matrix $A$ |

## Derivatives

| | |
|---|---|
| $\dfrac{\partial f}{\partial x_i}, f_{x_i}$ | Partial derivatives |
| $f'$ | Derivative, gradient |
| $f''$ | Second derivative, Hessian |

## Functional spaces

| | |
|---|---|
| $M_b(E)$ | Real bounded measurable functions on the topological space $E$ |
| $\\|f\\|_\infty$ | $= \sup_{x\in E} \|f(x)\|$ if $f \in M_b(E)$ |

| | |
|---|---|
| $C_b(\mathbb{R}^m)$ | The Banach space of real bounded continuous functions on $\mathbb{R}^m$ endowed with the norm $\| \|_\infty$ |
| $C_K(\mathbb{R}^m)$ | The subspace of $C_b(\mathbb{R}^m)$ of the continuous functions with compact support |
| $C^k(D)$ | $k$ times differentiable functions on the open set $D \subset \mathbb{R}^m$ |
| $C^{2,1}(D \times [0,T[)$ | Functions twice differentiable in $x \in D$ and once in $t \in [0,T[$ |
| $\mathscr{C}([0,T],\mathbb{R}^m)$ | The vector space of continuous maps $\gamma : [0,T] \to \mathbb{R}^m$ endowed with the sup norm $\| \|_\infty$. It is a Banach space |
| $\mathscr{C}(\mathbb{R}^+,\mathbb{R}^m)$ | The vector space of continuous maps $\gamma : \mathbb{R}^+ \to \mathbb{R}^m$, endowed with the topology of uniform convergence on compact sets of $\mathbb{R}^+$ |
| $\mathscr{C}_x([0,T],\mathbb{R}^m)$ | The vector space of continuous maps $\gamma : [0,T] \to \mathbb{R}^m$ such that $\gamma_0 = x$, endowed with the topology of uniform convergence |

# Chapter 1
# Elements of Probability

In this chapter we recall the basic facts in probability that are required for the investigation of the stochastic processes that are the object of the subsequent chapters.

## 1.1 Probability spaces, random variables

A *measurable space* is a pair $(E, \mathscr{E})$ where

- $E$ is a set;
- $\mathscr{E}$ is a $\sigma$-algebra of subsets of $E$.

A *probability space* is a triple $(\Omega, \mathscr{F}, \mathrm{P})$ where

- $\Omega$ is a set;
- $\mathscr{F}$ is a $\sigma$-algebra of subsets of $\Omega$ (i.e. $(\Omega, \mathscr{F})$ is a measurable space);
- $\mathrm{P}$ is a positive measure on $\mathscr{F}$ such that $\mathrm{P}(\Omega) = 1$.

The elements of $\mathscr{F}$ are called *events*, P a *probability* (measure).
If $(\Omega, \mathscr{F}, \mathrm{P})$ is a probability space and $(E, \mathscr{E})$ a measurable space, a *random variable* (r.v.) is a measurable function $X : \Omega \to E$, i.e. such that $X^{-1}(A) \in \mathscr{F}$ whenever $A \in \mathscr{E}$. It is a *real random variable* if $E = \mathbb{R}$ and $\mathscr{E} = \mathscr{B}(\mathbb{R})$ (Borel sets of $\mathbb{R}$).

For a real r.v. $X$ on $(\Omega, \mathscr{F}, \mathrm{P})$ we can speak of an integral. If $X$ is integrable we shall usually write $\mathrm{E}[X]$ instead of $\int X \, d\mathrm{P}$. $\mathrm{E}[X]$ is the *mathematical expectation*. Sometimes the terms *mean* or *mean value* are also used.

If $X = (X_1, \dots, X_m)$ is $(\mathbb{R}^m, \mathscr{B}(\mathbb{R}^m))$-valued we shall speak of $m$-dimensional r.v.'s and, if the components $X_i$ are integrable, we define

$$\mathrm{E}[X] = (\mathrm{E}[X_1], \dots, \mathrm{E}[X_m]) \ .$$

We say that two r.v.'s $X, Y$ are equivalent if $\mathrm{P}(X = Y) = 1$.

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_1

The spaces $L^p$, $1 \le p < +\infty$, and $L^\infty$ are defined as usual as well as the norms $\|X\|_p$ and $\|X\|_\infty$. Recall in particular that $L^p$ is the set of equivalence classes of r.v.'s $X$ such that $\|X\|_p = \mathrm{E}[|X|^p]^{1/p} < +\infty$. It is worth pointing out that $L^p$ is a set of equivalence classes and not of r.v.'s; this distinction will sometimes be necessary even if often, in order to simplify the statements, we shall identify a r.v. and its equivalence class.

For a real r.v. $X$, let us denote by $X = X^+ - X^-$ its decomposition into positive and negative parts. Recall that both $X^+ = X \vee 0$ and $X^- = (-X) \vee 0$ are positive r.v.'s. $X$ is said to be *lower semi-integrable* (l.s.i.) if $X^-$ is integrable. In this case it is possible to define the mathematical expectation

$$\mathrm{E}[X] = \mathrm{E}[X^+] - \mathrm{E}[X^-] ,$$

which is well defined, even if it can take the value $+\infty$. Of course a positive r.v. is always l.s.i.

The following classical inequalities hold.

*Jensen's inequality*. Let $X$ be an $m$-dimensional integrable r.v. and $\Phi : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ a lower semi-continuous convex function; then $\Phi(X)$ is lower semi-integrable and

$$\mathrm{E}[\Phi(X)] \ge \Phi(\mathrm{E}[X]) \tag{1.1}$$

(possibly one or both sides in the previous inequality can be equal to $+\infty$). Moreover, if $\Phi$ is strictly convex and $\Phi(\mathrm{E}[X]) < +\infty$, then the inequality is strict unless $X$ takes only one value a.s.

*Hölder's inequality*. Let $Z$, $W$ be real positive r.v.'s and $\alpha$, $\beta$ positive real numbers such that $\alpha + \beta = 1$. Then

$$\mathrm{E}[Z^\alpha W^\beta] \le \mathrm{E}[Z]^\alpha \mathrm{E}[W]^\beta . \tag{1.2}$$

From (1.2) it follows that if $X$, $Y$ are real r.v.'s and $p$, $q$ positive numbers such that $\frac{1}{p} + \frac{1}{q} = 1$, then, by setting $\alpha = \frac{1}{p}$, $\beta = \frac{1}{q}$, $Z = |X|^p$, $W = |Y|^q$, we have

$$\big|\mathrm{E}[XY]\big| \le \mathrm{E}[|X|^p]^{1/p}\mathrm{E}[|Y|^q]^{1/q} , \tag{1.3}$$

which also goes under the name of Hölder's inequality.

*Minkowski's inequality*. Let $X$, $Y$ be real r.v.'s and $p \geq 1$, then

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p \ . \tag{1.4}$$

Minkowski's inequality, in particular, implies that $L^p$, $p \geq 1$, is a vector space. If $p > q$, Jensen's inequality applied to the continuous convex function $\Phi(x) = |x|^{p/q}$ gives

$$\|X\|_p^p = \mathrm{E}[|X|^p] = \mathrm{E}[\Phi(|X|^q)] \geq \Phi(\mathrm{E}[|X|^q]) = \mathrm{E}[|X|^q]^{p/q}$$

and therefore, taking the $p$-th root,

$$\|X\|_p \geq \|X\|_q \ . \tag{1.5}$$

In particular, if $p \geq q$, $L^p \subset L^q$.

## 1.2 Variance, covariance, law of a r.v.

Let $X$ be a real square integrable r.v. (i.e. such that $\mathrm{E}[X^2] < +\infty$). Its *variance* is the quantity

$$\mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[X^2] - \mathrm{E}[X]^2 \ .$$

If $\alpha > 0$, the quantity $\int |x|^\alpha \mu(dx)$ is the *absolute moment of order* $\alpha$ of $\mu$. If $\alpha$ is a positive integer the quantity $\int x^\alpha \mu(dx)$, if it exists and is finite, is the *moment of order* $\alpha$ of $\mu$.

*Markov's inequality*. For every $\delta > 0$, $\beta > 0$ we have

$$\mathrm{P}(|X| \geq \delta) \leq \frac{\mathrm{E}[|X|^\beta]}{\delta^\beta} \ \cdot \tag{1.6}$$

It is apparent from its definition that the variance of a r.v. is so much larger as $X$ takes values far from its mean $\mathrm{E}[X]$. This intuitive fact is made precise by the following

*Chebyshev's inequality*. Let $X \in L^2$. Then for every $\alpha > 0$

$$\mathrm{P}(|X - \mathrm{E}[X]| \geq \alpha) \leq \frac{\mathrm{Var}(X)}{\alpha^2} \ \cdot$$

This is a very important inequality. It is, of course, a particular case of Markov's inequality.

The *covariance* of two r.v.'s $X$ and $Y$ is defined as

$$\mathrm{Cov}(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] .$$

If $\mathrm{Cov}(X,Y) = 0$, $X$ and $Y$ are said to be *uncorrelated*.
Let $X$ be a r.v. with values in the measurable space $(E, \mathscr{E})$. It is easy to see that the set function $\mu_X$ defined on $\mathscr{E}$ as

$$\mu_X(A) = \mathrm{P}(X^{-1}(A))$$

is itself a probability measure (on $\mathscr{E}$). $\mu_X$ is the *law* of $X$. It is the image (or pullback) of P through $X$. The following proposition provides a formula for the computation of integrals with respect to an image law. We shall make use of it throughout.

**Proposition 1.1** Let $X : (\Omega, \mathscr{F}, \mathrm{P}) \to (E, \mathscr{E})$ be a r.v., $\mu_X$ the law of $X$. Then a measurable function $f : (E, \mathscr{E}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is $\mu_X$-integrable if and only if $f(X)$ is P-integrable, and then we have

$$\int_E f(x)\,\mu_X(dx) = \int_\Omega f(X(\omega))\,\mathrm{P}(d\omega) . \tag{1.7}$$

In particular, if $X$ is a real r.v. then $X$ has finite mathematical expectation if and only if $x \mapsto x$ is $\mu_X$-integrable, and in this case

$$\mathrm{E}[X] = \int x\,\mu_X(dx) .$$

Also we have, for every $\alpha$,

$$\mathrm{E}[|X|^\alpha] = \int |x|^\alpha \mu_X(dx) .$$

Therefore $X \in L^p$ if and only if its law has a finite absolute moment of order $p$.

By the notations $X \sim \mu$, $X \sim Y$ we shall mean "$X$ has law $\mu$" and "$X$ and $Y$ have the same law", respectively.
Given two probabilities P and Q on $(\Omega, \mathscr{F})$, we say that Q is *absolutely continuous* with respect to P if $\mathrm{P}(A) = 0$ implies $\mathrm{Q}(A) = 0$. This relation is denoted $\mathrm{P} \gg \mathrm{Q}$. The Radon–Nikodym Theorem states that if $\mathrm{P} \gg \mathrm{Q}$ then there exists a r.v. $Z \geq 0$

such that

$$Q(A) = \int_A Z\,dP$$

(the converse is obvious). P and Q are said to be *equivalent* if both P $\gg$ Q and Q $\gg$ P. In other words, two probabilities are equivalent if and only if they have the same negligible events. Recall that an event $N \in \mathscr{F}$ is negligible if $P(N) = 0$. Conversely, P and Q are said to be *orthogonal* if there exists an event $A$ such that $P(A) = 1$ and $Q(A) = 0$. If P and Q are orthogonal, of course, it is not possible for one of them to be absolutely continuous with respect to the other one.

## 1.3 Independence, product measure

If $\mathscr{E}$, $\mathscr{E}'$ are $\sigma$-algebras of events of $E$ let us denote by $\mathscr{E} \vee \mathscr{E}'$ the smallest $\sigma$-algebra containing both $\mathscr{E}$ and $\mathscr{E}'$. This is well defined: it is immediate that the intersection of any non-empty family of $\sigma$-algebras is again a $\sigma$-algebra. We can therefore consider the family of all $\sigma$-algebras containing both $\mathscr{E}$ and $\mathscr{E}'$. This is non-empty as certainly $\mathscr{P}(E)$ (all subsets of $E$) is one of them. Then $\mathscr{E} \vee \mathscr{E}'$ will be the intersection of this family of $\sigma$-algebras. This argument will be used from now on to define $\sigma$-algebras as "the smallest $\sigma$-algebra such that…".

Let $X : (\Omega, \mathscr{F}, P) \to (E, \mathscr{E})$ be a r.v. and denote by $\sigma(X)$ the *$\sigma$-algebra generated by* $X$, i.e. the smallest sub-$\sigma$-algebra of $\mathscr{F}$ with respect to which $X$ is measurable. It is easy to see that $\sigma(X)$ is the family of the events of $\mathscr{F}$ of the form $X^{-1}(A')$ with $A' \in \mathscr{E}$.

The following lemma, characterizing the real $\sigma(X)$-measurable r.v.'s, is very useful.

**Lemma 1.1 (Doob's measurability criterion)** Let $X : (\Omega, \mathscr{F}, P) \to (E, \mathscr{E})$ be a r.v. Then every real $\sigma(X)$-measurable r.v. is of the form $f(X)$, where $f : (E, \mathscr{E}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is a measurable map.

The r.v.'s $X_1, \dots, X_m$, taking values respectively in $(E_1, \mathscr{E}_1), \dots, (E_m, \mathscr{E}_m)$, are said to be *independent* if, for every $A'_1 \in \mathscr{E}_1, \dots, A'_m \in \mathscr{E}_m$,

$$P(X_1 \in A'_1, \dots, X_m \in A'_m) = P(X_1 \in A'_1) \dots P(X_m \in A'_m)\ .$$

The events $A_1, \dots, A_m \in \mathscr{F}$ are said to be independent if and only if

$$P(A_{i_1} \cap \dots \cap A_{i_\ell}) = P(A_{i_1}) \dots P(A_{i_\ell})$$

for every choice of $1 \le \ell \le m$ and of $1 \le i_1 < i_2 < \cdots < i_\ell \le m$. A careful but straightforward computation shows that this definition is equivalent to the independence of the r.v.'s $1_{A_1}, \dots, 1_{A_m}$.

If $\mathscr{F}_1, \dots, \mathscr{F}_m$ are sub-$\sigma$-algebras of $\mathscr{F}$, we say that they are *independent* if, for every $A_1 \in \mathscr{F}_1, \dots, A_m \in \mathscr{F}_m$,

$$\mathrm{P}(A_1 \cap \cdots \cap A_m) = \mathrm{P}(A_1) \dots \mathrm{P}(A_m) \ .$$

It is immediate that the r.v.'s $X_1, \dots, X_m$ are independent if and only if so are the generated $\sigma$-algebras $\sigma(X_1), \dots, \sigma(X_m)$.

Let $(X_i)_{i \in I}$ be a (possibly infinite) family of r.v.'s; they are said to be independent if and only if the r.v.'s $X_{i_1}, \dots, X_{i_m}$ are independent for every choice of a finite number $i_1, \dots, i_m$ of distinct indices in $I$. A similar definition is made for an infinite family of $\sigma$-algebras.

We shall say, finally, that the r.v. $X$ is independent of the $\sigma$-algebra $\mathscr{G}$ if and only if the $\sigma$-algebras $\sigma(X)$ and $\mathscr{G}$ are independent. It is easy to see that this happens if and only if $X$ is independent of every $\mathscr{G}$-measurable r.v. $W$.

Let us now point out the relation between independence and laws of r.v.'s. If we denote by $\mu_i$ the law of $X_i$ and we define $E = E_1 \times \cdots \times E_m$, $\mathscr{E} = \mathscr{E}_1 \otimes \cdots \otimes \mathscr{E}_m$, on the product space $(E, \mathscr{E})$, we can consider the product measure $\mu = \mu_1 \otimes \cdots \otimes \mu_m$.

The following result will be of constant use in the sequel.

**Proposition 1.2** In the previous notations the r.v.'s $X_1, \dots, X_m$ are independent if and only if the law of $X = (X_1, \dots, X_m)$ on $(E, \mathscr{E})$ is the product law $\mu$.

The proof of this proposition makes use of the following theorem.

**Theorem 1.1 (Carathéodory's criterion)** Let $(E, \mathscr{E})$ be a measurable space, $\mu_1$, $\mu_2$ two finite measures on $(E, \mathscr{E})$; let $\mathscr{I} \subset \mathscr{E}$ be a family of subsets of $E$ stable with respect to finite intersections and generating $\mathscr{E}$. Then, if $\mu_1(E) = \mu_2(E)$ and $\mu_1$ and $\mu_2$ coincide on $\mathscr{I}$, they also coincide on $\mathscr{E}$.

*Proof* of Proposition 1.2. Let us denote by $\nu$ the law of $X$ and let $\mathscr{I}$ be the family of the sets of $\mathscr{E}$ of the form $A_1 \times \cdots \times A_m$ with $A_i \in \mathscr{E}_i$, $i = 1, \dots, m$. $\mathscr{I}$ is stable with respect to finite intersections and, by definition, generates $\mathscr{E}$. The definition of independence states exactly that $\mu$ and $\nu$ coincide on $\mathscr{I}$; therefore by Carathéodory's criterion they coincide on $\mathscr{E}$. $\square$

**Proposition 1.3** If $X$ and $Y$ are real independent integrable r.v.'s, then also their product is integrable and

$$\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y] .$$

Recalling that $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$, in particular we have that real integrable and independent r.v.'s are not correlated. Examples show that the converse is not true.

*Remark 1.1* A problem we shall often be confronted with is to prove that some $\sigma$-algebras are independent. Let us see how, in practice, to perform this task.

Let $\mathscr{F}_1$ and $\mathscr{F}_2$ be sub-$\sigma$-algebras of $\mathscr{F}$ and let $\mathscr{C}_1 \subset \mathscr{F}_1$, $\mathscr{C}_2 \subset \mathscr{F}_2$ be subclasses generating $\mathscr{F}_1$ and $\mathscr{F}_2$, respectively, and stable with respect to finite intersections. Then $\mathscr{F}_1$ and $\mathscr{F}_2$ are independent if and only if

$$\mathrm{P}(A_1 \cap A_2) = \mathrm{P}(A_1)\mathrm{P}(A_2) \tag{1.8}$$

for every $A_1 \in \mathscr{C}_1, A_2 \in \mathscr{C}_2$.

Actually, if $A_2 \in \mathscr{C}_2$ is fixed, the two measures on $\mathscr{F}_1$

$$A_1 \mapsto \mathrm{P}(A_1 \cap A_2) \qquad \text{and} \qquad A_1 \mapsto \mathrm{P}(A_1)\mathrm{P}(A_2)$$

are finite and coincide on $\mathscr{C}_1$; they also have the same total mass ($= \mathrm{P}(A_2)$). By Carathéodory's criterion, Theorem 1.1, they coincide on $\mathscr{F}_1$, hence (1.8) holds for every $A_1 \in \mathscr{F}_1$, $A_2 \in \mathscr{C}_2$. By a repetition of this argument with $A_1 \in \mathscr{F}_1$ fixed, (1.8) holds for every $A_1 \in \mathscr{F}_1$, $A_2 \in \mathscr{F}_2$, i.e. $\mathscr{F}_1$ and $\mathscr{F}_2$ are independent.

A case of particular interest appears when $\mathscr{F}_1 = \sigma(X, X \in \mathscr{I})$, $\mathscr{F}_2 = \sigma(Y, Y \in \mathscr{J})$ are $\sigma$-algebras generated by families $\mathscr{I}$ and $\mathscr{J}$ of r.v.'s respectively. If we assume that every $X \in \mathscr{I}$ is independent of every $Y \in \mathscr{J}$, is this enough to guarantee the independence of $\mathscr{F}_1$ and $\mathscr{F}_2$?

Thinking about this a bit it is clear that the answer is negative, as even in elementary classes in probability one deals with examples of r.v.'s that are independent pairwise but not globally. It might therefore happen that $X_1$ and $Y$ are independent, as well as $X_2$ and $Y$, whereas the pair $(X_1, X_2)$ is *not* independent of $Y$. $\{X_1, X_2\} = \mathscr{I}$ and $\{Y\} = \mathscr{J}$ therefore provides a counterexample.

If, however, we assume that for every choice of $X_1, \dots, X_n \in \mathscr{I}$ and $Y_1, \dots, Y_k \in \mathscr{J}$ the r.v.'s $(X_1, \dots, X_n)$ and $(Y_1, \dots, Y_k)$ are independent, then necessarily the two generated $\sigma$-algebras $\mathscr{F}_1$ and $\mathscr{F}_2$ are independent.

(continued)

*Remark 1.1* (continued)

Let us denote by $(E_i, \mathscr{E}_i)$ and $(G_j, \mathscr{G}_j)$ the measurable spaces where the r.v.'s $X_i, Y_j$ take the their values respectively. Let us consider the class $\mathscr{C}_1$ of the events of the form

$$\{X_1 \in A'_1, \dots, X_n \in A'_n\}$$

for some $n \geq 1, X_1, \dots, X_n \in \mathscr{I}, A'_1 \in \mathscr{E}_1, \dots, A'_n \in \mathscr{E}_n$ and similarly define $\mathscr{C}_2$ as the class of the events of the form

$$\{Y_1 \in B'_1, \dots, Y_k \in B'_k\}$$

for some $k \geq 1, Y_1, \dots, Y_k \in \mathscr{J}, B'_1 \in \mathscr{G}_1, \dots, B'_k \in \mathscr{G}_k$. One verifies that $\mathscr{C}_1$ and $\mathscr{C}_2$ are stable with respect to finite intersections and generate $\sigma(X, X \in \mathscr{I})$ and $\sigma(Y, Y \in \mathscr{J})$, respectively. Moreover, (1.8) is satisfied for every $A_1 \in \mathscr{C}_1, A_2 \in \mathscr{C}_2$. Actually, as the r.v.'s $(X_1, \dots, X_n)$ and $(Y_1, \dots, Y_k)$ are assumed to be independent, the two events

$$A_1 = \{X_1 \in A'_1, \dots, X_n \in A'_n\} = \{(X_1, \dots, X_n) \in A'_1 \times \dots \times A'_n\}$$
$$A_2 = \{Y_1 \in B'_1, \dots, Y_k \in B'_k\} = \{(Y_1, \dots, Y_k) \in B'_1 \times \dots \times B'_k\}$$

are independent. Therefore, by the previous criterion, $\mathscr{F}_1$ and $\mathscr{F}_2$ are independent. In other words, $\mathscr{F}_1$ and $\mathscr{F}_2$ are independent if and only if the r.v.'s $(X_1, \dots, X_n)$ and $(Y_1, \dots, Y_k)$ are independent for every choice of $n, k$ and of $X_1, \dots, X_n \in \mathscr{I}$ and $Y_1, \dots, Y_k \in \mathscr{J}$.

We shall see later that if the r.v.'s of $\mathscr{I}$ and $\mathscr{J}$ form together a Gaussian family, then this criterion can be considerably simplified.

The following result will be of constant use in the sequel.

**Theorem 1.2 (Fubini)** Let $\mu_1, \mu_2$ be measures respectively on the measurable spaces $(E_1, \mathscr{E}_1)$, $(E_2, \mathscr{E}_2)$ and let $f : E_1 \times E_2 \to \mathbb{R}$ be an $\mathscr{E}_1 \otimes \mathscr{E}_2$-measurable function such that at least one of the following is true:

1) $f$ is integrable with respect to $\mu_1 \otimes \mu_2$
2) $f$ is positive.

Then

a) the functions

$$x_1 \mapsto \int f(x_1, z)\, \mu_2(z), \qquad x_2 \mapsto \int f(z, x_2)\, \mu_1(dz) \tag{1.9}$$

are respectively $\mathscr{E}_1$-measurable and $\mathscr{E}_2$-measurable.

(continued)

**Theorem 1.2** (continued)
b)

$$\int_{E_1 \times E_2} f \, d\mu_1 \otimes \mu_2 = \int \mu_1(dx_1) \int f(x_1, x_2) \, \mu_2(dx_2)$$
$$= \int \mu_2(dx_2) \int f(x_1, x_2) \, \mu_1(dx_1) \, .$$

## 1.4 Probabilities on $\mathbb{R}^m$

Let $\mu$ be a probability on $(\mathbb{R}^m, \mathscr{B}(\mathbb{R}^m))$ and let us denote by $\pi_i : \mathbb{R}^m \to \mathbb{R}$ its projection on the $i$-th coordinate: we call the image of $\mu$ through $\pi_i$ the *i-th marginal law* of $\mu$. By the integration rule with respect to an image law, (1.7), the $i$-th marginal law is therefore given by

$$\mu_i(A) = \int_{\mathbb{R}^m} 1_A(x_i) \, \mu(dx_1, \dots, dx_m), \qquad A \in \mathscr{B}(\mathbb{R}) \, . \tag{1.10}$$

If $X = (X_1, \dots, X_m)$ is an $m$-dimensional r.v. with law $\mu$ it is clear that its $i$-th marginal $\mu_i$ coincides with the law of $X_i$.

We say that a probability $\mu$ on $\mathbb{R}^m$ admits a *density* (with respect to Lebesgue measure) if there exists a Borel function $f \geq 0$ such that for every $A \in \mathscr{B}(\mathbb{R}^m)$

$$\mu(A) = \int_A f(x) \, dx \, .$$

If $\mu$ admits a density $f$ then its $i$-th marginal $\mu_i$ also admits a density $f_i$, given by

$$f_i(x) = \int f(y_1, \dots, y_{i-1}, x, y_{i+1}, \dots, y_m) \, dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_m$$

(the existence of the integral and the fact that such an $f_i$ is a measurable function are consequences of Fubini's Theorem 1.2, see (1.9)).

In any case the previous formulas show that, given $\mu$, it is possible to determine its marginal distributions. The converse is not true: it is not possible, in general, knowing just the laws of the $X_i$'s, to deduce the law of $X$. Unless, of course, the r.v.'s $X_i$ are independent as, in this case, by Proposition 1.2 the law of $X$ is the product of its marginals.

Let $X$ be an $\mathbb{R}^m$-valued r.v. with density $f$, $a \in \mathbb{R}^m$ and $A$ an $m \times m$ *invertible* matrix; it is easy to see that the r.v. $Y = AX + a$ has also a density $g$ given by

$$g(y) = |\det A|^{-1} f(A^{-1}(y - a)) \, . \tag{1.11}$$

If $X = (X_1, \dots, X_m)$ is an $m$-dimensional r.v., its *covariance matrix* is the matrix $C = (c_{ij})_{i,j}$ defined as

$$c_{ij} = \mathrm{Cov}(X_i, X_j) = \mathrm{E}[(X_i - \mathrm{E}[X_i])(X_j - \mathrm{E}[X_j])] = \mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\mathrm{E}[X_j] \ .$$

It is the matrix having on the diagonal the variances of the components $X_1, \dots, X_m$ and outside the diagonal their covariances. Another more compact way of expressing the covariance matrix is the following. Let us assume $X$ centered, then

$$C = \mathrm{E}[XX^*] \ , \tag{1.12}$$

where we consider $X$ as a column vector; $X^*$ ($^*$ denotes the transpose) is a row vector and the result of the product $XX^*$ is an $m \times m$ matrix whose $i, j$-th entry is precisely $X_i X_j$. If, conversely, $X$ is not centered, we have $C = \mathrm{E}[(X - \mathrm{E}(X))(X - \mathrm{E}[X])^*]$. This formulation allows us, for instance, to derive immediately how the covariance matrix transforms with respect to linear transformations: if $A$ is a $k \times m$ matrix and $Y = AX$, then $Y$ is a $k$-dimensional r.v. and (assume $X$ centered for simplicity)

$$C_Y = \mathrm{E}[AX(AX)^*] = \mathrm{E}[AXX^*A^*] = A\mathrm{E}[XX^*]A^* = ACA^* \ . \tag{1.13}$$

If $X = (X_1, \dots, X_m)$ is an $m$-dimensional r.v., the covariance matrix provides information concerning the correlation among the r.v.'s $X_i, i = 1, \dots, m$. In particular, if these are pairwise independent the correlation matrix is diagonal. The converse is not true in general.

An important property of every covariance matrix $C$ is that it is positive definite, i.e. that for every vector $\xi \in \mathbb{R}^m$

$$\langle C\xi, \xi\rangle = \sum_{i,j=1}^m c_{ij}\xi_i\xi_j \ge 0 \ . \tag{1.14}$$

Assuming for simplicity that $X$ is centered, we have $c_{ij} = \mathrm{E}[X_i X_j]$ and

$$\langle C\xi, \xi\rangle = \sum_{i,j=1}^m c_{ij}\xi_i\xi_j = \mathrm{E}\Big[\sum_{i,j=1}^m X_i X_j \xi_i \xi_j\Big] = \mathrm{E}\Big[\Big(\sum_{i=1}^m X_i\xi_i\Big)^2\Big] \ge 0 \ , \tag{1.15}$$

as the r.v. inside the rightmost expectation is positive.

Let now $X$ and $Y$ be independent $m$-dimensional r.v.'s with laws $\mu$ and $\nu$ respectively. We call the law of $X + Y$, denoted by $\mu * \nu$, the *convolution product* of $\mu$ and $\nu$. This definition actually only depends on $\mu$ and $\nu$ and not on $X$ and $Y$: $\mu * \nu$ is actually the image on $\mathbb{R}^m$ through the map $(x, y) \mapsto x + y$ of the law of $(X, Y)$ on $\mathbb{R}^m \times \mathbb{R}^m$ that, as we have seen, is the product law $\mu \otimes \nu$ and does not depend on the particular choice of $X$ and $Y$.

**Proposition 1.4** If $\mu$ and $\nu$ are probabilities on $\mathbb{R}^m$ having density $f$ and $g$ respectively with respect to Lebesgue measure, then $\mu * \nu$ also has density $h$ with respect to Lebesgue measure given by

$$h(x) = \int_{\mathbb{R}^m} f(z)g(x-z)\,dz\,.$$

*Proof* If $A \in \mathscr{B}(\mathbb{R}^m)$, thanks to the theorem of integration with respect to an image law, Proposition 1.1,

$$\begin{aligned}\mu * \nu(A) &= \int 1_A(z)\,\mu * \nu(dz) = \int 1_A(x+y)\,\mu(dx)\,\nu(dy)\\ &= \int 1_A(x+y)f(x)g(y)\,dx\,dy = \int_A dx \int_{\mathbb{R}^m} f(z)g(x-z)\,dz\,,\end{aligned}$$

which allows us to conclude the proof. □

## 1.5 Convergence of probabilities and random variables

In this section $(E, \mathscr{B}(E))$ denotes a measurable space formed by a topological space $E$ with its Borel $\sigma$-algebra $\mathscr{B}(E)$.

- Let $(\mu_n)_n$ be a sequence of finite measures on $(E, \mathscr{B}(E))$. We say that it converges to $\mu$ *weakly* if for every continuous bounded function $f : E \to \mathbb{R}$

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu\,. \tag{1.16}$$

Note that if $\mu_n \to_{n\to\infty} \mu$ weakly, in general we do not have $\mu_n(A) \to_{n\to\infty} \mu(A)$ for $A \in \mathscr{E}$, as the indicator function $1_A$ is not, in general, continuous. It can be proved, however, that $\mu_n(A) \to_{n\to\infty} \mu(A)$ if $\mu(\partial A) = 0$ and that (1.16) also holds for functions $f$ such that the set of their points of discontinuity is negligible with respect to the limit measure $\mu$.

Let $X_n$, $n \in \mathbb{N}$, and $X$ be r.v.'s on $(\Omega, \mathscr{F}, \mathrm{P})$ taking values in $(E, \mathscr{B}(E))$.

- We say that $(X_n)_n$ converges to $X$ *almost surely (a.s.)*, denoted $X_n \overset{\text{a.s.}}{\to} X$, if there exists a negligible event $N \in \mathscr{F}$ such that

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \quad \text{for every } \omega \notin N\,.$$

- If $E = \mathbb{R}^m$ we say that $(X_n)_n$ converges to $X$ *in* $L^p$, denoted $X_n \xrightarrow{L^p} X$, if $X \in L^p$ and

$$\lim_{n\to\infty} \mathrm{E}[|X_n - X|^p] = 0 \,.$$

  Thanks to (1.5), $L^p$-convergence implies $L^q$-convergence if $p > q$.
- If the sequence $(X_n)_n$ takes its values in a metric space $E$ with a distance denoted by $d$, we say that it converges to $X$ *in probability*, denoted $X_n \xrightarrow{\mathrm{P}} X$, if, for every $\delta > 0$,

$$\lim_{n\to\infty} \mathrm{P}(d(X_n, X) > \delta) = 0 \,.$$

- We say that $(X_n)_n$ converges to $X$ *in law*, denoted $X_n \xrightarrow{\mathscr{L}} X$, if $\mu_n \to \mu$ weakly, $\mu_n$, $\mu$ denoting respectively the laws of $X_n$ and $X$. Note that for this kind of convergence it is not necessary for the r.v.'s $X, X_n, n = 1, 2, \dots,$ to be defined on the same probability space.

The following proposition summarizes the general comparison results between these notions of convergence.

**Proposition 1.5** If $X_n \xrightarrow{L^p} X$ then $X_n \xrightarrow{\mathrm{P}} X$. If $X_n \xrightarrow{a.s.} X$ then $X_n \xrightarrow{\mathrm{P}} X$. If $X_n \xrightarrow{\mathrm{P}} X$ then $X_n \xrightarrow{\mathscr{L}} X$. If $X_n \xrightarrow{\mathrm{P}} X$ then there exists a subsequence $(X_{n_k})_k$ converging to $X$ a.s.

In particular, the last statement of Proposition 1.5 implies the uniqueness of the limit in probability.

**Proposition 1.6 (Cauchy criterion of convergence in probability)** Let $(X_n)_n$ be a sequence of r.v.'s with values in the complete metric space $E$ and let us assume that for every $\lambda > 0$, $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$ such that for every $n, m > n_0$

$$\mathrm{P}(d(X_n, X_m) > \lambda) < \varepsilon \,.$$

Then there exists an $E$-valued r.v. $X$ such that $X_n \xrightarrow{\mathrm{P}} X$ as $n \to \infty$.

Let us consider on the probability space $(\Omega, \mathscr{F}, \mathrm{P})$ a sequence $(A_n)_n$ of events. Let us define an event $A$ (the *superior limit of* $(A_n)_n$) through

$$\overline{\lim_{n\to\infty}}\, A_n = \bigcap_{n=1}^{\infty} \bigcup_{k\geq n} A_k = A\ .$$

It is easy to see that, equivalently,

$$1_A = \overline{\lim_{n\to\infty}}\, 1_{A_n}$$

or

$$A = \{\omega; \omega \in A_n \text{ for infinitely many indices } n\}\ .$$

We shall often be led to the computation of the probability of the superior limit of a sequence of events. To this end, the key argument is the following.

**Proposition 1.7 (Borel–Cantelli lemma)** Let $(A_n)_n$, and $A = \overline{\lim}_{n\to\infty} A_n$ as above. Then

a) if $\sum_{n=1}^{\infty} \mathrm{P}(A_n) < +\infty$, then $\mathrm{P}(A) = 0$.
b) If $\sum_{n=1}^{\infty} \mathrm{P}(A_n) = +\infty$ and, moreover, the events $(A_n)_n$ are pairwise independent, then $\mathrm{P}(A) = 1$.

Usually Proposition 1.7 b) is stated under assumption of global independence of the events $(A_n)_n$. Actually, pairwise independence is sufficient.

## 1.6 Characteristic functions

If $X$ is an $m$-dimensional r.v. and $\mu$ denotes its law, let us define

$$\widehat{\mu}(\theta) = \int e^{i\langle\theta,x\rangle}\mu(dx) = \mathrm{E}[e^{i\langle\theta,X\rangle}]\ , \quad \theta \in \mathbb{R}^m\ .$$

$\widehat{\mu}$ is the *characteristic function* (very much similar to the Fourier transform) of $\mu$. It is defined for every probability $\mu$ on $\mathbb{R}^m$ and enjoys the following properties, some of them being immediate.

1. $\widehat{\mu}(0) = 1$ and $|\widehat{\mu}(\theta)| \leq 1$, for every $\theta \in \mathbb{R}^m$.
2. If $X$ and $Y$ are independent r.v.'s with laws $\mu$ and $\nu$ respectively, then we have $\widehat{\mu}_{X+Y}(\theta) = \widehat{\mu}(\theta)\widehat{\nu}(\theta)$.

3. $\widehat{\mu}$ is uniformly continuous.
4. If $\mu$ has finite mathematical expectation then $\widehat{\mu}$ is differentiable and

$$\frac{\partial \widehat{\mu}}{\partial \theta_j}(\theta) = i \int x_j e^{i\langle \theta, x\rangle} \mu(dx) .$$

In particular,

$$\frac{\partial \widehat{\mu}}{\partial \theta_j}(0) = i \int x_j \mu(dx) ,$$

i.e. $\widehat{\mu}'(0) = i\mathrm{E}[X]$.

5. If $\mu$ has finite moment of order 2, $\widehat{\mu}$ is twice differentiable and

$$\frac{\partial^2 \widehat{\mu}}{\partial \theta_k \partial \theta_j}(\theta) = - \int x_k x_j e^{i\langle \theta, x\rangle} \mu(dx) . \tag{1.17}$$

In particular,

$$\frac{\partial^2 \widehat{\mu}}{\partial \theta_k \partial \theta_j}(0) = - \int x_k x_j \mu(dx) . \tag{1.18}$$

For $m = 1$ this relation becomes $\widehat{\mu}''(0) = -\mathrm{E}[X^2]$. If $X$ is centered, then the second derivative at 0 is equal to $-1$ times the variance. If $m \geq 2$ and $X$ is centered, then (1.18) states that the Hessian of $\widehat{\mu}$ at the origin is equal to $-1$ times the covariance matrix.

Similar statements hold for higher-order moments: if $\alpha = (i_1, \dots, i_m)$ is a multi-index and $|\alpha| = i_1 + \dots + i_m$ and if the absolute moment of order $|\alpha|$, $\mathrm{E}[|X|^{|\alpha|}]$, is finite, then $\widehat{\mu}$ is $\alpha$ times differentiable and

$$\frac{\partial^{|\alpha|} \widehat{\mu}}{\partial x^\alpha}(\theta) = \int i^{|\alpha|} x^\alpha \, d\mu(x) . \tag{1.19}$$

These results of differentiability are an immediate consequence of the theorem of differentiation of integrals depending on a parameter.

Conversely it is not true, in general, that if all the derivatives of order $|\alpha|$ of $\widehat{\mu}$ are finite, then the moment of order $|\alpha|$ is finite. It can be shown, however, that if $|\alpha|$ is an even number and $\widehat{\mu}$ is differentiable up to the order $|\alpha|$ at 0, then $X$ has a finite moment of order $|\alpha|$ and therefore (1.19) holds. Thus, in particular, in dimension 1, if $\widehat{\mu}$ is twice differentiable at the origin, then $X$ has finite variance and (1.17) and (1.18) hold.

The previous formulas are very useful as they allow us to obtain the mean, variance, covariance, moments… of a r.v. simply by computing the derivatives of its characteristic function at 0.

6. If $\widehat{\mu}(\theta) = \widehat{\nu}(\theta)$ for every $\theta \in \mathbb{R}^m$, then $\mu = \nu$. This very important property explains the name "characteristic function".
7. If $X_1, \dots, X_m$ are r.v.'s respectively $\mu_1, \dots, \mu_n$-distributed, then they are independent if and only if, denoting by $\mu$ the law of $X = (X_1, \dots, X_m)$,

$$\widehat{\mu}(\theta_1, \dots, \theta_m) = \widehat{\mu}_1(\theta_1) \dots \widehat{\mu}_m(\theta_m)$$

(the "only if part" is a consequence of the definitions, the "if" part follows from 6 above).
8. If $\mu_k$ is the $k$-th marginal of $\mu$ then

$$\widehat{\mu}_k(\theta) = \widehat{\mu}(0, \dots, 0, \underset{\substack{\uparrow \\ k-th}}{\theta}, 0, \dots, 0)\ .$$

9. Let $b \in \mathbb{R}^k$ and $A$ a $k \times m$ matrix. Then $Y = AX + b$ is a $\mathbb{R}^k$-valued r.v.; if $\nu$ denotes its law, for $\theta \in \mathbb{R}^k$,

$$\widehat{\nu}(\theta) = \mathrm{E}[\mathrm{e}^{i\langle \theta, AX+b\rangle}] = \mathrm{e}^{i\langle \theta, b\rangle}\mathrm{E}[\mathrm{e}^{i\langle A^*\theta, X\rangle}] = \widehat{\mu}(A^*\theta)\,\mathrm{e}^{i\langle \theta, b\rangle} \tag{1.20}$$

(again $A^*$ is the transpose of $A$).
10. Clearly if $\mu_n \to_{n\to\infty} \mu$ weakly then $\widehat{\mu}_n(\theta) \to_{n\to\infty} \widehat{\mu}(\theta)$ for every $\theta$; indeed, $x \mapsto \mathrm{e}^{i\langle \theta, x\rangle}$ is a continuous function having bounded real and imaginary parts. Conversely, (P. Lévy's theorem) if

$$\lim_{n\to\infty} \widehat{\mu}_n(\theta) = \psi(\theta) \qquad \text{for every } \ \theta \in \mathbb{R}^m$$

and if $\psi$ is continuous at the origin, then $\psi$ is the characteristic function of some probability law $\mu$ and $\mu_n \to_{n\to\infty} \mu$ weakly.

## 1.7 Gaussian laws

A probability $\mu$ on $\mathbb{R}$ is said to be $N(a, \sigma^2)$ (*normal*, or *Gaussian*, with mean $a$ and variance $\sigma^2$), where $a \in \mathbb{R}, \sigma > 0$, if it has density with respect to Lebesgue measure given by

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\Big[-\frac{(x-a)^2}{2\sigma^2}\Big]\ .$$

Let us compute its characteristic function. We have at first, with the change of variable $x = y - a$,

$$\widehat{\mu}(\theta) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{+\infty} e^{i\theta y} \exp\Big[-\frac{(y-a)^2}{2\sigma^2}\Big]\,dy = \frac{e^{i\theta a}}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{+\infty} e^{i\theta x} e^{-\frac{x^2}{2\sigma^2}}\,dx$$

and if we set

$$u(\theta) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{+\infty} e^{i\theta x} e^{-\frac{x^2}{2\sigma^2}}\,dx\ ,$$

we have $\widehat{\mu}(\theta) = u(\theta)\, e^{i\theta a}$. Integrating by parts we have

$$\begin{aligned} u'(\theta) &= \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{+\infty} ix\, e^{i\theta x} e^{-\frac{x^2}{2\sigma^2}}\,dx \\ &= \frac{1}{\sqrt{2\pi}\,\sigma}(-i\sigma^2)\, e^{i\theta x} e^{-\frac{x^2}{2\sigma^2}}\Big|_{-\infty}^{+\infty} - \frac{\sigma^2\theta}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{+\infty} e^{i\theta x} e^{-\frac{x^2}{2\sigma^2}}\,dx \\ &= -\sigma^2\theta u(\theta)\ . \end{aligned}$$

This is a first order differential equation. Its general solution is $u(\theta) = c \cdot e^{-\frac{1}{2}\sigma^2\theta^2}$ and, recalling the condition $u(0) = 1$, we find

$$u(\theta) = e^{-\frac{1}{2}\sigma^2\theta^2}, \qquad \widehat{\mu}(\theta) = e^{i\theta a} e^{-\frac{1}{2}\sigma^2\theta^2}\ .$$

By points 4. and 5. of the previous section one easily derives, by taking the derivative at 0, that $\mu$ has mean $a$ and variance $\sigma^2$. If now $\nu \sim N(b, \tau^2)$ then

$$(\mu * \nu)\widehat{\ }(\theta) = \widehat{\mu}(\theta)\widehat{\nu}(\theta) = e^{i\theta(a+b)} \exp\Big[-\frac{(\sigma^2 + \tau^2)\theta^2}{2}\Big]\ .$$

$\mu * \nu$ therefore has the same characteristic function as an $N(a+b, \sigma^2+\tau^2)$ law and, by 6. of the previous section, $\mu * \nu \sim N(a + b, \sigma^2 + \tau^2)$. In particular, if $X$ and $Y$ are independent normal r.v.'s, then $X + Y$ is also normal.

Let $X_1, \dots, X_m$ be independent $N(0, 1)$-distributed r.v.'s and let $X = (X_1, \dots, X_m)$; then the vector $X$ has density

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2} \dots \frac{1}{\sqrt{2\pi}} e^{-x_m^2/2} = \frac{1}{(2\pi)^{m/2}} e^{-\frac{1}{2}|x|^2} \tag{1.21}$$

and, by 7. of the previous section, its characteristic function is given by

$$e^{-\theta_1^2/2} \dots e^{-\theta_m^2/2} = e^{-\frac{1}{2}|\theta|^2}\ .$$

If $a \in \mathbb{R}^m$ and $A$ is an $m \times m$ matrix then, recalling (1.20), $AX + a$ has characteristic function

$$e^{i\langle \theta,a\rangle} e^{-\frac{1}{2}\langle A^*\theta,A^*\theta\rangle} = e^{i\langle \theta,a\rangle} e^{-\frac{1}{2}\langle \Gamma\theta,\theta\rangle} , \tag{1.22}$$

where $\Gamma = AA^*$.

Let $a \in \mathbb{R}^m$ and let $\Gamma$ be an $m \times m$ positive semi-definite matrix. A law $\mu$ on $\mathbb{R}^m$ is said to be $N(a,\Gamma)$ (*normal with mean a and covariance matrix* $\Gamma$) if its characteristic function is given by (1.22). This is well defined as we can prove that (1.22) is certainly the characteristic function of a r.v.

It is actually well-known that if $\Gamma$ is an $m \times m$ positive semi-definite matrix, then there always exists a matrix $A$ such that $AA^* = \Gamma$; this matrix is unique under the additional requirement of being symmetric and in this case we will denote it by $\Gamma^{1/2}$. Therefore (1.22) is the characteristic function of $\Gamma^{1/2}X + a$, where $X$ has density given by (1.21). In particular, (1.21) is the density of an $N(0,I)$ law ($I$ denotes the identity matrix).

In particular, we have seen that every $X \sim N(a,\Gamma)$-distributed r.v. can be written as $X = a + \Gamma^{1/2}Z$, where $Z \sim N(0,I)$. We shall often take advantage of this property, which allows us to reduce computations concerning $N(a,\Gamma)$-distributed r.v.'s to $N(0,I)$-distributed r.v.'s, usually much simpler to deal with. This is also very useful in dimension 1: a r.v. $X \sim N(a,\sigma^2)$ can always be written as $X = a + \sigma Z$ with $Z \sim N(0,1)$.

Throughout the computation of the derivatives at 0, as indicated in 4. and 5. of the previous section, $a$ is the mean and $\Gamma$ the covariance matrix. Similarly as in the one-dimensional case we find that if $\mu$ and $\nu$ are respectively $N(a,\Gamma)$ and $N(b,\Lambda)$, then $\mu * \nu$ is $N(a+b,\Gamma+\Lambda)$ and that the sum of independent normal r.v.'s is also normal.

If the covariance matrix $\Gamma$ is invertible and $\Gamma = AA^*$, then it is not difficult to check that $A$ also must be invertible and therefore, thanks to (1.20), an $N(a,\Gamma)$-distributed r.v. has density $g$ given by (1.11), where $f$ is the $N(0,I)$ density defined in (1.21). Developing this relation and noting that $\det\Gamma = (\det A)^2$, we find, more explicitly, that

$$g(y) = \frac{1}{(2\pi)^{m/2}(\det\Gamma)^{1/2}}\, e^{-\frac{1}{2}\langle \Gamma^{-1}(y-a),y-a\rangle} .$$

If, conversely, the covariance matrix $\Gamma$ is not invertible, it can be shown that the $N(a,\Gamma)$ law does not have a density (see Exercise 1.4 for example). In particular, the $N(a,0)$ law is also defined: it is the law having characteristic function $\theta \mapsto e^{i\langle\theta,a\rangle}$, and is therefore the Dirac mass at $a$.

Let $X$ be a Gaussian $\mathbb{R}^m$-valued $N(a,\Gamma)$-distributed r.v., $A$ a $k\times m$ matrix and $b \in \mathbb{R}^k$. Let us consider the r.v. $Y = AX + b$, which is $\mathbb{R}^k$-valued. By (1.20)

$$\widehat{\nu}(\theta) = \widehat{\mu}(A^*\theta)\, e^{i\langle\theta,b\rangle} = e^{i\langle\theta,b\rangle} e^{i\langle A^*\theta,a\rangle} e^{-\frac{1}{2}\langle\Gamma A^*\theta,A^*\theta\rangle} = e^{i\langle\theta,b+Aa\rangle} e^{-\frac{1}{2}\langle A\Gamma A^*\theta,\theta\rangle} .$$

Therefore $Y$ is Gaussian with mean $b + Aa$ and variance $A\Gamma A^*$. This fact can be summarized by saying that

**Theorem 1.3** Linear-affine transformations map normal laws into normal laws.

This is a fundamental property. In particular, it implies that if $X$ and $Y$ are real r.v.'s with a normal joint law, then $X + Y$ is normal. It suffices to observe that $X + Y$ is a linear function of the vector $(X, Y)$.

The same argument gives that if $\mu = N(a, \Gamma)$ and $\mu_k$ is the $k$-th marginal distribution of $\mu$, then $\mu_k$ is also normal (the projection on the $k$-th coordinate is a linear function). This can also be deduced directly from 8. of the previous section:

$$\widehat{\mu}_k(\theta) = \mathrm{e}^{-\frac{1}{2}\Gamma_{kk}\theta^2}\mathrm{e}^{i\theta a_k} \tag{1.23}$$

and therefore $\mu_k \sim N(a_k, \Gamma_{kk})$.

If $\Gamma$ is diagonal, then using (1.22) and (1.23),

$$\widehat{\mu}(\theta_1, \dots, \theta_m) = \widehat{\mu}_1(\theta_1) \dots \widehat{\mu}_m(\theta_m) . \tag{1.24}$$

By 7. of Sect. 1.6 we see then that $\mu = \mu_1 \otimes \dots \otimes \mu_m$. Therefore

if $X = (X_1, \dots, X_m)$ is normal and has a diagonal covariance matrix, its components $X_1, \dots X_m$ are independent r.v.'s.

Therefore for two real *jointly* Gaussian r.v.'s $X$ and $Y$, if they are uncorrelated, they are also independent. This is a specific property of jointly normal r.v.'s which, as already remarked, is false in general.

A similar criterion can be stated if $X_1, \dots, X_m$ are themselves multidimensional: if the covariances between the components of $X_h$ and $X_k$, $h \neq k, 1 \leq h, k \leq m$, vanish, then $X_1, \dots, X_m$ are independent. Actually, if we denote by $\Gamma_h$ the covariance matrix of $X_h$, the covariance matrix $\Gamma$ of $X = (X_1, \dots, X_m)$ turns out to be block diagonal, with the blocks $\Gamma_h$ on the diagonal. It is not difficult therefore to repeat the previous argument and show that the relation (1.24) holds between the characteristic functions $\widehat{\mu}$ of $X$ and those, $\widehat{\mu}_h$, of the $X_h$, which implies the independence of $X_1, \dots, X_m$.

**Definition 1.1** A family $\mathscr{I}$ of $d$-dimensional r.v.'s defined on $(\Omega, \mathscr{F}, \mathrm{P})$ is said to be a *Gaussian family* if, for every choice of $X_1, \dots, X_m \in \mathscr{I}$, the $dm$-dimensional r.v. $X = (X_1, \dots, X_m)$ is Gaussian.

*Remark 1.2* Let us go back to Remark 1.1: let $\mathscr{F}_1 = \sigma(X, X \in \mathscr{I})$ and $\mathscr{F}_2 = \sigma(Y, Y \in \mathscr{J})$ be $\sigma$-algebras generated respectively by the families $\mathscr{I}$ and $\mathscr{J}$ of real r.v.'s. Let us assume, moreover, that $\mathscr{I} \cup \mathscr{J}$ is a Gaussian family. Then it is immediate that $\mathscr{F}_1$ and $\mathscr{F}_2$ are independent if and only if

$$\mathrm{Cov}(X, Y) = 0 \qquad \text{for every } X \in \mathscr{I}, Y \in \mathscr{J} \,. \tag{1.25}$$

In fact this condition guarantees that the r.v.'s $(X_1, \dots, X_n)$ and $(Y_1, \dots, Y_k)$ are independent for every choice of $X_1, \dots, X_n \in \mathscr{I}$ and $Y_1, \dots, Y_k \in \mathscr{J}$ and therefore the criterion of Remark 1.1 is satisfied. Let us recall again that (1.25) implies the independence of the generated $\sigma$-algebras only under the assumption that the r.v.'s $X \in \mathscr{I}, Y \in \mathscr{J}$ are *jointly* Gaussian.

**Proposition 1.8** Let $\mathscr{I}$ be a family of $d$-dimensional r.v.'s. Then it is a Gaussian family if and only if for every $X_1, \dots, X_m \in \mathscr{I}$ and $\gamma_1, \dots, \gamma_m \in \mathbb{R}^d$, the r.v. $\langle \gamma_1, X_1 \rangle + \dots + \langle \gamma_m, X_m \rangle$ is Gaussian.

*Proof* We shall assume $d = 1$ in order to make things simple.

Let us assume that every finite linear combination of the r.v.'s of $\mathscr{I}$ is Gaussian. Hence, in particular, every r.v. of $\mathscr{I}$ is Gaussian. This implies that if $X, Y \in \mathscr{I}$ then

$$\mathrm{E}[|XY|] \le \mathrm{E}[X^2]^{1/2} \mathrm{E}[Y^2]^{1/2} < +\infty \,.$$

So that $\mathrm{Cov}(X, Y)$ is well defined. Let $X_1, \dots, X_m \in \mathscr{I}$ and $X = (X_1, \dots, X_m)$ and let us denote by $\mu$ the law of $X$. Let us compute its characteristic function $\widehat{\mu}$. Let $\gamma = (\gamma_1, \dots, \gamma_m) \in \mathbb{R}^m$. By hypothesis $\gamma_1 X_1 + \dots + \gamma_m X_m = \langle \gamma, X \rangle$ is $N(a, \sigma^2)$-distributed for some $\sigma^2 \ge 0$, $a \in \mathbb{R}$. Now we know that $X$ has finite expectation, $z$, and covariance matrix $\Gamma$, hence

$$\begin{aligned} a &= \mathrm{E}[\langle \gamma, X \rangle] = \sum_{i=1}^m \gamma_i \mathrm{E}[X_i] = \langle \gamma, z \rangle \\ \sigma^2 &= \mathrm{E}[\langle \gamma, X \rangle^2] - a^2 = \sum_{i,j=1}^m \gamma_i \gamma_j \big( \mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\mathrm{E}[X_j] \big) = \sum_{i,j=1}^m \Gamma_{ij} \gamma_i \gamma_j \,. \end{aligned}$$

Therefore we have

$$\widehat{\mu}(\gamma) = \mathrm{E}[\mathrm{e}^{i\langle \gamma, X \rangle}] = \mathrm{e}^{ia} \mathrm{e}^{-\sigma^2/2} = \mathrm{e}^{i\langle \gamma, z \rangle} \mathrm{e}^{-\frac{1}{2}\langle \Gamma \gamma, \gamma \rangle}$$

so $X$ is Gaussian.

Conversely, if $X = (X_1, \dots, X_m)$ is Gaussian and $\gamma \in \mathbb{R}^m$, then $\gamma_1 X_1 + \dots + \gamma_m X_m = \langle \gamma, X \rangle$ is Gaussian, being a linear function of a Gaussian vector (see Theorem 1.3). □

The following property of the Gaussian laws, which we shall use quite often, is, together with the invariance with respect to affine transformations of Theorem 1.3, the most important.

**Proposition 1.9** Let $(X_n)_n$ be a sequence of $m$-dimensional Gaussian r.v.'s converging in law to a r.v. $X$. Then $X$ is also Gaussian.

*Proof* Thanks to Proposition 1.8 it is sufficient to prove that $Y = \langle \gamma, X \rangle$ is Gaussian for every $\gamma \in \mathbb{R}^m$.

The r.v.'s $Y_n = \langle \gamma, X_n \rangle$ are Gaussian, being linear functions of Gaussian vectors. Denoting by $m_n, \sigma_n^2$ respectively the mean and variance of $Y_n$, the characteristic function of $Y_n$ is

$$\phi_n(\theta) = \mathrm{e}^{i\theta m_n} \mathrm{e}^{-\frac{1}{2}\sigma_n^2 \theta^2} .$$

By hypothesis $\phi_n(\theta) \to \phi(\theta)$ as $n \to \infty$, where $\phi$ is the characteristic function of $Y$. Taking the modulus we have, for every $\theta \in \mathbb{R}$,

$$\mathrm{e}^{-\frac{1}{2}\sigma_n^2 \theta^2} \underset{n\to\infty}{\to} |\phi(\theta)| .$$

This proves that the sequence $(\sigma_n^2)_n$ is bounded. Actually, if there existed a subsequence converging to $+\infty$ we would have $|\phi(\theta)| = 0$ for $\theta \neq 0$ and $|\phi(0)| = 1$, which is not possible, $\phi$ being continuous. If $\sigma^2$ denotes the limit of a subsequence of $(\sigma_n^2)_n$, then necessarily $\mathrm{e}^{-\frac{1}{2}\sigma_n^2 \theta^2} \to \mathrm{e}^{-\frac{1}{2}\sigma^2 \theta^2}$.

Let us prove now that the sequence of the means, $(m_n)_n$, is also bounded. Note that, $Y_n$ being Gaussian, we have $\mathrm{P}(Y_n \ge m_n) = \frac{1}{2}$ if $\sigma_n^2 > 0$ and $\mathrm{P}(Y_n \ge m_n) = 1$ if $\sigma_n^2 = 0$, as in this case the law of $Y_n$ is the Dirac mass at $m_n$. In any case, $\mathrm{P}(Y_n \ge m_n) \ge \frac{1}{2}$.

Let us assume that $(m_n)_n$ is unbounded. Then there would exist a subsequence $(m_{n_k})_k$ converging to $+\infty$ (this argument is easily adapted to the case $m_{n_k} \to -\infty$). For every $M \in \mathbb{R}$ we would have, as $m_{n_k} \ge M$ for $k$ large,

$$\mathrm{P}(Y \ge M) \ge \varlimsup_{n\to\infty} \mathrm{P}(Y_{n_k} \ge M) \ge \varlimsup_{n\to\infty} \mathrm{P}(Y_{n_k} \ge m_{n_k}) = \frac{1}{2} ,$$

which is not possible, as necessarily $\lim_{M\to+\infty} \mathrm{P}(Y \ge M) = 0$. As $(m_n)_n$ is bounded, there exists a convergent subsequence $(m_{n_k})_k$, $m_{n_k} \to m$ say. Therefore we have, for every $\theta \in \mathbb{R}$,

$$\phi(\theta) = \lim_{k\to\infty} \mathrm{e}^{i\theta\mu_{n_k}} \mathrm{e}^{-\frac{1}{2}\sigma_{n_k}^2\theta^2} = \mathrm{e}^{i\theta m}\mathrm{e}^{-\frac{1}{2}\sigma^2\theta^2}$$

hence $Y$ is Gaussian. □

## 1.8 Simulation

Often one needs to compute a probability or an expectation for which there is no explicit formula. In this case simulation is natural option.

For instance, if $X$ is a r.v. and $f$ a real Borel function, bounded to make things simple, and we must know the value of the expectation

$$\mathrm{E}[f(X)] ,$$

then the Law of Large Numbers states that if $(X_n)_n$ is a sequence of independent identically distributed r.v.'s with the same law as $X$, then

$$\frac{1}{N}\sum_{i=1}^{N} f(X_i) \underset{N\to\infty}{\to} \mathrm{E}[f(X)] .$$

A first question in this direction is to be able to simulate a r.v. with a given distribution, i.e. to instruct a computer to generate random numbers with a given distribution.

In this section we shall see how to simulate Gaussian distributed random numbers.

We will assume that the researcher has access to a programming language with a random generator uniformly distributed on $[0, 1]$, i.e. a command that produces a random number with this law and such that, moreover, repeated calls to it give rise to random numbers that are independent.

These are the so-called *pseudo-random numbers*. We shall not approach the question of how they are obtained: we shall just be satisfied with the fact that all the common programming languages (FORTRAN, C, …) comply with our needs. We shall see later, however, that it is always wise to doubt of the quality of these generators…

The first problem is: how is it possible, starting from a random generator uniform on $[0, 1]$, to obtain a generator producing a r.v. with a Gaussian distribution? Various algorithms are available (see also Exercise 1.21); curiously most of them produce simultaneously *two* independent random numbers with an $N(0, 1)$ distribution. The one provided by the following proposition is possibly the simplest.

**Proposition 1.10 (The Box–Müller algorithm)** Let $W, Z$ be independent r.v.'s respectively exponential with parameter $\frac{1}{2}$ (see Exercise 1.2) and uniform on $[0, 2\pi]$. Then $X = \sqrt{W} \cos Z$ and $Y = \sqrt{W} \sin Z$ are independent and $N(0, 1)$-distributed.

*Proof* First, let us compute the density of $R = \sqrt{W}$. The partition function of $W$ is $F(t) = 1 - e^{-t/2}$ (see also Exercise 1.2). Therefore the partition function of $R$ is, for $r > 0$,

$$F_R(r) = P(\sqrt{W} \le r) = P(W \le r^2) = 1 - e^{-r^2/2}$$

and $F_R(r) = 0$ for $r \le 0$. Therefore, taking the derivative, its density is $f_R(r) = re^{-r^2/2}$, $r > 0$. As the density of $Z$ is equal to $\frac{1}{2\pi}$ on the interval $[0, 2\pi]$, the joint density of $\sqrt{W}$ and $Z$ is

$$f(r, z) = \frac{1}{2\pi} re^{-r^2/2}, \quad \text{for } r > 0,\ 0 \le z \le 2\pi,$$

and $f(r, z) = 0$ otherwise. Let us compute the joint density, $g$ say, of $X = \sqrt{W} \cos Z$ and $Y = \sqrt{W} \sin Z$: $g$ is characterized by the relation

$$E[\Phi(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi(x, y) g(x, y)\, dx\, dy \tag{1.26}$$

for every bounded Borel function $\Phi : \mathbb{R}^2 \to \mathbb{R}$. Recalling the formulas of integration in polar coordinates,

$$\begin{aligned} E[\Phi(X, Y)] = E[\Phi(\sqrt{W} \cos Z, \sqrt{W} \sin Z)] &= \int\int \Phi(r \cos z, r \sin z) f(r, z)\, dr\, dz \\ &= \frac{1}{2\pi} \int_0^{2\pi} dz \int_0^{+\infty} \Phi(r \cos z, r \sin z)\, re^{-r^2/2}\, dr \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi(x, y) e^{-\frac{1}{2}(x^2+y^2)}\, dx\, dy \end{aligned}$$

and, comparing with (1.26), we derive that the joint density $g$ of $X, Y$ is

$$g(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)},$$

which proves simultaneously that $X$ and $Y$ are independent and $N(0, 1)$-distributed. □

Proposition 1.10 suggests the following recipe for the simulation of an $N(0, I)$-distributed r.v.: let $Z$ be a r.v. uniform on $[0, 2\pi]$ and $W$ an exponential r.v. with parameter $\frac{1}{2}$. They can be obtained from a uniform distribution as explained in Exercise 1.2). Then the r.v.'s

$$X = \sqrt{W} \cos Z, \quad Y = \sqrt{W} \sin Z \tag{1.27}$$

are Gaussian $N(0, I)$-distributed and independent. To be more explicit the steps are:

- Simulate two independent r.v.'s $U_1$, $U_2$ uniform on $[0, 1]$ (these are provided by the random number generator of the programming language that you are using);
- set $Z = 2\pi U_1$, $W = 2 \log(1 - U_2)$; then $Z$ is uniform on $[0, 2\pi]$ and $W$ is exponential with parameter $\frac{1}{2}$ (see Exercise 1.2);
- then the r.v.'s $X$, $Y$ as in (1.27) are $N(0, 1)$-distributed and independent.

This algorithm produces an $N(0, 1)$-distributed r.v., but of course from this we can easily obtain an $N(m, \sigma^2)$-distributed one using the fact that if $X$ is $N(0, 1)$-distributed, then $m + \sigma X$ is $N(m, \sigma^2)$-distributed.

*Remark 1.3* a) The Box–Müller algorithm introduced in Proposition 1.10 is necessary when using low level programming languages such as FORTRAN or C. High level languages such as Matlab, Mathematica or scilab, in fact already provide routines that directly produce Gaussian-distributed r.v.'s and, in fact, the simulated paths of the figures of the forthcoming chapters have been produced with one of them.

However, these high-level languages are interpreted, which means that they are relatively slow when dealing with a real simulation, requiring the production of possibly millions of random numbers. In these situations a compiled language such as FORTRAN or C is necessary.

b) The huge number of simulations that are required by a real life application introduces another caveat. The random number generator that is usually available in languages such as FORTRAN or C is based on arithmetic procedures *that have a cycle*, i.e. the generator produces numbers $a_0, a_1, \dots$ that can be considered uniformly distributed on $[0, 1]$ and independent, but after $N$ numbers ($N$ =the cycle length) it goes back to the beginning of the cycle, repeating again the same numbers $a_0, a_1, \dots$

This means that it is not fit for your simulation if your application requires the production of more than $N$ random numbers. In particular, the C command `rnd()` has a cycle $N \sim 2 \cdot 10^9$, which can very well be insufficient (we shall see that the simulation of a single path may require many random numbers).

Luckily more efficient random number simulators, with a reassuringly large cycle length, have been developed and are available in all these languages and should be used instead of the language default.

## 1.9 Measure-theoretic arguments

Very often in the sequel we shall be confronted with the problem of proving that a certain statement is true for a large class of functions. Measure theory provides several tools in order to deal with this kind of question, all based on the same idea: just prove the statement for a smaller class of functions (for which the check is easy) and then show that necessarily it must be true for the larger class. In this section we give, without proof, three results that will be useful in order to produce this kind of argument.

In general, if $E$ is a set and $\mathscr{C}$ a class of parts of $E$, by $\sigma(\mathscr{C})$ we denote the $\sigma$-algebra generated by $\mathscr{C}$, i.e. the smallest $\sigma$-algebra of parts of $E$ containing $\mathscr{C}$.

**Proposition 1.11** Let $(E, \mathscr{E})$ be a measurable space and $f$ a positive real measurable function on $E$. Then there exists an increasing sequence $(f_n)_n$ of functions of the form

$$f_n(x) = \sum_{i=1}^{m} \alpha_i 1_{A_i}(x) \tag{1.28}$$

such that $f_n \nearrow f$.

Functions of the form appearing on the right-hand side of (1.28) are called *elementary*. Therefore Proposition 1.11 states that every measurable positive function is the increasing limit of a sequence of elementary functions.

**Theorem 1.4** Let $E$ be a set, $\mathscr{C}$ a class of parts of $E$ stable with respect to finite intersections and $\mathscr{H}$ a vector space of bounded functions such that:

i) if $(f_n)_n$ is an increasing sequence of elements of $\mathscr{H}$ all bounded above by the same element of $\mathscr{H}$ then $\sup_n f_n \in \mathscr{H}$; and
ii) $\mathscr{H}$ contains the function 1 and the indicator functions of the elements of $\mathscr{C}$.

Then $\mathscr{H}$ contains every bounded $\sigma(\mathscr{C})$-measurable function.

**Theorem 1.5** Let $\mathscr{H}$ be a vector space of bounded real functions on a set $E$ such that:

(continued)

**Theorem 1.5** (continued)
i) $1 \in \mathscr{H}$;
ii) if $(f_n)_n \subset \mathscr{H}$ and $f_n \to_{n\to\infty} f$ uniformly, then $f \in \mathscr{H}$; and
iii) if $(f_n)_n \subset \mathscr{H}$ is an increasing sequence of equi-bounded positive functions and $\lim_{n\to\infty} f_n = f$, then $f \in \mathscr{H}$.

Let $L$ be a subspace of $\mathscr{H}$ stable under multiplication. Then $\mathscr{H}$ contains every bounded $\sigma(L)$-measurable function.

Let us prove Theorem 1.1 (Carathéodory's criterion) as an example of application of Theorem 1.4. Let $\mathscr{H}$ be the set of the real bounded measurable functions $f$ on $E$ such that

$$\int f \, d\mu_1 = \int f \, d\mu_2 \; .$$

$\mathscr{H}$ is a vector space and satisfies i) of Theorem 1.4 thanks to Lebesgue's theorem. Moreover, of course, $\mathscr{H}$ contains the function 1 and the indicators of the elements of $\mathscr{I}$. Therefore $\mathscr{H}$ contains every bounded $\sigma(\mathscr{I})$-measurable function. As by assumption $\sigma(\mathscr{I}) = \mathscr{E}$, $\mu_1$ and $\mu_2$ coincide on $\mathscr{E}$.

This example of application of Theorem 1.4 is typical when one must prove in general a property that is initially only known for a particular class of events.

Similarly, Theorem 1.5 will be used, typically, in order to extend to every bounded measurable function a property which is known at first only for a particular class of functions: for instance, the class of continuous functions on $(E, \mathscr{B}(E))$, if $E$ is a metric space.

## Exercises

**1.1** (p. 437) Given a real r.v. $X$, its *partition function* (p.f.) is the function

$$F(t) = \mathrm{P}(X \le t) \; .$$

Show that two real r.v.'s $X$ and $Y$ have the same p.f. if and only if they have the same distribution.

Use Carathéodory's criterion, Theorem 1.1.

**1.2** (p. 437) a) A r.v. $X$ has *exponential law* with parameter $\lambda$ if it has density

$$f(x) = \lambda e^{-\lambda x} 1_{[0,+\infty[}(x) \; .$$

What is the p.f. $F$ of $X$ (see the definition in Exercise 1.1)? Compute the mean and variance of $X$.

b) Let $U$ be a r.v. uniformly distributed on $[0, 1]$, i.e. having density

$$f(x) = 1_{[0,1]}(x).$$

Compute the mean and variance of $U$.

c) Let $U$ be as in b).

c1) Compute the law of $Z = \alpha U$, $\alpha > 0$.

c2) Compute the law of $W = -\frac{1}{\lambda}\log U$, $\lambda > 0$.

**1.3** (p. 439) a) Let $X$ be a positive r.v. and $f : \mathbb{R}^+ \to \mathbb{R}$ a differentiable function with continuous derivative and such that $f(X)$ is integrable. Then

$$\mathrm{E}[f(X)] = f(0) + \int_0^{+\infty} f'(t)\mathrm{P}(X \geq t)\, dt\ .$$

b) Let $X$ be a positive integer-valued r.v. Then

$$\mathrm{E}[X] = \sum_{k=1}^{\infty} \mathrm{P}(X \geq k)\ .$$

a) If $\mu$ is the law of $X$, then $\int_0^{+\infty} f'(t)\, dt \int_t^{+\infty} \mu(dx) = \int_0^{+\infty} \mu(dx) \int_0^x f'(t)\, dt$ by Fubini's theorem.

**1.4** (p. 439) Let $X$ be an $m$-dimensional r.v. and let us denote by $C$ its covariance matrix.

a) Prove that if $X$ is centered then $\mathrm{P}(X \in \mathrm{Im}\, C) = 1$ ($\mathrm{Im}\, C$ is the image of the matrix $C$).

b) Deduce that if the covariance matrix $C$ of a r.v. $X$ is not invertible, then the law of $X$ cannot have a density.

**1.5** (p. 439) Let $X, X_n, n = 1, 2, \dots,$ be $\mathbb{R}^m$-valued r.v.'s. Prove that if from every subsequence of $(X_n)_n$ we can extract a further subsequence convergent to $X$ in $L^p$ (resp. in probability) then $(X_n)_n$ converges to $X$ in $L^p$ (resp. in probability). Is this also true for a.s. convergence?

**1.6** (p. 440) Given an $m$-dimensional r.v. $X$, its *Laplace transform* is the function

$$\mathbb{R}^m \ni \theta \to \mathrm{E}[\mathrm{e}^{\langle \theta, X\rangle}]$$

(possibly $= +\infty$). Prove that, if $X \sim N(b, \Gamma)$, then

$$\mathrm{E}[\mathrm{e}^{\langle \theta, X\rangle}] = \mathrm{e}^{i\langle \theta, b\rangle}\mathrm{e}^{\frac{1}{2}\langle \Gamma\theta, \theta\rangle}.$$

**1.7** (p. 440) Prove points 2 and 8 of Sect. 1.6.

**1.8** (p. 440) Let $X$ be a real r.v. having a *Laplace law* with parameter $\lambda$, i.e. having density

$$f(x) = \frac{\lambda}{2}\,\mathrm{e}^{-\lambda|x|}\,.$$

Compute its characteristic function and its Laplace transform.

**1.9** (p. 441) Let $X$, $Y$ be independent $N(0,1)$-distributed r.v.'s. Determine the laws of the two-dimensional r.v.'s $(X, X+Y)$ and $(X, \sqrt{2}X)$. Show that these two laws have the same marginals.

**1.10** (p. 442) Let $X_1, X_2$ be independent $N(0,1)$-distributed r.v.'s. If $Y_1 = X_1 - X_2$, $Y_2 = X_1 + X_2$, show that $Y_1$ and $Y_2$ are independent. And if it were $Y_1 = \frac{1}{2}X_1 - \frac{\sqrt{3}}{2}X_2$, $Y_2 = \frac{1}{2}X_1 + \frac{\sqrt{3}}{2}X_2$?

**1.11** (p. 443) a) Let $X$ be an $N(\mu, \sigma^2)$-distributed r.v. Compute the density of $\mathrm{e}^X$ (*lognormal law* of parameters $\mu$ and $\sigma^2$).

b) Show that a lognormal law has finite moments of all orders and compute them. What are the values of its mean and variance?

**1.12** (p. 443) Let $X$ be an $N(0, \sigma^2)$-distributed r.v. Compute, for $t \in \mathbb{R}$, $\mathrm{E}[\mathrm{e}^{tX^2}]$.

**1.13** (p. 444) Let $X$ be an $N(0,1)$-distributed r.v., $\sigma, b$ real numbers and $x, K > 0$. Show that

$$\mathrm{E}[(x\mathrm{e}^{b+\sigma X} - K)^+] = x\mathrm{e}^{b+\frac{1}{2}\sigma^2}\Phi(-\zeta + \sigma) - K\Phi(-\zeta)\,,$$

where $\zeta = \frac{1}{\sigma}(\log \frac{K}{x} - b)$ and $\Phi$ denotes the partition function of an $N(0,1)$-distributed r.v. This quantity appears naturally in many questions in mathematical finance, see Sect. 13.6. ($x^+$ denotes the positive part function, $x^+ = x$ if $x \geq 0$, $x^+ = 0$ if $x < 0$.)

**1.14** (p. 444) a) Let $(X_n)_n$ be a sequence of $m$-dimensional Gaussian r.v.'s respectively with mean $b_n$ and covariance matrix $\Gamma_n$. Let us assume that

$$\lim_{n\to\infty} b_n := b, \qquad \lim_{n\to\infty} \Gamma_n := \Gamma.$$

Show that $X_n \xrightarrow{\mathcal{L}} N(b, \Gamma)$ as $n \to \infty$.

b1) Let $(Z_n)_n$ be a sequence of $N(0, \sigma^2)$-distributed real independent r.v.'s. Let $(X_n)_n$ be the sequence defined recursively by

$$X_0 = x \in \mathbb{R}, \qquad X_{n+1} = \alpha X_n + Z_{n+1}$$

where $|\alpha| < 1$ (i.e., possibly, $X_1, \dots, X_n, \dots$ represent the subsequent positions of a moving object that at every iteration moves from the actual position $X_n$ to the position $\alpha X_n$ but also undergoes a perturbation $Z_{n+1}$). What is the law of $X_1$? And of $X_2$? Prove that, as $n \to \infty$, $(X_n)_n$ converges in law and determine the limit law.

b2) Prove that, as $n \to \infty$, the sequence of two-dimensional r.v.'s $((X_n, X_{n+1}))_n$ converges in law and determine the limit law.

**1.15** (p. 446) a) Prove that, for every $p \geq 0$, there exists a constant $c_{p,m} > 0$ such that $\mathrm{E}(|X|^p) \leq c_{p,m} \mathrm{E}(|X|^2)^{p/2}$ for every Gaussian $m$-dimensional centered r.v. $X$.

b) Let $(X_n)_n$ be a sequence of Gaussian r.v.'s converging to 0 in $L^2$. Show that the convergence also takes place in $L^p$ for every $p > 0$.

Recall the inequality, for $x = (x_1, \dots, x_m) \in \mathbb{R}^m$,

$$\sum_{i=1}^m |x_i|^p \leq |x|^p \leq m^{\frac{p-2}{2}} \sum_{i=1}^m |x_i|^p .$$

**1.16** (p. 446) (Example of a pair of Gaussian r.v.'s whose joint law it is not Gaussian) Let $X, Z$ be independent r.v.'s with $X \sim N(0, 1)$ and $\mathrm{P}(Z = 1) = \mathrm{P}(Z = -1) = \frac{1}{2}$. Let $Y = XZ$.

a) Prove that $Y$ is itself $N(0, 1)$.

b) Prove that $X + Y$ is not Gaussian. Does $(X, Y)$ have a joint Gaussian law?

**1.17** (p. 447) a) Let $(X_n)_n$ be a sequence of independent $N(0, 1)$-distributed r.v.'s. Prove that, for every $\alpha > 2$,

$$\mathrm{P}\big(X_n > (\alpha \log n)^{1/2} \text{ for infinitely many indices } n\big) = 0 . \tag{1.29}$$

b) Prove that

$$\mathrm{P}\big(X_n > (2 \log n)^{1/2} \text{ for infinitely many indices } n\big) = 1 . \tag{1.30}$$

c) Show that the sequence $((\log n)^{-1/2} X_n)_n$ tends to 0 in probability but not a.s.

Use the following inequalities that will be proved later (Lemma 3.2)

$$\Big(x + \frac{1}{x}\Big)^{-1} \mathrm{e}^{-x^2/2} \leq \int_x^{+\infty} \mathrm{e}^{-z^2/2}\, dz \leq \frac{1}{x}\, \mathrm{e}^{-x^2/2} .$$

**1.18** (p. 448) Prove Proposition 1.1 (the integration rule with respect to an image probability).

**1.19** (p. 449) (A very useful measurability criterion) Let $X$ be a map $(\Omega, \mathscr{F}) \to (E, \mathscr{E})$, $\mathscr{D} \subset \mathscr{E}$ a family of subsets of $E$ such that $\sigma(\mathscr{D}) = \mathscr{E}$ and let us assume that $X^{-1}(A) \in \mathscr{F}$ for every $A \in \mathscr{D}$. Show that $X$ is measurable.

**1.20** (p. 449) (A special trick for the $L^1$ convergence of densities) Let $Z_n, Z$ be positive r.v.'s such that $Z_n \to Z$ a.s. as $n \to \infty$ and $\mathrm{E}[Z_n] = \mathrm{E}[Z] < +\infty$ for every $n$. We want to prove that the convergence also takes place in $L^1$.

a) Let $H_n = \min(Z_n, Z)$. Prove that $\lim_{n\to\infty} \mathrm{E}[H_n] = \mathrm{E}[Z]$.

b) Note that $|Z_n - Z| = (Z - H_n) + (Z_n - H_n)$ and deduce that $Z_n \to Z$ also in $L^1$.

**1.21** (p. 449) In the FORTRAN libraries in use in the 70s (but also nowadays ...), in order to generate an $N(0, 1)$-distributed random number the following procedure was implemented. If $X_1, \dots, X_{12}$ are independent r.v.'s uniformly distributed on $[0, 1]$, then the number

$$W = X_1 + \cdots + X_{12} - 6 \tag{1.31}$$

is (approximately) $N(0, 1)$-distributed.

a) Can you give a justification of this procedure?

b) Let $Z$ be a $N(0, 1)$-distributed r.v. What is the value of $\mathrm{E}[Z^4]$? What is the value of $\mathrm{E}[W^4]$ for the r.v. given by (1.31)? What do you think of this procedure?

# Chapter 2
# Stochastic Processes

## 2.1 General facts

A stochastic process is a mathematical object that is intended to model the evolution in time of a random phenomenon. As will become clear in the sequel the appropriate setting is the following.

> A *stochastic process* is an object of the form
>
> $$X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in T}, (X_t)_{t\in T}, \mathrm{P}) ,$$
>
> where
>
> - $(\Omega, \mathscr{F}, \mathrm{P})$ is a probability space;
> - $T$ (the times) is a subset of $\mathbb{R}^+$;
> - $(\mathscr{F}_t)_{t\in T}$ is a *filtration*, i.e. an increasing family of sub-$\sigma$-algebras of $\mathscr{F}$: $\mathscr{F}_s \subset \mathscr{F}_t$ whenever $s \le t$; and
> - $(X_t)_{t\in T}$ is a family of r.v.'s on $(\Omega, \mathscr{F})$ taking values in a measurable space $(E, \mathscr{E})$ such that, for every $t$, $X_t$ is $\mathscr{F}_t$-measurable. This fact is also expressed by saying that $(X_t)_t$ is *adapted* to the filtration $(\mathscr{F}_t)_t$.

Let us look at some aspects of this definition.

Of course $X_t$ will represent the random position of our object at time $t$.

At this point the meaning of the $\sigma$-algebras $\mathscr{F}_t$ and the role that they are going to play may not be evident; we will understand them well only later. For now let us just say that $\mathscr{F}_t$ is "the amount of information that is available at time $t$", i.e. it is the family of the events for which, at time $t$, one knows whether they happened or not.

Note that, if $s \le t$, as $X_s$ must be $\mathscr{F}_s$-measurable and $\mathscr{F}_s \subset \mathscr{F}_t$, then all the r.v.'s $X_s$ for $s \le t$, must be $\mathscr{F}_t$ measurable, so that $\mathscr{F}_t$ necessarily contains the $\sigma$-algebra $\mathscr{G}_t = \sigma(X_s, s \le t)$. These $\sigma$-algebras form the *natural filtration* $(\mathscr{G}_t)_t$, which is therefore the smallest possible filtration with respect to which $X$ is adapted. As $X_s$

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_2

for $s \le t$ is an $\mathscr{F}_t$ measurable r.v., this means intuitively that at time $t$ we know the positions of the process at the times before time $t$.

In general, if $(\mathscr{F}_t)_t$ is a filtration, the family $\bigcup_t \mathscr{F}_t$ is not necessarily a $\sigma$-algebra. By $\mathscr{F}_\infty$ we shall denote the smallest $\sigma$-algebra of parts of $\Omega$ containing $\bigcup_t \mathscr{F}_t$. This is a $\sigma$-algebra that can be strictly smaller than $\mathscr{F}$; it is also denoted $\bigvee_t \mathscr{F}_t$.

Another filtration that we shall often be led to consider is the *augmented natural filtration*, $(\overline{\mathscr{G}}_t)_t$, where $\overline{\mathscr{G}}_t$ is the $\sigma$-algebra obtained by adding to $\mathscr{G}_t = \sigma(X_u, u \le t)$ the negligible events of $\mathscr{F}$, i.e.

$$\mathscr{G}_t = \sigma(X_u, u \le t), \qquad \overline{\mathscr{G}}_t = \sigma(\mathscr{G}_t, \mathscr{N}) ,$$

where $\mathscr{N} = \{A; A \in \mathscr{F}, \mathrm{P}(A) = 0\}$.

The space $E$ appearing above (in which the process takes its values) is called the *state space*. We shall mainly be interested in the case $(E, \mathscr{E}) = (\mathbb{R}^m, \mathscr{B}(\mathbb{R}^m))$; in general, however, it might also be a more involved object (a group of matrices, a manifold,…).

Via the map $\omega \mapsto (t \mapsto X_t(\omega))$ one can always think of $\Omega$ as a subset of $E^T$ ($E^T$ is the set of all functions $T \to E$); for this reason $\Omega$ is also called the *space of paths*. This point of view will be developed in Sect. 3.2.
Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in T}, (X_t)_{t\in T}, \mathrm{P})$ and $(\Omega', \mathscr{F}', (\mathscr{F}'_t)_{t\in T}, (X'_t)_{t\in T}, \mathrm{P}')$ be two processes.

They are said to be *equivalent* if for every choice of $t_1, \dots, t_m \in T$ the r.v.'s $(X_{t_1}, \dots, X_{t_m})$ and $(X'_{t_1}, \dots, X'_{t_m})$ have the same law.

As we shall see in Sect. 2.3, the notion of equivalence of processes is very important: in a certain sense two equivalent processes "are the same process", at least in the sense that they model the same situation.

One process is said to be a *modification* (or a *version*) of the other if $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in T}, \mathrm{P}) = (\Omega', \mathscr{F}', (\mathscr{F}'_t)_{t\in T}, \mathrm{P}')$ and if, for every $t \in T$, $X_t = X'_t$ P-a.s. In this case, in particular, they are equivalent.

They are said to be *indistinguishable* if

$$(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in T}, \mathrm{P}) = (\Omega', \mathscr{F}', (\mathscr{F}'_t)_{t\in T}, \mathrm{P}')$$

and

$$\mathrm{P}(X_t = X'_t \text{ for every } t \in T) = 1 .$$

Two indistinguishable processes are clearly modifications of one another. The converse is false, as shown in the following example.

*Example 2.1* If $\Omega = [0, 1]$, $\mathscr{F} = \mathscr{B}([0, 1])$ and P =Lebesgue measure, let

$$X_t(\omega) = 1_{\{\omega\}}(t), \qquad X'_t(\omega) \equiv 0 .$$

$X'$ is a modification of $X$, because the event $\{X_t \neq X'_t\}$ is formed by the only element $t$, and therefore has probability 0. But the two processes are not indistinguishable, and the event $\{\omega; X_t(\omega) = X'_t(\omega) \text{ for every } t\}$ is even empty.

Let us assume from now on that the state space $E$ is a topological space endowed with its Borel $\sigma$-algebra $\mathscr{B}(E)$ and that $T$ is an interval of $\mathbb{R}^+$.

A process is said to be *continuous* (resp. *a.s. continuous*) if for every $\omega$ (resp. for almost every $\omega$) the map $t \mapsto X_t(\omega)$ is continuous. The definitions of a *right-continuous* process, an *a.s. right-continuous* process, etc., are quite similar.

Note that the processes $X$ and $X'$ of Example 2.1 are modifications of each other but, whereas $X'$ is continuous, $X$ is not. Therefore, in general, the property of being continuous is not preserved when passing from a process to a modification.

$X$ is said to be *measurable* if the map $(t, \omega) \mapsto X_t(\omega)$ is measurable $(T \times \Omega, \mathscr{B}(T) \otimes \mathscr{F}) \to (E, \mathscr{B}(E))$. It is said to be *progressively measurable* if for every $u \in T$ the map $(t, \omega) \mapsto X_t(\omega)$ is measurable $([0, u] \times \Omega, \mathscr{B}([0, u]) \otimes \mathscr{F}_u) \to (E, \mathscr{B}(E))$. A progressively measurable process is obviously measurable.

It is crucial to assume that the processes we are working with are progressively measurable (see Examples 2.2 and 2.3 below). Luckily we will be dealing with continuous processes and the following proposition guarantees that they are progressively measurable.

**Proposition 2.1** Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be a right-continuous process. Then it is progressively measurable.

*Proof* For a fixed $u \in T$ let us consider the piecewise constant process

$$X_s^{(n)} = \begin{cases} X_{\frac{k+1}{2^n} u} & \text{if } s \in [\frac{k}{2^n} u, \frac{k+1}{2^n} u[ \\ X_u & \text{if } s = u . \end{cases}$$

Note that we can write $X_s^{(n)} = X_{s_n}$ where $s_n > s$ is a time such that $|s_n - s| \le u2^{-n}$ ($s_n = (k + 1)u2^{-n}$ if $s \in [\frac{k}{2^n} u, \frac{k+1}{2^n} u[$). Hence $s_n \searrow s$ as $n \to \infty$ and, as $X$ is assumed to be right-continuous, $X_s^{(n)} \to X_s$ as $n \to \infty$ for every $s \le u$.

Let us prove now that $X^{(n)}$ is progressively measurable, i.e. that if $\Gamma \in \mathscr{B}(E)$ then the event $\{(s, \omega); s \le u, X_s^{(n)}(\omega) \in \Gamma\}$ belongs to $\mathscr{B}([0, u]) \otimes \mathscr{F}_u$. This follows from the relation

$$\{(s,\omega); s \le u, X_s^{(n)}(\omega) \in \Gamma\}$$
$$= \bigcup_{k<2^n} \Big(\big[\tfrac{k}{2^n}\,u, \tfrac{k+1}{2^n}\,u\big[ \times \big\{X_{\frac{k+1}{2^n}u} \in \Gamma\big\}\Big) \cup \big(\{u\} \times \{X_u \in \Gamma\}\big)$$

and the right-hand side is an element of $\mathscr{B}([0,u]) \otimes \mathscr{F}_u$.

The map $(s,\omega) \mapsto X_s(\omega)$ is now $\mathscr{B}([0,u]) \otimes \mathscr{F}_u$-measurable, being the limit of $\mathscr{B}([0,u]) \otimes \mathscr{F}_u$-measurable functions. □

*Example 2.2* Let us see a situation where the assumption of progressive measurability is needed. Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be a process. We shall often be led to the consideration of the process $Y_t(\omega) = \int_0^t X_s(\omega)\,ds$ (assuming that $t \mapsto X_t$ is integrable for almost every $\omega$). Is the new process $(Y_t)_t$ adapted to the filtration $(\mathscr{F}_t)_t$? That is, is $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (Y_t)_t, \mathrm{P})$ a process?

Let us fix $t > 0$. If $X$ is progressively measurable, then the map $(\omega, s) \mapsto X_s(\omega), \omega \in \Omega, s \le t$, is $\mathscr{F}_t \otimes \mathscr{B}([0,t])$-measurable. Therefore by the first part of Fubini's theorem (Theorem 1.2) the map $\omega \mapsto \int_0^t X_s(\omega)\,ds = Y_t$ is $\mathscr{F}_t$-measurable. Therefore $Y$ is adapted and, being continuous, also progressively measurable. Without the assumption of progressive measurability for $X$, $Y$ might not be adapted.

*Example 2.3* Let $X$ be a stochastic process and $\zeta$ a positive r.v. Sometimes we shall be led to the consideration of the r.v. $X_\zeta$, i.e. the position of $X$ at the random time $\zeta$. Can we say that $X_\zeta$ is a r.v.?

The answer is positive if we assume that $X$ is a measurable process: in this case $(\omega, t) \mapsto X_t(\omega)$ is measurable and $X_\zeta$ is obtained as the composition of the measurable maps $\omega \mapsto (\omega, \zeta(\omega))$ and $(\omega, t) \mapsto X_t(\omega)$.

In order to be rigorous we will sometimes need to make some assumptions concerning the filtration. These will not produce difficulties, as we shall soon see that they are satisfied for our processes of interest.

Let $\mathscr{F}_{t+} = \bigcap_{\varepsilon>0} \mathscr{F}_{t+\varepsilon}$. Clearly $\mathscr{F}_{t+}$ is a $\sigma$-algebra (the intersection of any family of $\sigma$-algebras is always a $\sigma$-algebra) and $\mathscr{F}_t \subset \mathscr{F}_{t+}$; we say that the filtration is *right-continuous* if $\mathscr{F}_{t+} = \mathscr{F}_t$ for every $t$.

A process is said to be *standard* if

a) the filtration $(\mathscr{F}_t)_t$ is right-continuous;
b) for every $t$, $\mathscr{F}_t$ contains the negligible events of $\mathscr{F}$.

Note that the assumption that a process is standard concerns only the filtration and not, for instance, the r.v.'s $X_t$.

A situation where this kind of assumption is needed is the following: let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be a process and $(Y_t)_t$ a family of r.v.'s such that $X_t = Y_t$ a.s. for every $t$. In general this does not imply that $Y = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (Y_t)_t, \mathrm{P})$ is a process. Actually, $Y_t$ might not be $\mathscr{F}_t$-measurable because the negligible event $N_t = \{X_t \neq Y_t\}$ might not belong to $\mathscr{F}_t$. This problem does not appear if the space is standard. Moreover, in this case every a.s. continuous process has a continuous modification.

The fact that the filtration $(\mathscr{F}_t)_t$ is right-continuous is also a technical assumption that is often necessary; this explains why we shall prove, as soon as possible, that we can assume that the processes we are dealing with are standard.

## 2.2 Kolmogorov's continuity theorem

We have seen in Example 2.1 that a non-continuous process can have a continuous modification. The following classical theorem provides a simple criterion in order to ensure the existence of such a continuous version.

**Theorem 2.1 (Kolmogorov's continuity theorem)** Let $D \subset \mathbb{R}^m$ be an open set and $(X_y)_{y \in D}$ a family of $d$-dimensional r.v.'s on $(\Omega, \mathscr{F}, \mathrm{P})$ such that there exist $\alpha > 0$, $\beta > 0$, $c > 0$ satisfying

$$\mathrm{E}[|X_y - X_z|^\beta] \leq c|y - z|^{m+\alpha} .$$

Then there exists a family $(\widetilde{X}_y)_y$ of $\mathbb{R}^d$-valued r.v.'s such that

$$X_y = \widetilde{X}_y \qquad \text{a.s. for every } y \in D$$

(i.e. $\widetilde{X}$ is a modification of $X$) and that, for every $\omega \in \Omega$, the map $y \mapsto \widetilde{X}_y(\omega)$ is continuous and even Hölder continuous with exponent $\gamma$ for every $\gamma < \frac{\alpha}{\beta}$ on every compact subset of $D$.

In the proof we shall assume $D = ]0, 1[^m$. The key technical point is the following lemma.

**Lemma 2.1** Under the assumptions of Theorem 2.1, let $D = ]0,1[^m$ and let us denote by $D_B$ the set, which is dense in $D$, of the dyadic points (i.e. points having as coordinates fractions with powers of 2 in the denominator). Then, for every $\gamma < \frac{\alpha}{\beta}$, there exists a negligible event $N$ such that the restriction of $X$ to $D_B$ is Hölder continuous with exponent $\gamma$ on $N^c$.

*Proof* For a fixed $n$ let $A_n \subset D_B$ be the set of the points $y \in D$ whose coordinates are of the form $k2^{-n}$. Let $\gamma < \frac{\beta}{\alpha}$ and let

$$\Gamma_n = \{|X_y - X_z| > 2^{-n\gamma} \text{ for some } z, y \in A_n \text{ with } |y-z| = 2^{-n}\} .$$

If $y, z \in A_n$ are such that $|y-z| = 2^{-n}$, then, by Markov's inequality (1.6),

$$\mathrm{P}(|X_y - X_z| > 2^{-n\gamma}) \le 2^{n\beta\gamma}\mathrm{E}(|X_y - X_z|^\beta) \le 2^{n(\beta\gamma-\alpha-m)}.$$

As the set of the pairs $y, z \in A_n$ such that $|y-z| = 2^{-n}$ has cardinality $2m \cdot 2^{nm}$,

$$\mathrm{P}(\Gamma_n) \le 2m2^{nm+n(\beta\gamma-\alpha-m)} \le \text{ const } 2^{-n\mu}$$

where $\mu = \alpha - \beta\gamma > 0$. This is the general term of a convergent series and therefore, by the Borel–Cantelli lemma, there exists a negligible event $N$ such that, if $\omega \in N^c$, we have $\omega \in \Gamma_n^c$ eventually. Let us fix now $\omega \in N^c$ and let $n = n(\omega)$ be such that $\omega \in \Gamma_k^c$ for every $k > n$. Let us assume at first $m = 1$. Let $y \in D_B$: if $\nu > n$ and $y \in [i\,2^{-\nu}, (i+1)2^{-\nu}[$ then

$$y = i2^{-\nu} + \sum_{\ell=\nu+1}^{r} \alpha_\ell 2^{-\ell} ,$$

where $\alpha_\ell = 0$ or 1. By the triangle inequality

$$\begin{aligned} |X_y - X_{i2^{-\nu}}| &\le \sum_{k=\nu+1}^{r} \Big|X\Big(i2^{-\nu} + \sum_{\ell=\nu+1}^{k+1} \alpha_\ell 2^{-\ell}\Big) - X\Big(i2^{-\nu} + \sum_{\ell=\nu+1}^{k} \alpha_\ell 2^{-\ell}\Big)\Big| \\ &\le \sum_{k=1}^{r} 2^{-(\nu+k)\gamma} \le \frac{1}{1-2^{-\gamma}}\, 2^{-\nu\gamma} . \end{aligned}$$

Let now $y, z \in D_B$ be such that $|y-z| \le 2^{-\nu}$; there are two possibilities: if there exists an $i$ such that $(i-1)2^{-\nu} \le y < i2^{-\nu} \le z < (i+1)2^{-\nu}$ then

$$|X_y - X_z| \le |X_z - X_{i2^{-\nu}}| + |X_y - X_{(i-1)2^{-\nu}}| + |X_{i2^{-\nu}} - X_{(i-1)2^{-\nu}}| \le \Big(1 + \frac{2}{1-2^{-\gamma}}\Big) 2^{-\nu\gamma}.$$

Otherwise $y, z \in [i2^{-\nu}, (i+1)2^{-\nu}[$ and then

$$|X_y - X_z| \le |X_y - X_{i2^{-\nu}}| + |X_z - X_{i2^{-\nu}}| \le \frac{2}{1-2^{-\gamma}} 2^{-\nu\gamma}\ .$$

Thus we have proved that if $y, z \in D_B$ and $|y-z| \le 2^{-\nu}$ then $|X_y - X_z| \le k 2^{-\nu\gamma}$, where $k$ does not depend on $\nu$ or, equivalently, that if $y, z \in D_B$ and $|y-z| \le 2^{-\nu}$ then

$$|X_y - X_z| \le k|y-z|^\gamma$$

for every $\nu > n$. The lemma is therefore proved if $m = 1$. Let us now consider the case $m > 1$. We can repeat the same argument as in dimension 1 and derive that the previous relation holds as soon as $y$ and $z$ differ at most by one coordinate. Let us define $x^{(i)} \in \mathbb{R}^m$, for $i = 0, \dots, m$, by

$$x_j^{(i)} = \begin{cases} y_i & \text{if} \quad j \le i \\ z_j & \text{if} \quad j > 1\ . \end{cases}$$

Therefore $x^{(0)} = z$, $x^{(m)} = y$ and $x^{(i)}$ and $x^{(i+1)}$ have all but one of their coordinates equal, and then

$$|X_y - X_z| \le \sum_{i=1}^m |X_{x^{(i)}} - X_{x^{(i-1)}}| \le k \sum_{i=1}^m |x^{(i)} - x^{(i-1)}|^\gamma \le mk|y-z|^\gamma\ ,$$

which allows us to conclude the proof.

□

*Proof* of Theorem 2.1 By Lemma 2.1, if $\gamma < \frac{\alpha}{\beta}$ then there exists a negligible event $N$ such that if $\omega \notin N$, the restriction of $y \mapsto X_y(\omega)$ to $D_B$ is Hölder continuous with exponent $\gamma$ and hence uniformly continuous. Let us denote by $\widetilde{X}_.(\omega)$ its unique continuous extension to $D$. It suffices now to prove that $\widetilde{X}_y = X_y$ a.s. This fact is obvious if $y \in D_B$. Otherwise let $(y_n)_n \subset D_B$ be a sequence converging to $y$. The assumption and Markov's inequality (1.6) imply that $X_{y_n} \to_{n\to\infty} X_y$ in probability and therefore, possibly taking a subsequence, $X_{y_n} \to_{n\to\infty} X_y$ a.s. As $\widetilde{X}_{y_n} = X_{y_n}$ a.s. for every $n$, this implies that $X_y = \widetilde{X}_y$ a.s. □

**Corollary 2.1** Let $X$ be an $\mathbb{R}^d$-valued process such that there exist $\alpha > 0$, $\beta > 0$, $c > 0$ satisfying, for every $s, t$,

$$\mathrm{E}[|X_t - X_s|^\beta] \le c|t-s|^{1+\alpha} .$$

Then there exists a modification $Y$ of $X$ that is continuous. Moreover, for every $\gamma < \frac{\alpha}{\beta}$ the paths of $Y$ are Hölder continuous with exponent $\gamma$ on every bounded time interval.

*Example 2.4* In the next chapter we shall see that a *Brownian motion* is a real-valued process $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\ge 0}, (B_t)_{t\ge 0}, \mathrm{P})$ such that

i) $B_0 = 0$ a.s.;
ii) for every $0 \le s \le t$ the r.v. $B_t - B_s$ is independent of $\mathscr{F}_s$;
iii) for every $0 \le s \le t$ $B_t - B_s$ is $N(0, t-s)$-distributed.

Let us show that a Brownian motion has a continuous modification. It is sufficient to check the condition of Corollary 2.1. Let $t > s$; as $B_t - B_s \sim N(0, t-s)$, we have $B_t - B_s = (t-s)^{1/2} Z$ with $Z \sim N(0,1)$. Therefore

$$\mathrm{E}[|B_t - B_s|^\beta] = (t-s)^{\beta/2}\mathrm{E}[|Z|^\beta] .$$

As $\mathrm{E}[|Z|^\beta] < +\infty$ for every $\beta > 0$, we can apply Corollary 2.1 with $\alpha = \frac{\beta}{2} - 1$. Hence a Brownian motion has a continuous version, which is also Hölder continuous with exponent $\gamma$ for every $\gamma < \frac{1}{2} - \frac{1}{\beta}$; i.e., $\beta$ being arbitrary, for every $\gamma < \frac{1}{2}$.

## 2.3 Construction of stochastic processes

Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in T}, (X_t)_{t\in T}, \mathrm{P})$ be a process taking values in the topological space $(E, \mathscr{B}(E))$ and $\pi = (t_1, \dots, t_n)$ an $n$-tuple of elements of $T$ with $t_1 < \dots < t_n$. Then we can consider the r.v.

$$X_\pi = (X_{t_1}, \dots, X_{t_n}) : \Omega \to E^n = E \times \dots \times E$$

and denote by $\mu_\pi$ its distribution. The probabilities $\mu_\pi$ are called the *finite-dimensional distributions* of the process $X$.

Note that two processes have the same finite-dimensional distributions if and only if they are equivalent.

This family of probabilities is quite important: the finite-dimensional distributions characterize P in the following sense.

**Proposition 2.2** Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in T}, (X_t)_{t\in T}, \mathrm{P})$ and $X' = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in T}, (X_t)_{t\in T}, \mathrm{P}')$ be two processes (defined on the same $(\Omega, \mathscr{F})$) having the same finite-dimensional distributions. Then P and P′ coincide on the $\sigma$-algebra $\sigma(X_t, t \in T)$.

*Proof* If $\Gamma_1, \dots, \Gamma_n \in \mathscr{B}(E)$, as $X$ and $X'$ have the same finite-dimensional distributions,

$$\mathrm{P}(X_{t_1} \in \Gamma_1, \dots, X_{t_n} \in \Gamma_n) = \mu_\pi(\Gamma_1 \times \cdots \times \Gamma_n) = \mathrm{P}'(X_{t_1} \in \Gamma_1, \dots, X_{t_n} \in \Gamma_n) ,$$

i.e. P and P′ coincide on the events of the type

$$\{X_{t_1} \in \Gamma_1, \dots, X_{t_n} \in \Gamma_n\} .$$

As these events form a class that is stable with respect to finite intersections, thanks to Carathéodory's criterion, Theorem 1.1, P and P′ coincide on the generated $\sigma$-algebra, i.e. $\sigma(X_t, t \in T)$.

□

A very important problem which we are going to be confronted with later is the converse: given a topological space $E$, a time span $T$ and a family $(\mu_\pi)_{\pi\in\Pi}$ of finite-dimensional distributions ($\Pi$ = all possible $n$-tuples of distinct elements of $T$ for $n$ ranging over the positive integers), does an $E$-valued stochastic process having $(\mu_\pi)_{\pi\in\Pi}$ as its family of finite-dimensional distributions exist?

It is clear, however, that the $\mu_\pi$'s cannot be anything. For instance, if $\pi = \{t_1, t_2\}$ and $\pi' = \{t_1\}$, then if the $(\mu_\pi)_{\pi\in\Pi}$ were the finite-dimensional distributions of some process $(X_t)_t$, $\mu_\pi$ would be the law of $(X_{t_1}, X_{t_2})$ and $\mu_{\pi'}$ the law of $X_{t_1}$. Therefore $\mu_{\pi'}$ would necessarily be the first marginal of $\mu_\pi$. This can also be stated by saying that $\mu_{\pi'}$ is the image of $\mu_\pi$ through the map $p : E\times E \to E$ given by $p(x_1, x_2) = x_1$.

More generally, in order to be the family of finite-dimensional distributions of some process $X$, the family $(\mu_\pi)_{\pi\in\Pi}$ must necessarily satisfy the following consistency condition.

**Condition 2.1** Let $\pi = (t_1, \dots, t_n)$ and $\pi' = (t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$. Let $p_i : E^n \to E^{n-1}$ be the map defined as $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$. Then the image of $\mu_\pi$ through $p_i$ is equal to $\mu_{\pi'}$.

The next theorem states that Condition 2.1 is also sufficient for $(\mu_\pi)_{\pi \in \Pi}$ to be the system of finite-dimensional distributions of some process $X$, at least if the topological space $E$ is sufficiently regular.

**Theorem 2.2 (Kolmogorov's existence theorem)** Let $E$ be a complete separable metric space, $(\mu_\pi)_{\pi \in \Pi}$ a system of finite-dimensional distributions on $E$ satisfying Condition 2.1. Let $\Omega = E^T$ ($E^T$ is the set of all paths $T \to E$) and let us define $X_t(\omega) = \omega(t)$, $\mathscr{F}_t = \sigma(X_s, s \le t)$, $\mathscr{F} = \sigma(X_t, t \in T)$. Then there exists on $(\Omega, \mathscr{F})$ a unique probability P such that $(\mu_\pi)_{\pi \in \Pi}$ is the family of the finite-dimensional distributions of the process $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \in T}, (X_t)_{t \in T}, \mathrm{P})$.

Uniqueness is of course a consequence of Proposition 2.2.

*Example 2.5 (Gaussian processes)* An $\mathbb{R}^m$-valued process $(X_t)_t$ is said to be *Gaussian* if it is a Gaussian family i.e. if its finite-dimensional distributions are Gaussian. If we define

$$\begin{aligned} b_t &= \mathrm{E}(X_t), \text{ its } \textit{mean function} \\ K^{ij}_{s,t} &= \mathrm{Cov}(X_i(s), X_j(t)), 1 \le i,j \le m, \text{ its } \textit{covariance function}\ , \end{aligned}$$

then the finite-dimensional distributions of $(X_t)_t$ are completely determined by these two quantities.

Let us, for instance, determine the distribution of the $2m$-dimensional r.v. $(X_s, X_t)$, $s < t$. By assumption it is Gaussian (Proposition 1.8) and has mean $(b_s, b_t)$. As $K_{t,t}$ is the $m \times m$ covariance matrix of $X_t$, the covariance matrix of $(X_s, X_t)$ is the $2m \times 2m$ block matrix

$$\begin{pmatrix} K_{t,t} & K_{s,t} \\ K_{t,s} & K_{s,s} \end{pmatrix}\ .$$

The law of $(X_s, X_t)$ is therefore determined by the functions $K$ and $b$. Similarly, if $0 \le t_1 \le \dots \le t_n$, then the r.v. $(X_{t_1}, \dots, X_{t_n})$ is Gaussian, has a covariance matrix with blocks $K_{t_i,t_i}$ on the diagonal and blocks $K_{t_i,t_j}$ outside the diagonal and mean $(b_{t_1}, \dots, b_{t_n})$.

(continued)

(continued)

Conversely, given a mean and covariance functions, does an associated Gaussian process exist?

We must first point out that the covariance function must satisfy an important property. Let us consider for simplicity the case $m = 1$, i.e. of a real-valued process $X$. A real function $(s, t) \mapsto C(s, t)$ is said to be a *positive definite kernel* if, for every choice of $t_1, \dots, t_n \in \mathbb{R}^+$ and $\xi_1, \dots, \xi_n \in \mathbb{R}$,

$$\sum_{i,j=1}^{n} C(t_i, t_j)\xi_i\xi_j \ge 0 \tag{2.1}$$

i.e. if the matrix $(C(t_i, t_j))_{i,j}$ is positive definite.

The covariance function $K$ is necessarily a positive definite kernel: actually $(K_{t_i,t_j})_{ij}$ is the covariance matrix of the random vector $(X_{t_1}, \dots, X_{t_n})$ and every covariance matrix is positive definite (as was shown in (1.14)).

Conversely, let us prove that if $K$ is a positive definite kernel, then there exists a Gaussian process $X$ associated to it. This is a simple application of Kolmogorov's existence Theorem 2.2.

If $\pi = (t_1, \dots, t_m)$, let us define the finite-dimensional distribution $\mu_\pi$ of $X$ to be Gaussian with covariance matrix

$$\Gamma_{ij} = K_{t_i,t_j}\ .$$

Let us check that this family of finite-dimensional distributions satisfies the coherence Condition 2.1. If $\pi' = (t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_m)$ then $\mu_{\pi'}$ is Gaussian with covariance matrix obtained by removing the $i$-th row and the $i$-th column from $\Gamma$. Let us denote by $\Gamma'$ such a matrix. We must check that $\mu_{\pi'}$ coincides with $p_i(\mu_\pi)$, where $p_i$ is the map $p_i(x_1, \dots, x_n) = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$.

But this is immediate: $p_i(\mu_\pi)$ is also a Gaussian law, being the image of a Gaussian distribution through the linear mapping $p_i$ (recall Theorem 1.3). Moreover, the covariance matrix of $p_i(\mu_\pi)$ is equal to $\Gamma'$. To see this simply observe that if $(X_1, \dots, X_m)$ is a r.v. having law $\mu_\pi$ and therefore covariance matrix $\Gamma$, then $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_m)$ has distribution

$$p_i(\mu_\pi)$$

and covariance matrix $\Gamma'$.

Therefore if $K$ is a positive definite kernel, then there exists a stochastic process $X$ which is Gaussian with covariance function $K$ and mean function $b_t \equiv 0$. Now, for every mean function $b$ the process $Y_t = X_t + b_t$ is Gaussian with covariance function $K$ and mean function $b$.

## 2.4 Next...

In this chapter we have already met some of the relevant problems which arise in the investigation of stochastic processes:

a) the construction of processes satisfying particular properties (that can be reduced to finite-dimensional distributions); for instance, in the next chapter we shall see that it is immediate, from its definition, to determine the finite-dimensional distributions of a Brownian motion;
b) the regularity of the paths (continuity, ...);
c) the determination of the probability P of the process, i.e. the computation of the probability of events connected to it. For instance, for a Brownian motion $B$, what is the value of $\mathrm{P}(\sup_{0\le s\le t} B_s \ge 1)$?

Note again that, moving from a process to one of its modifications, the finite-dimensional distributions do not change, whereas other properties, such as regularity of the paths, can turn out to be very different, as in Example 2.1.

In the next chapters we investigate a particular class of processes: diffusions. We shall be led to the development of particular techniques (stochastic integral) that, together with the two Kolmogorov's theorems, will allow us first to prove their existence and then to construct continuous versions. The determination of P, besides some particular situations, will in general not be so simple. We shall see, however, that the probability of certain events or the expectations of some functionals of the process can be obtained by solving suitable PDE problems. Furthermore these quantities can be computed numerically by methods of simulation.

These processes (i.e. diffusions) are very important

a) first because there are strong links with other areas of mathematics (for example, the theory of PDEs, but in other fields too)
b) but also because they provide models in many applications (control theory, filtering, finance, telecommunications, ...). Some of these aspects will be developed in the last chapter.

## Exercises

**2.1** (p. 451) Let $X$ and $Y$ be two processes that are modifications of one another.
  a) Prove that they are equivalent.
  b) Prove that if the time set is $\mathbb{R}^+$ or a subinterval of $\mathbb{R}^+$ and $X$ and $Y$ are both a.s. continuous, then they are indistinguishable.

**2.2** (p. 451) Let $(X_t)_{0\le t\le T}$ be a *continuous* process and $D$ a dense subset of $[0,T]$.
  a) Show that $\sigma(X_t, t\le T) = \sigma(X_t, t\in D)$.
  b) What if $(X_t)_{0\le t\le T}$ was only right-continuous?

**2.3** (p. 452) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t$ be a progressively measurable process with values in the measurable space $(E, \mathscr{E})$. Let $\psi : E \to G$ be a measurable function into the measurable space $(G, \mathscr{G})$. Prove that the $G$-valued process $t \mapsto \psi(X_t)$ is also progressively measurable.

**2.4** (p. 452) a) Let $Z_\infty, Z_n$, $n = 1, \dots,$ be r.v.'s on some probability space $(\Omega, \mathscr{F}, \mathrm{P})$. Prove that the event $\{\lim_{n\to\infty} Z_n \neq Z_\infty\}$ is equal to

$$\bigcup_{m=1}^{\infty} \bigcap_{n_0=1}^{\infty} \bigcup_{n\ge n_0} \{|Z_n - Z_\infty| \ge \tfrac{1}{m}\} . \tag{2.2}$$

b) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in T}, (X_t)_{t\in T}, \mathrm{P})$ and $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, (\widetilde{\mathscr{F}}_t)_{t\in T}, (\widetilde{X}_t)_{t\in T}, \widetilde{\mathrm{P}})$ be equivalent processes.

b1) Let $(t_n)_n \subset T$. Let us assume that there exists a number $\ell \in \mathbb{R}$ such that

$$\lim_{n\to\infty} X_{t_n} = \ell \qquad \text{a.s.}$$

Then also

$$\lim_{n\to\infty} \widetilde{X}_{t_n} = \ell \qquad \text{a.s.}$$

b2) Let $t \in \overline{T}$. Then if the limit

$$\lim_{s\to t, s\in\mathbb{Q}} X_s = \ell$$

exists for some $\ell \in \mathbb{R}$ then also

$$\lim_{s\to t, s\in\mathbb{Q}} \widetilde{X}_s = \ell . \tag{2.3}$$

**2.5** (p. 453) An example of a process that comes to mind quite naturally is so-called "white noise", i.e. a process $(X_t)_t$ defined for $t \in [0, 1]$, say, and such that the r.v.'s $X_t$ are identically distributed centered and square integrable and $X_t$ and $X_s$ are independent for every $s \neq t$.

In this exercise we prove that a white noise cannot be a measurable process, unless it is $\equiv 0$ a.s. Let therefore $(X_t)_t$ be a measurable white noise.

a) Prove that, for every $a, b \in [0, 1]$, $a \le b$,

$$\int_a^b \int_a^b \mathrm{E}(X_s X_t)\, ds\, dt = 0 . \tag{2.4}$$

b) Show that for every $a, b \in [0, 1]$

$$\mathrm{E}\Big[\Big(\int_a^b X_s\,ds\Big)^2\Big] = 0\,.$$

c) Deduce that necessarily $X_t(\omega) = 0$ a.e. in $t$ for almost every $\omega$.

**2.6** (p. 454) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (Z_t)_t, \mathrm{P})$ be an $\mathbb{R}^m$-valued continuous process. Let $\psi_Z : \omega \to \{t \to Z_t(\omega)\}$ be the application associating to every $\omega \in \Omega$ the corresponding path of the process. $\psi_Z$ is a map with values in the space $\mathscr{C} = \mathscr{C}(\mathbb{R}^+, \mathbb{R}^m)$. Let us consider on $\mathscr{C}$ the topology of uniform convergence on compact sets and denote by $\mathscr{M}$ the associated Borel $\sigma$-algebra. We want to prove that $\psi_Z$ is measurable.

a) Let, for $\gamma \in \mathscr{C}$, $t > 0$, $\varepsilon > 0$, $A_{\gamma,t,\varepsilon} = \{w \in \mathscr{C}; |\gamma_t - w_t| \le \varepsilon\}$. Prove that $\psi_Z^{-1}(A_{\gamma,t,\varepsilon}) \in \mathscr{F}$.

b) For $\gamma \in \mathscr{C}, T > 0, \varepsilon > 0$, let

$$\overline{U}_{\gamma,T,\varepsilon} = \{w \in \mathscr{C}; |\gamma_t - w_t| \le \varepsilon \text{ for every } t \in [0, T]\} \tag{2.5}$$

(a closed tube of radius $\varepsilon$ around the path $\gamma$). Prove that $\psi_Z^{-1}(\overline{U}_{\gamma,T,\varepsilon}) \in \mathscr{F}$.

c) Prove that $\psi_Z$ is a $(\mathscr{C}, \mathscr{M})$-valued r.v.

b) As the paths of $\mathscr{C}$ are continuous, $\overline{U}_{\gamma,T,\varepsilon} = \{w \in \mathscr{C}; |\gamma_r - w_r| \le \varepsilon \text{ for every } r \in [0, T] \cap \mathbb{Q}\}$, which is a countable intersection of events of the form $A_{\gamma,t,\varepsilon}$. c) Recall Exercise 1.19.

# Chapter 3
# Brownian Motion

Brownian motion is a particular stochastic process which is the prototype of the class of processes which will be our main concern. Its investigation is the object of this chapter.

## 3.1 Definition and general facts

We already know from the previous chapter the definition of a Brownian motion.

**Definition 3.1** A real-valued process $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\ge 0}, (B_t)_{t\ge 0}, \mathrm{P})$ is a Brownian motion if

i) $B_0 = 0$ a.s.;
ii) for every $0 \le s \le t$ the r.v. $B_t - B_s$ is independent of $\mathscr{F}_s$;
iii) for every $0 \le s \le t$ $B_t - B_s$ has law $N(0, t-s)$.

*Remarks 3.1*

a) ii) of Definition 3.1 implies that $B_t - B_s$ is independent of $B_u$ for every $u \le s$ and even from $\sigma(B_u, u \le s)$, which is a $\sigma$-algebra that is contained in $\mathscr{F}_s$. Intuitively this means that the increments of the process after time $s$ are independent of the path of the process up to time $s$.
b) A Brownian motion is a Gaussian process, i.e. the joint distributions of $B_{t_1}, \dots, B_{t_m}$ are Gaussian. Let $\alpha_1, \dots, \alpha_m \in \mathbb{R}$, $0 \le t_1 < t_2 < \dots < t_m$: we must prove that $\alpha_1 B_{t_1} + \dots + \alpha_m B_{t_m}$ is a normal r.v., so that we can apply Proposition 1.8. This is obvious if $m = 1$, as Definition 3.1 with $s = 0$

(continued)

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_3

*Remarks 3.1* (continued)

states that $B_t \sim N(0,t)$. Let us assume this fact true for $m-1$ and let us prove it for $m$: we can write

$$\alpha_1 B_{t_1} + \cdots + \alpha_m B_{t_m} = [\alpha_1 B_{t_1} + \cdots + (\alpha_{m-1} + \alpha_m) B_{t_{m-1}}] + \alpha_m (B_{t_m} - B_{t_{m-1}}) .$$

This is a normal r.v., as we have seen in Sect. 1.7, being the sum of two *independent* normal r.v.'s (the r.v. between [ ] is $\mathscr{F}_{t_{m-1}}$-measurable whereas $B_{t_m} - B_{t_{m-1}}$ is independent of $\mathscr{F}_{t_{m-1}}$, thanks to ii) of Definition 3.1).

c) For every $0 \le t_0 < \cdots < t_m$ the real r.v.'s $B_{t_k} - B_{t_{k-1}}, k = 1, \dots, m$, are independent: they are actually jointly Gaussian and pairwise uncorrelated.

d) Sometimes it will be important to specify with respect to which filtration a Brownian motion is considered. When the probability space is fixed we shall say that $B$ is an $(\mathscr{F}_t)_t$-Brownian motion in order to specify that $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ is a Brownian motion. Of course, for every $t$ the $\sigma$-algebra $\mathscr{F}_t$ must necessarily contain the $\sigma$-algebra $\mathscr{G}_t = \sigma(B_s, s \le t)$ (otherwise $(B_t)_t$ would not be adapted to $(\mathscr{F}_t)_t$). It is also clear that if $B$ is an $(\mathscr{F}_t)_t$-Brownian motion it is *a fortiori* a Brownian motion with respect to every other filtration $(\mathscr{F}'_t)_t$ that is smaller than $(\mathscr{F}_t)_t$, (i.e. such that $\mathscr{F}'_t \subset \mathscr{F}_t$ for every $t \ge 0$) provided that $B$ is adapted to $(\mathscr{F}'_t)_t$ i.e provided that $(\mathscr{F}'_t)_t$ contains the natural filtration (see p. 31). Actually if $B_t - B_s$ is independent of $\mathscr{F}_s$, *a fortiori* it will be independent of $\mathscr{F}'_s$.

We shall speak of *natural Brownian motion* when $(\mathscr{F}_t)_t$ is the natural filtration.

**Proposition 3.1** If $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ is a Brownian motion then

1) $B_0 = 0$ a.s.;
2) for every $0 \le t_1 < \cdots < t_m$, $(B_{t_1}, \dots, B_{t_m})$ is an $m$-dimensional centered normal r.v.;
3) $\mathrm{E}[B_s B_t] = s \wedge t$.

Conversely, properties 1), 2) and 3) imply that $B$ is a *natural* Brownian motion.

*Proof* If $B$ is a Brownian motion, 1) is obvious; 2) is Remark 3.1 b). As for 3), if $s \le t$, as $B_t - B_s$ and $B_s$ are centered and independent

$$\mathrm{E}[B_t B_s] = \mathrm{E}[\{(B_t - B_s) + B_s\} B_s] = \mathrm{E}[(B_t - B_s) B_s] + \mathrm{E}[B_s^2] = s = s \wedge t .$$

Conversely, if $B$ satisfies 1), 2) and 3), then i) of Definition 3.1 is obvious. Moreover, for $0 \le s < t$, $B_t - B_s$ is a normal r.v., being a linear function of $(B_s, B_t)$, and is centered since both $B_t$ and $B_s$ are so; as, for $s \le t$,

$$\mathrm{E}[(B_t - B_s)^2] = \mathrm{E}[B_t^2] + \mathrm{E}[B_s^2] - 2\mathrm{E}[B_t B_s] = t + s - 2s = t - s \; ,$$

$B_t - B_s$ has an $N(0, t-s)$ distribution and iii) is satisfied. To complete the proof we must show that $B_t - B_s$ is independent of $\sigma(B_u, u \le s)$; Brownian motion being a Gaussian process, thanks to Remark 1.2 it is sufficient to prove that $B_t - B_s$ is uncorrelated with $B_u$ for every $u \le s$. Actually, if $u \le s \le t$,

$$\mathrm{E}[(B_t - B_s)B_u] = t \wedge u - s \wedge u = 0$$

and $B_t - B_s$ is therefore independent of $\sigma(B_u, u \le s)$. □

We have not yet proved the existence of a process satisfying the conditions of Definition 3.1, but this is a simple application of Remark 2.5. Actually it is sufficient to verify that, if $t_1 \le t_2 \le \dots \le t_m$,

$$K(t_i, t_j) = t_i \wedge t_j$$

is a positive definite kernel, i.e. that the matrix $\Gamma$ with entries $\Gamma_{ij} = t_i \wedge t_j$ is positive definite. The simplest way to check this fact is to produce a r.v. having $\Gamma$ as a covariance matrix, every covariance matrix being positive definite as pointed out on p. 10.

Let $Z_1, \dots, Z_m$ be independent centered Gaussian r.v.'s with $\mathrm{Var}(Z_i) = t_i - t_{i-1}$, with the understanding that $t_0 = 0$. Then it is immediate that the r.v. $(X_1, \dots, X_m)$ with $X_i = Z_1 + \dots + Z_i$ has covariance matrix $\Gamma$: as the r.v.'s $Z_k$ are independent we have, for $i \le j$,

$$\begin{aligned} \mathrm{Cov}(X_i, X_j) &= \mathrm{Cov}(Z_1 + \dots + Z_i, Z_1 + \dots + Z_j) = \mathrm{Var}(Z_1) + \dots + \mathrm{Var}(Z_i) \\ &= t_1 + (t_2 - t_1) + \dots + (t_i - t_{i-1}) = t_i = t_i \wedge t_j \; . \end{aligned}$$

The next statement points out that Brownian motion is invariant with respect to certain transformations.

**Proposition 3.2** Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion. Then

$$X_t = B_{t+s} - B_s, \qquad -B_t, \qquad cB_{t/c^2},$$

$$Z_t = \begin{cases} tB_{1/t} & t > 0 \\ 0 & t = 0 \end{cases}$$

are also Brownian motions, the first one with respect to the filtration $(\mathscr{F}_{t+s})_t$, the second one with respect to $(\mathscr{F}_t)_t$, and the third one with respect to $(\mathscr{F}_{t/c^2})_t$. $(Z_t)_t$ is a natural Brownian motion.

*Proof* Let us prove that $X$ is a Brownian motion. The condition $X_0 = 0$ is immediate as is the fact that, if $v < u$, then $X_u - X_v = B_{u+s} - B_s - B_{v+s} + B_s = B_{u+s} - B_{v+s}$ is $N(0, u - v)$-distributed. This also gives that the increment $X_u - X_v$ is independent of $\mathscr{F}_{v+s}$, so that $X$ is an $(\mathscr{F}_{t+s})_t$-Brownian motion.

Let us now prove that $Z$ is a Brownian motion, the other situations being similar to the case of $X$. The proof amounts to checking that the three conditions of Proposition 3.1 are satisfied.

1) of Proposition 3.1 is immediate, 2) follows from the fact that $Z$ is also a Gaussian family; actually $(Z_{t_1}, \dots, Z_{t_m})$ can be written as a linear transformation of $(B_{1/t_1}, \dots, B_{1/t_m})$, which is Gaussian. Finally, if $s < t$ then $\frac{1}{s} \wedge \frac{1}{t} = \frac{1}{t}$ and

$$\mathrm{E}[Z_s Z_t] = st\, \mathrm{E}[B_{1/s} B_{1/t}] = st\, (\tfrac{1}{s} \wedge \tfrac{1}{t}) = s = s \wedge t$$

and therefore 3) is satisfied. □

We shall speak of the invariance properties of Proposition 3.2 as the *scaling properties* of the Brownian motion.

**Definition 3.2** An $\mathbb{R}^m$-valued process $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ is an $m$-dimensional Brownian motion if

a) $X_0 = 0$ a.s.;
b) for every $0 \le s \le t$, the r.v. $X_t - X_s$ is independent of $\mathscr{F}_s$;
c) for every $0 \le s \le t$, $X_t - X_s$ is $N(0, (t - s)I)$-distributed ($I$ is the $m \times m$ identity matrix).

*Remarks 3.2* Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion.

a) The real processes $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_i(t))_t, \mathrm{P})$ are Brownian motions for $i = 1, \dots, m$. Definition 3.1 is immediately verified.
b) The $\sigma$-algebras $\sigma(X_i(t), t \ge 0)$, $i = 1, \dots, m$, are independent. Keeping in mind Remark 1.2 c), it is sufficient to check that $\mathrm{E}[X_i(t) X_j(s)] = 0$, for every $s \le t$, $i \ne j$. But

$$\mathrm{E}[X_i(t) X_j(s)] = \mathrm{E}[(X_i(t) - X_i(s)) X_j(s)] + \mathrm{E}[X_i(s) X_j(s)]$$

and the first term on the right-hand side vanishes, $X_j(s)$ and $X_i(t) - X_i(s)$ being independent and centered, the second one vanishes too as the covariance matrix of $X_s$ is diagonal.

Therefore the components $(X_i(t))_t$, $i = 1, \dots, m$, of an $m$-dimensional Brownian motion are *independent* real Brownian motions.

*Example 3.1* Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion. Let us prove that if $z \in \mathbb{R}^m$ and $|z| = 1$, then the process $X_t = \langle z, B_t \rangle$ is a real Brownian motion.

Since, for $t \geq s$, $B_t - B_s$ is independent of $\mathscr{F}_s$, the r.v.

$$X_t - X_s = \langle z, B_t \rangle - \langle z, B_s \rangle = \langle z, B_t - B_s \rangle$$

is itself independent of $\mathscr{F}_s$. Moreover, $X_t - X_s = \langle z, B_t - B_s \rangle$ is Gaussian, being a linear combination of $B_1(t) - B_1(s), \dots, B_m(t) - B_m(s)$, which are jointly Gaussian. The variance of $X_t - X_s$ is computed immediately, being a linear function of $B_t - B_s$: (1.13) gives

$$\mathrm{Var}(X_t - X_s) = (t - s) z^* z = (t - s)|z|^2 = t - s\ .$$

As a) of Definition 3.1 is obvious, this proves that $X$ is a Brownian motion with respect to the filtration $(\mathscr{F}_t)_t$. Conditions 1), 2), 3) of Proposition 3.1 are also easily checked. Note. however, that these only ensure that $X$ is a *natural* Brownian motion, which is a weaker result.

This, similarly to Proposition 3.2, is a typical exercise: given a Brownian motion $B$, prove that some transformation of it gives rise to a Brownian motion. By now we have the choice of checking that the conditions of Definition 3.1 or of Proposition 3.1 are satisfied. Note that Proposition 3.1 is useful only if we have to prove that our candidate process is a natural Brownian motion. Later (starting from Chap. 8) we shall learn more powerful methods in order to prove that a given process is a Brownian motion.

Of course a) of Remark 3.2 is a particular case of this example: just choose as $z$ the unit vector along the $i$-th axis.

We already know (Example 2.4) that a Brownian motion has a continuous modification. Note that the argument of Example 2.4 also works for an $m$-dimensional Brownian motion. From now on, by "Brownian motion" we shall always understand a Brownian motion that is continuous. Figure 3.1 provides a typical example of a path of a two-dimensional Brownian motion.

If we go back to Proposition 3.2, if $B$ is a continuous Brownian motion, then the "new" Brownian motions $X$, $-B$, $(cB_{t/c^2})_t$ are also obviously continuous. For $Z$ instead a proof is needed in order to have continuity at 0. In order to do this, note that the processes $(B_t)_t$ and $(Z_t)_t$ are equivalent. Therefore, as $B$ is assumed to be continuous,

$$\lim_{t \to 0+} B_t = 0\ ,$$

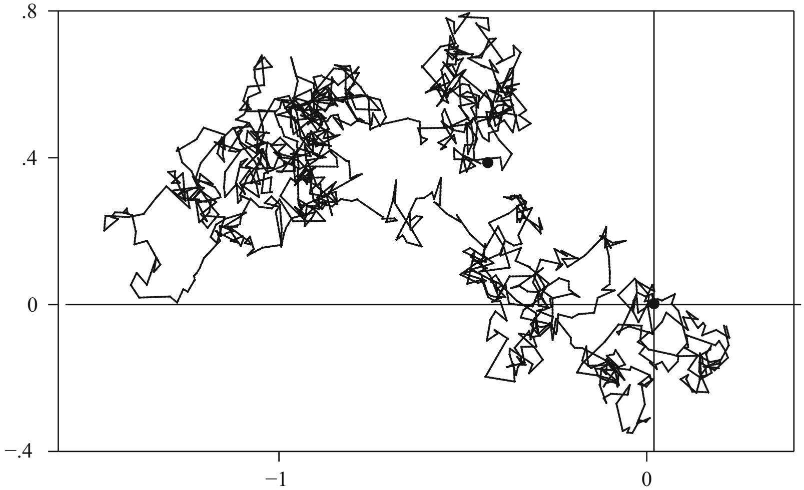


**Fig. 3.1** A typical image of a path of a two-dimensional Brownian motion for $0 \le t \le 1$ (a black small circle denotes the origin and the position at time 1). For information about the simulation of Brownian motion see Sect. 3.7

therefore (see Exercise 2.4 b2))

$$\lim_{t\to 0+, t\in\mathbb{Q}} Z_t = 0$$

from which, the process $(Z_t)_{t,t>0}$ being continuous, we derive

$$\lim_{t\to 0+} Z_t = 0\,.$$

It will be apparent in the sequel that it is sometimes important to specify the filtration with respect to which a process $B$ is a Brownian motion. The following remark points out a particularly important typical filtration. Exercise 3.5 deals with a similar question.

*Remark 3.1 (The augmented natural filtration)* Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and let $\overline{\mathscr{F}}_t$ be the $\sigma$-algebra that is obtained by adding to $\mathscr{F}_t$ the negligible events of $\mathscr{F}$. Let us prove that $B$ is a Brownian motion also with respect to $(\overline{\mathscr{F}}_t)_t$.

Actually, we must only prove that the increment $B_t - B_s$, $s \le t$, is independent of $\overline{\mathscr{F}}_s$. Let us denote by $\mathscr{N}$ the family of the negligible events of $\mathscr{F}$. Let $\mathscr{C}$ be the class of the events of the form $A \cap G$ with $A \in \mathscr{F}_s$ and $G \in \mathscr{N}$ or $G = \Omega$. $\mathscr{C}$ is stable with respect to finite intersections and

(continued)

*Remark 3.1* (continued)
contains $\mathscr{F}_s$ (one chooses $G = \Omega$) and $\mathscr{N}$, and therefore generates $\overline{\mathscr{F}}_s$. By Remark 1.1, in order to prove that $B_t - B_s$ and $\overline{\mathscr{F}}_s$ are independent we just have to check that, for every Borel set $C \subset \mathbb{R}$ and $A$ and $G$ as above,

$$\mathrm{P}(\{B_t - B_s \in C\} \cap A \cap G) = \mathrm{P}(\{B_t - B_s \in C\})\mathrm{P}(A \cap G) ,$$

which is immediate. If $G = \Omega$ the relation above is true because then $A \cap G = A$ and $B$ is an $(\mathscr{F}_t)_t$-Brownian motion, whereas if $G \in \mathscr{N}$, both members are equal to 0.

In particular, $B$ is a Brownian motion with respect to the augmented natural filtration, i.e. with respect to the filtration $(\overline{\mathscr{G}}_t)_t$ that is obtained by adding the negligible events of $\mathscr{F}$ to the natural filtration $(\mathscr{G}_t)_t$.

We shall see in Sect. 4.5 that $(\overline{\mathscr{G}}_t)_t$ is also right-continuous, so that the Brownian motion with respect to the natural augmented filtration is a standard process (see the definition on p. 34).

*Remark 3.2* Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion. We know that each of the increments $B_t - B_s$, $t > s$, is independent of $\mathscr{F}_s$. Actually a stronger statement holds: the $\sigma$-algebra $\sigma(B_t - B_s, t \geq s)$ is independent of $\mathscr{F}_s$.

This fact is almost obvious if $\mathscr{F}_s = \sigma(B_u, u \leq s)$, i.e. if $B$ is a natural Brownian motion, as the families of r.v.'s $\{B_u, u \leq s\}$ and $\{B_v - B_s, v \geq s\}$ are jointly Gaussian and pairwise non-correlated, so that we can apply the criterion of Remark 1.2. It requires some care if the filtration is not the natural one. The proof is left as an exercise (Exercise 3.4).

Note that this is a useful property. For instance, it allows us to state that the r.v. $(B_{t_1} - B_s) + (B_{t_2} - B_s)$, $t_1, t_2 > s$, is independent of $\mathscr{F}_s$. Recall that if $X, Y$ are r.v.'s that are independent of a $\sigma$-algebra $\mathscr{G}$, it is not true in general that their sum is also independent of $\mathscr{G}$. Here, however, we can argue that $(B_{t_1} - B_s) + (B_{t_2} - B_s)$ is $\sigma(B_t - B_s, t \geq s)$-measurable and therefore independent of $\mathscr{F}_s$. Similarly, if $b > s$, the r.v.

$$Y = \int_s^b (B_u - B_s)\, du$$

for $t > s$ is independent of $\mathscr{F}_s$: writing the integral as the limit of its Riemann sums, it is immediate that $Y$ is $\sigma(B_t - B_s, t \geq s)$-measurable and therefore independent of $\mathscr{F}_s$.

*Remark 3.3* Computations concerning Brownian motion repeatedly require a certain set of formulas typical of Gaussian distributions. Let us recall them (they are all based on the relation $B_t \sim \sqrt{t}\,Z$ with $Z \sim N(0,1)$):

$$\text{a)} \quad \mathrm{E}[\mathrm{e}^{\theta B_t}] = \mathrm{e}^{\frac{1}{2}t\theta^2};$$

$$\text{b)} \quad \mathrm{E}[\mathrm{e}^{\theta B_t^2}] = \begin{cases} \frac{1}{\sqrt{1-2t\theta}} & \text{if } t\theta < \frac{1}{2} \\ +\infty & \text{if } t\theta \geq \frac{1}{2}\,. \end{cases}$$

a) (the Laplace transform) is computed in Exercise 1.6, b) in Exercise 1.12.

## 3.2 The law of a continuous process, Wiener measure

Let $\xi = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (\xi_t)_t, \mathrm{P})$ be a continuous $\mathbb{R}^d$-valued process and let $\psi_\xi : \omega \mapsto \{t \mapsto \xi_t(\omega)\}$ be the map mapping every $\omega$ to its associated path. This map takes its values in $\mathscr{C} = \mathscr{C}(\mathbb{R}^+, \mathbb{R}^d)$, the space of continuous maps $\mathbb{R}^+ \to \mathbb{R}^d$, endowed with the topology of uniform convergence on compact sets. Let us denote by $\mathscr{M}$ the corresponding Borel $\sigma$-algebra on $\mathscr{C}$.

**Proposition 3.3** $\psi_\xi : (\Omega, \mathscr{F}) \to (\mathscr{C}, \mathscr{M})$ is measurable.

The proof of Proposition 3.3 is rather straightforward and we shall skip it (see, however, Exercise 2.6).

Proposition 3.3 authorizes us to consider on the space $(\mathscr{C}, \mathscr{M})$ the image probability of P through $\psi_\xi$, called *the law of the process* $\xi$. The law of a continuous process is therefore a probability on the space $\mathscr{C}$ of continuous paths.

Let us denote by $\mathrm{P}^\xi$ the law of a process $\xi$. If we consider the coordinate r.v.'s $X_t : \mathscr{C} \to \mathbb{R}^d$ defined as $X_t(\gamma) = \gamma(t)$ (recall that $\gamma \in \mathscr{C}$ is a continuous function) and define $\mathscr{M}_t = \sigma(X_s, s \leq t)$, then

$$X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, \mathrm{P}^\xi)$$

is itself a stochastic process. By construction this new process has the same finite-dimensional distributions as $\xi$. Let $A_1, \dots, A_m \in \mathscr{B}(\mathbb{R}^d)$, then it is immediate that

$$\{\xi_{t_1} \in A_1, \dots, \xi_{t_m} \in A_m\} = \psi_\xi^{-1}\big(X_{t_1} \in A_1, \dots, X_{t_m} \in A_m\big)$$

so that

$$\mathrm{P}^\xi\big(X_{t_1} \in A_1, \dots, X_{t_m} \in A_m\big) = \mathrm{P}\big(\psi_\xi^{-1}\big(X_{t_1} \in A_1, \dots, X_{t_m} \in A_m\big)\big)$$
$$= \mathrm{P}\big(\xi_{t_1} \in A_1, \dots, \xi_{t_m} \in A_m\big)$$

and the two processes $\xi$ and $X$ have the same finite-dimensional distributions and are equivalent. This also implies that if $\xi$ and $\xi'$ are equivalent processes, then $\mathrm{P}^\xi = \mathrm{P}^{\xi'}$, i.e. they have the same law.

In particular, given two (continuous) Brownian motions, they have the same law. Let us denote this law by $\mathrm{P}^W$ (recall that this a probability on $\mathscr{C}$). $\mathrm{P}^W$ is the *Wiener measure* and the process $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, \mathrm{P}^W)$, having the same finite-dimensional distributions, is also a Brownian motion: it is the *canonical Brownian motion*.

## 3.3 Regularity of the paths

We have seen that a Brownian motion always admits a continuous version which is, moreover, $\gamma$-Hölder continuous for every $\gamma < \frac{1}{2}$. It is possible to provide a better description of the regularity of the paths, in particular showing that, in some sense, this estimate cannot be improved.

From now on $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ will denote a (continuous) Brownian motion. Let us recall that if $I \subset \mathbb{R}$ is an interval and $f : I \to \mathbb{R}$ is a continuous function, its *modulus of continuity* is the function

$$w(\delta) = \sup_{x,y\in I, |x-y|\le\delta} |f(x) - f(y)| \ .$$

The regularity of a function is determined by the behavior of its modulus of continuity as $\delta \to 0+$. In particular,

- $f$ is uniformly continuous if and only if $\lim_{\delta\to 0+} w(\delta) = 0$.
- $f$ is Hölder continuous with exponent $\alpha$ if and only if $w(\delta) \le c\delta^\alpha$ for some $c > 0$.
- $f$ is Lipschitz continuous if and only if $w(\delta) \le c\delta$ for some $c > 0$.

**Theorem 3.1 (P. Lévy's modulus of continuity)** Let $X$ be a real Brownian motion. For every $T > 0$

$$\mathrm{P}\Big(\lim_{\delta\to 0+} \frac{1}{\big(2\delta \log \frac{1}{\delta}\big)^{1/2}} \sup_{\substack{0\le s<t\le T\\ t-s\le\delta}} |X_t - X_s| = 1\Big) = 1 \ .$$

We skip the proof of Theorem 3.1, which is somewhat similar to the proof of the Iterated Logarithm Law, Theorem 3.2, that we shall see soon.

P. Lévy's theorem asserts that if $w(\cdot, \omega)$ is the modulus of continuity of $X_t(\omega)$ for $t \in [0, T]$, then P-a.s.

$$\lim_{\delta \to 0+} \frac{w(\delta, \omega)}{\left(2\delta \log \frac{1}{\delta}\right)^{1/2}} = 1 .$$

Note that this relation holds *for every* $\omega$ a.s. and does not depend on $T$.

As $\frac{w(\delta)}{\delta^{1/2}} \to +\infty$ as $\delta \to 0+$, Theorem 3.1 specifies that the paths of a Brownian motion cannot be Hölder continuous of exponent $\frac{1}{2}$ on the interval $[0, T]$ for every $T$ (Fig. 3.2). More precisely

**Corollary 3.1** Outside a set of probability 0 no path is Hölder continuous with exponent $\gamma \geq \frac{1}{2}$ in any time interval $I \subset \mathbb{R}^+$ having non-empty interior.

*Proof* It suffices to observe that for $q, r \in \mathbb{Q}^+$, $0 \leq q < r$,

$$\lim_{\delta \to 0+} \sup_{\substack{q \leq s < t \leq r \\ t-s \leq \delta}} \frac{|X_t - X_s|}{\left(2\delta \log \frac{1}{\delta}\right)^{1/2}} = 1 \qquad \text{a.s.} \tag{3.1}$$

thanks to Theorem 3.1 applied to the Brownian motion $(X_{t+q} - X_q)_t$. Therefore if $N_{q,r}$ is the negligible event on which (3.1) is not satisfied, $N = \bigcup_{q,r \in \mathbb{Q}^+} N_{q,r}$ is still negligible. Since an interval $I \subset \mathbb{R}^+$ having non-empty interior necessarily contains

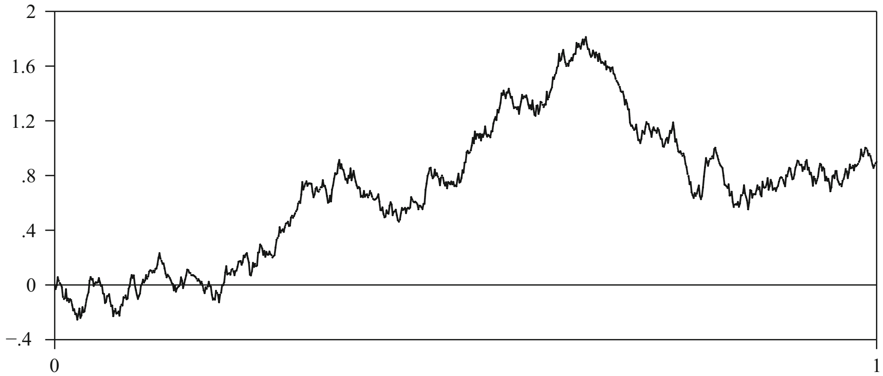


**Fig. 3.2** Example of the path of a real Brownian motion for $0 \leq t \leq 1$ (here the $x$ axis represents time). As in Fig. 3.1, the lack of regularity is evident as well as the typical oscillatory behavior, which will be better understood with the help of Theorem 3.2

an interval of the form $[q, r]$ with $q < r$, $q, r \in \mathbb{Q}^+$, no path outside $N$ can be Hölder continuous with exponent $\gamma \geq \frac{1}{2}$ in any time interval $I \subset \mathbb{R}^+$ having non-empty interior.

□

Let us recall that, given a function $f : \mathbb{R} \to \mathbb{R}$, its *variation* in the interval $[a, b]$ is the quantity

$$V_b^a f = \sup_\pi \sum_{i=1}^n \big| f(t_{i+1}) - f(t_i) \big| \, ,$$

the supremum being taken among all finite partitions $a = t_0 < t_1 < \cdots < t_{n+1} = b$ of the interval $[a, b]$. $f$ is said to have *finite variation* if $V_b^a f < +\infty$ for every $a$, $b \in \mathbb{R}$.

Note that a Lipschitz continuous function $f$ is certainly with finite variation: if we denote by $L$ the Lipschitz constant of $f$ then

$$\sum_{i=1}^n \big| f(t_{i+1}) - f(t_i) \big| \leq \sum_{i=1}^n L\, |t_{i+1} - t_i| = L \sum_{i=1}^n (t_{i+1} - t_i) = L(b - a) \, .$$

Therefore differentiable functions have finite variation on every bounded interval. It is also immediate that monotone functions have finite variation.

**Proposition 3.4** Let $X$ be a real Brownian motion. Let $\pi = \{t_0, \dots, t_m\}$ with $s = t_0 < t_1 \cdots < t_m = t$ be a partition of the interval $[s, t]$, $|\pi| = \max_{0 \leq k \leq m-1} |t_{k+1} - t_k|$. Then if

$$S_\pi = \sum_{k=0}^{m-1} |X_{t_{k+1}} - X_{t_k}|^2$$

we have

$$\lim_{|\pi| \to 0+} S_\pi = t - s \qquad \text{in } L^2 \, . \tag{3.2}$$

In particular, the paths of a Brownian motion do not have finite variation in any time interval a.s.

*Proof* We have $\sum_{k=0}^{m-1}(t_{k+1}-t_k) = (t_1-s)+(t_2-t_1)+\cdots+(t-t_{m-1}) = t-s$ so we can write

$$S_\pi - (t-s) = \sum_{k=0}^{m-1}\big[(X_{t_{k+1}}-X_{t_k})^2-(t_{k+1}-t_k)\big]\,.$$

We must prove that $\mathrm{E}[(S_\pi-(t-s))^2]\to 0$ as $|\pi|\to 0$. Note that $(X_{t_{k+1}}-X_{t_k})^2-(t_{k+1}-t_k)$ are independent (the increments of a Brownian motion over disjoint intervals are independent) and centered; therefore, if $h\neq k$, the expectation of the product

$$\big[(X_{t_{h+1}}-X_{t_h})^2-(t_{h+1}-t_h)\big]\big[(X_{t_{k+1}}-X_{t_k})^2-(t_{k+1}-t_k)\big]\Big)$$

vanishes so that

$$\begin{aligned}
&\mathrm{E}[(S_\pi-(t-s))^2]\\
&= \mathrm{E}\Big(\sum_{k=0}^{m-1}\big[(X_{t_{k+1}}-X_{t_k})^2-(t_{k+1}-t_k)\big]\times\sum_{h=0}^{m-1}\big[(X_{t_{h+1}}-X_{t_h})^2-(t_{h+1}-t_h)\big]\Big)\\
&= \sum_{k=0}^{m-1}\mathrm{E}\Big(\big[(X_{t_{k+1}}-X_{t_k})^2-(t_{k+1}-t_k)\big]^2\Big) = \sum_{k=0}^{m-1}(t_{k+1}-t_k)^2\mathrm{E}\Big[\Big(\frac{(X_{t_{k+1}}-X_{t_k})^2}{t_{k+1}-t_k}-1\Big)^2\Big]\,.
\end{aligned}$$

But for every $k$ the r.v. $\frac{X_{t_{k+1}}-X_{t_k}}{\sqrt{t_{k+1}-t_k}}$ is $N(0,1)$-distributed and the quantities

$$c = \mathrm{E}\Big[\Big(\frac{(X_{t_{k+1}}-X_{t_k})^2}{t_{k+1}-t_k}-1\Big)^2\Big]$$

are finite and *do not depend on* $k$ ($c=2$, if you really want to compute it...). Therefore, as $|\pi|\to 0$,

$$\mathrm{E}[(S_\pi-(t-s))^2] = c\sum_{k=0}^{m-1}(t_{k+1}-t_k)^2 \le c|\pi|\sum_{k=0}^{m-1}|t_{k+1}-t_k| = c|\pi|(t-s)\to 0\,,$$

which proves (3.2). Moreover,

$$S_\pi = \sum_{k=0}^{m-1}\big|X_{t_{k+1}}-X_{t_k}\big|^2 \le \max_{0\le i\le m-1}\big|X_{t_{i+1}}-X_{t_i}\big|\sum_{k=0}^{m-1}\big|X_{t_{k+1}}-X_{t_k}\big|\,. \tag{3.3}$$

As the paths are continuous, $\max_{0\le i\le m-1}|X_{t_{i+1}} - X_{t_i}| \to 0$ as $|\pi| \to 0+$ and therefore if the paths were with finite variation on $[s,t]$ for $\omega$ in some event $A$ of positive probability, then on $A$ we would have

$$\lim_{|\pi|\to 0+} \sum_{k=0}^{m-1} \big|X_{t_{k+1}} - X_{t_k}\big| < +\infty$$

and therefore, taking the limit in (3.3), we would have $\lim_{|\pi|\to 0+} S_\pi(\omega) = 0$ on $A$, in contradiction with the first part of the statement. □

Let us recall that if $f$ has finite variation, then it is possible to define the integral

$$\int_0^T \phi(t)\,df(t)$$

for every bounded Borel function $\phi$. Later we shall need to define an integral of the type

$$\int_0^T \phi(t)\,dX_t(\omega)\;,$$

which will be a key tool for the construction of new processes starting from Brownian motion. Proposition 3.4 states that this cannot be done $\omega$ by $\omega$, as the paths of a Brownian motion do not have finite variation. In order to perform this program we shall construct an *ad hoc* integral (the stochastic integral).

## 3.4 Asymptotics

We now present a classical result that gives very useful information concerning the behavior of the paths of a Brownian motion as $t \to 0+$ and as $t \to +\infty$.

**Theorem 3.2 (Iterated logarithm law)** Let $X$ be a Brownian motion. Then

$$\mathrm{P}\Big(\overline{\lim_{t\to 0+}}\, \frac{X_t}{\big(2t\log\log\frac{1}{t}\big)^{1/2}} = 1\Big) = 1\;. \tag{3.4}$$

Before giving the proof, let us point out some consequences.

**Corollary 3.2**

$$\underline{\lim}_{t\to 0+} \frac{X_t}{\big(2t\log\log\frac{1}{t}\big)^{1/2}} = -1 \qquad \text{a.s.} \tag{3.5}$$

$$\overline{\lim}_{s\to +\infty} \frac{X_s}{\big(2s\log\log s\big)^{1/2}} = 1 \qquad \text{a.s.} \tag{3.6}$$

$$\underline{\lim}_{s\to +\infty} \frac{X_s}{\big(2s\log\log s\big)^{1/2}} = -1 \qquad \text{a.s.} \tag{3.7}$$

*Proof* Let us prove (3.6). We know from Proposition 3.2 that $Z_t = tX_{1/t}$ is a Brownian motion. Theorem 3.2 applied to this Brownian motion gives

$$\overline{\lim}_{t\to 0+} \frac{tX_{1/t}}{\big(2t\log\log\frac{1}{t}\big)^{1/2}} = 1 \quad a.s.$$

Now, with the change of variable $s = \frac{1}{t}$,

$$\overline{\lim}_{t\to 0+} \frac{tX_{1/t}}{\big(2t\log\log\frac{1}{t}\big)^{1/2}} = \overline{\lim}_{t\to 0+} \sqrt{t}\,\frac{X_{1/t}}{\big(2\log\log\frac{1}{t}\big)^{1/2}} = \overline{\lim}_{s\to +\infty} \frac{X_s}{\big(2s\log\log s\big)^{1/2}}\,\cdot$$

Similarly, (3.5) and (3.7) follow from Theorem 3.2 applied to the Brownian motions $-X$, and $(-tX_{1/t})_t$. □

*Remark 3.4* (3.6) and (3.7) give important information concerning the asymptotic of the Brownian motion as $t \to +\infty$. Indeed they imply the existence of two sequences of times $(t_n)_n$, $(s_n)_n$, with $\lim_{n\to\infty} t_n = \lim_{n\to\infty} s_n = +\infty$ and such that

$$X_{t_n} \ge (1-\varepsilon)\sqrt{2t_n\log\log t_n}$$
$$X_{s_n} \le -(1-\varepsilon)\sqrt{2s_n\log\log s_n}\,.$$

This means that, as $t \to +\infty$, the Brownian motion takes arbitrarily large positive and negative values infinitely many times. It therefore exhibits larger and larger oscillations. As the paths are continuous, in particular, it visits every real number infinitely many times.

(3.6) and (3.7) also give a bound on how fast a Brownian motion moves away from the origin. In particular, (3.6) implies that, for $t$ large,

(continued)

*Remark 3.4* (continued)
$X_t(\omega) \le (1+\varepsilon)\sqrt{2t\log\log t}$. Similarly, by (3.7), for $t$ large, $X_t(\omega) \ge -(1+\varepsilon)\sqrt{2t\log\log t}$, so that, still for large $t$,

$$|X_t(\omega)| \le (1+\varepsilon)\sqrt{2t\log\log t}\ . \tag{3.8}$$

To be precise, there exists a $t_0 = t_0(\omega)$ such that (3.8) holds for every $t \ge t_0$.

Similarly, by (3.4) and (3.5), there exist two sequences $(t_n)_n$, $(s_n)_n$ decreasing to 0 and such that a.s. for every $n$,

$$X_{t_n} \ge (1-\varepsilon)\sqrt{2t_n\log\log\tfrac{1}{t_n}}$$
$$X_{s_n} \le -(1-\varepsilon)\sqrt{2s_n\log\log\tfrac{1}{s_n}}\ \cdot$$

In particular, $X_{s_n} < 0 < X_{t_n}$. By the intermediate value theorem the path $t \mapsto X_t$ crosses 0 infinitely many times in the time interval $[0,\varepsilon]$ for every $\varepsilon > 0$. This gives a hint concerning the oscillatory behavior of the Brownian motion.

In order to prove Theorem 3.2 we need some preliminary estimates.

**Lemma 3.1** Let $X$ be a continuous Brownian motion. If $x > 0$, $T > 0$, then

$$\mathrm{P}\Big(\sup_{0\le t\le T} X_t > x\Big) \le 2\mathrm{P}(X_T > x)\ .$$

*Proof* Let $t_0 < t_1 < \dots < t_n = T$, $I = \{t_0, \dots, t_n\}$ and let $\tau = \inf\{j; X_{t_j} > x\}$. Note that if $X_T(\omega) > x$, then $\tau \le T$, i.e. $\{X_T > x\} \subset \{\tau \le T\}$. Moreover, we have $X_{t_j} \ge x$ on $\{\tau = t_j\}$. Hence

$$\mathrm{P}(X_T > x) = \mathrm{P}(\tau \le T, X_T > x) = \sum_{j=0}^{n} \mathrm{P}(\tau = t_j, X_T > x) \ge \sum_{j=0}^{n} \mathrm{P}(\tau = t_j, X_T - X_{t_j} \ge 0)\ .$$

We have $\{\tau = t_j\} = \{X_{t_1} \le x, \dots, X_{t_{j-1}} \le x, X_{t_j} > x\}$. Hence $\{\tau = t_j\}$ is $\mathscr{F}_{t_j}$-measurable and independent of $\{X_T - X_{t_j} \ge 0\}$; moreover, $\mathrm{P}(X_T - X_{t_j} \ge 0) = \frac{1}{2}$, as $X_T - X_{t_j}$ is Gaussian and centered; therefore

$$\mathrm{P}(X_T > x) \ge \sum_{j=0}^{n} \mathrm{P}(\tau = t_j)\mathrm{P}(X_T - X_{t_j} \ge 0) = \frac{1}{2}\sum_{j=0}^{n} \mathrm{P}(\tau = t_j) = \frac{1}{2}\mathrm{P}\Big(\sup_{t\in I} X_t > x\Big)\ .$$

Let now $I_n$ be a sequence of finite subsets of $[0,T]$ increasing to $[0,T] \cap \mathbb{Q}$. By the previous inequality and taking the limit on the increasing sequence of events

$$\mathrm{P}\Big(\sup_{t\le T, t\in\mathbb{Q}} X_t > x\Big) = \sup_n \mathrm{P}\Big(\sup_{t\in I_n} X_t > x\Big) \le 2\mathrm{P}(X_T > x)\ .$$

As the paths are continuous, $\sup_{t\le T,t\in\mathbb{Q}} X_t = \sup_{t\le T} X_t$, and the statement is proved. □

**Lemma 3.2** If $x > 0$ then

$$\Big(x + \frac{1}{x}\Big)^{-1} \mathrm{e}^{-x^2/2} \le \int_x^{+\infty} \mathrm{e}^{-z^2/2}\, dz \le \frac{1}{x}\, \mathrm{e}^{-x^2/2}\ .$$

*Proof* We have

$$\int_x^{+\infty} \mathrm{e}^{-z^2/2}\, dz \le \frac{1}{x} \int_x^{+\infty} z\, \mathrm{e}^{-z^2/2}\, dz = \frac{1}{x}\, \mathrm{e}^{-x^2/2}$$

and the inequality on the right-hand side is proved. Moreover,

$$\frac{d}{dx}\frac{1}{x}\, \mathrm{e}^{-x^2/2} = -\Big(1 + \frac{1}{x^2}\Big)\, \mathrm{e}^{-x^2/2}$$

and therefore

$$\frac{1}{x}\, \mathrm{e}^{-x^2/2} = \int_x^{+\infty} \Big(1 + \frac{1}{z^2}\Big)\, \mathrm{e}^{-z^2/2}\, dz \le \Big(1 + \frac{1}{x^2}\Big) \int_x^{+\infty} \mathrm{e}^{-z^2/2}\, dz\ .$$

□

*Proof of Theorem 3.2* Let us prove first that

$$\overline{\lim_{t\to 0+}} \frac{X_t}{\big(2t \log\log \frac{1}{t}\big)^{1/2}} \le 1 \quad \text{a.s.} \tag{3.9}$$

Let $\phi(t) = \big(2t \log\log \frac{1}{t}\big)^{1/2}$. Let $(t_n)_n$ be a sequence decreasing to 0, let $\delta > 0$ and consider the event

$$A_n = \big\{X_t > (1+\delta)\phi(t) \text{ for some } t \in [t_{n+1}, t_n]\big\}\ .$$

$\delta$ being arbitrary, (3.9) follows if we can prove that $\mathrm{P}(\overline{\lim}_{n\to\infty} A_n) = 0$. Indeed, recall that $\overline{\lim}_{n\to\infty} A_n$ is the event that is formed by those $\omega$'s that belong to infinitely many $A_n$. If this set has probability 0, this means that $X_t > (1+\delta)\phi(t)$ for some $t \in [t_{n+1}, t_n]$ only for finitely many $n$ and therefore that $\overline{\lim}_{t\to 0+} \frac{X_t}{(2t\log\log\frac{1}{t})^{1/2}} \le (1+\delta)$, and $\delta$ being arbitrary, this implies (3.9).

By the Borel–Cantelli lemma it suffices to prove that the series $\sum_{n=1}^{\infty} \mathrm{P}(A_n)$ is convergent; we need, therefore, a good upper bound for $\mathrm{P}(A_n)$. First, as $\phi$ is increasing,

$$A_n \subset \Big\{ \sup_{0\le t\le t_n} X_t > (1+\delta)\phi(t_{n+1}) \Big\} ,$$

and by Lemmas 3.1 and 3.2, as $X_{t_n}/\sqrt{t_n} \sim N(0,1)$,

$$\begin{aligned}
\mathrm{P}(A_n) &= \mathrm{P}\Big( \sup_{0\le t\le t_n} X_t \ge (1+\delta)\phi(t_{n+1})\Big) \le 2\mathrm{P}\big(X_{t_n} \ge (1+\delta)\phi(t_{n+1})\big)\\
&= 2\mathrm{P}\Big(\frac{X_{t_n}}{\sqrt{t_n}} \ge (1+\delta)\Big(2\,\frac{t_{n+1}}{t_n}\log\log\frac{1}{t_{n+1}}\Big)^{1/2}\Big)\\
&= \frac{2}{\sqrt{2\pi}}\int_{x_n}^{+\infty} \mathrm{e}^{-z^2/2}\,dz \le \sqrt{\frac{2}{\pi}}\,\frac{1}{x_n}\,\mathrm{e}^{-x_n^2/2} ,
\end{aligned}$$

where $x_n = (1+\delta)\big(2\frac{t_{n+1}}{t_n}\log\log\frac{1}{t_{n+1}}\big)^{1/2}$. Let us choose now $t_n = q^n$ with $0 < q < 1$, but such that $\lambda = q(1+\delta)^2 > 1$. Now if we write $\alpha = \log\frac{1}{q} > 0$, then

$$x_n = (1+\delta)\big(2q\log\big[(n+1)\log\tfrac{1}{q}\big]\big)^{1/2} = \big[2\lambda\log(\alpha(n+1))\big]^{1/2} .$$

Therefore

$$\mathrm{P}(A_n) \le \sqrt{\frac{2}{\pi}}\,\frac{1}{x_n}\,\mathrm{e}^{-x_n^2/2} \le \sqrt{\frac{2}{\pi}}\,\mathrm{e}^{-\lambda\log(\alpha(n+1))} = \frac{c}{(n+1)^{\lambda}} \cdot$$

As $\lambda > 1$, the rightmost term is the general term of a convergent series, hence $\sum_{n=1}^{\infty} \mathrm{P}(A_n) < +\infty$ and, by the Borel–Cantelli Lemma, $\mathrm{P}(\overline{\lim}_{n\to\infty} A_n) = 0$, which completes the proof of (3.9).

Let us prove now the reverse inequality of (3.9). This will require the use of the converse part of the Borel–Cantelli Lemma, which holds under the additional assumption that the events involved are independent. For this reason, we shall first investigate the behavior of the increments of the Brownian motion. Let again $(t_n)_n$ be a sequence decreasing to 0 and let $Z_n = X_{t_n} - X_{t_{n+1}}$. The r.v.'s $Z_n$ are independent, being the increments of $X$. Then for every $x > 1$, $\varepsilon > 0$, we have

$$\begin{aligned}
&\mathrm{P}\big(Z_n > x\sqrt{t_n - t_{n+1}}\big) = \mathrm{P}\Big(\frac{X_{t_n} - X_{t_{n+1}}}{\sqrt{t_n - t_{n+1}}} > x\Big)\\
&= \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} \mathrm{e}^{-z^2/2}\,dz \ge \frac{x}{x^2+1}\,\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2} \ge \frac{1}{2x\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2} ,
\end{aligned} \tag{3.10}$$

where we have taken advantage of the left-hand side inequality of Lemma 3.2. Let $t_n = q^n$ with $0 < q < 1$ and put $\beta = \frac{2(1-\varepsilon)^2}{1-q}$, $\alpha = \log \frac{1}{q}$. Then

$$\begin{aligned} x = (1-\varepsilon)\frac{\phi(t_n)}{\sqrt{t_n - t_{n+1}}} &= \frac{1-\varepsilon}{\sqrt{1-q}}\sqrt{2\log\big(n\log\tfrac{1}{q}\big)} \\ &= \sqrt{\frac{2(1-\varepsilon)^2}{1-q}\log\big(n\log\tfrac{1}{q}\big)} = \sqrt{\beta\log(\alpha n)} \end{aligned}$$

so that (3.10) becomes

$$\mathrm{P}(Z_n > (1-\varepsilon)\phi(t_n)) \geq \frac{c}{n^{\beta/2}\sqrt{\log n}} \cdot$$

We can choose $q$ small enough so that $\beta < 2$ and the left-hand side becomes the general term of a divergent series. Moreover, as the r.v.'s $Z_n$ are independent, these events are independent themselves and by the Borel–Cantelli lemma we obtain

$$Z_n > (1-\varepsilon)\phi(t_n) \text{ infinitely many times a.s.}$$

On the other hand the upper bound (3.9), which has already been proved, applied to the Brownian motion $-X$ implies that a.s. we have eventually

$$X_{t_n} > -(1+\varepsilon)\phi(t_n)\ .$$

Putting these two relations together we have that a.s. for infinitely many indices $n$

$$\begin{aligned} X_{t_n} = Z_n + X_{t_{n+1}} &> (1-\varepsilon)\phi(t_n) - (1+\varepsilon)\phi(t_{n+1}) \\ &= \phi(t_n)\Big[1-\varepsilon-(1+\varepsilon)\frac{\phi(t_{n+1})}{\phi(t_n)}\Big]\ . \end{aligned} \tag{3.11}$$

Note that, as $\log\log\frac{1}{q^n} = \log n + \log\log\frac{1}{q}$ and $\lim_{n\to\infty}\frac{\log(n+1)}{\log n} = 1$,

$$\lim_{n\to\infty}\frac{\phi(t_{n+1})}{\phi(t_n)} = \lim_{n\to\infty}\frac{\sqrt{2q^{n+1}\log\log\frac{1}{q^{n+1}}}}{\sqrt{2q^n\log\log\frac{1}{q^n}}} = \sqrt{q}\lim_{n\to\infty}\frac{\sqrt{\log\log\frac{1}{q^{n+1}}}}{\sqrt{\log\log\frac{1}{q^n}}} = \sqrt{q}\ .$$

For every fixed $\delta > 0$ we can choose $\varepsilon > 0$ and $q > 0$ small enough so that

$$1-\varepsilon-(1+\varepsilon)\sqrt{q} > 1-\delta\ ,$$

hence from (3.11)

$$X_{t_n} > (1-\delta)\phi(t_n) \quad \text{for infinitely many times a.s.}$$

which completes the proof of the theorem. □

## 3.5 Stopping times

Sometimes, in the investigation of a stochastic process $X$, we shall consider the value of $X$ at some random time $\sigma$, i.e. of the r.v. $X_{\sigma(\omega)}(\omega)$. A random time is, in general, just a r.v. $\sigma : \Omega \to [0, +\infty[$. Among random times the stopping times of the next definition will play a particular role.

**Definition 3.3** Let $(\mathscr{F}_t)_{t\in T}$ be a filtration. A r.v. $\tau : \Omega \to T \cup \{+\infty\}$ is said to be a *stopping time* if, for every $t \in T$, $\{\tau \le t\} \in \mathscr{F}_t$. Associated to a stopping time $\tau$ let us define the $\sigma$-algebra

$$\mathscr{F}_\tau = \{A \in \mathscr{F}_\infty, A \cap \{\tau \le t\} \in \mathscr{F}_t \text{ for every } t \in T\}$$

where, as usual, $\mathscr{F}_\infty = \bigvee_t \mathscr{F}_t$.

Note that, in general, a stopping time is allowed to take the value $+\infty$. Intuitively the condition $\{\tau \le t\} \in \mathscr{F}_t$ means that at time $t$ we should be able to say whether $\tau \le t$ or not. For instance. we shall see that *the first time* at which a Brownian motion $B$ comes out of an open set $D$ is a stopping time. Intuitively, at time $t$ we know the values of $B_s$ for $s \le t$ and we are therefore able to say whether $B_s \notin D$ for some $s \le t$. Conversely, the last time of visit of $B$ to an open set $D$ is not a stopping time as in order to say whether some time $t$ is actually the last time of visit we also need to know the positions of $B_s$ for $s > t$.

$\mathscr{F}_\tau$ is, intuitively, the $\sigma$-algebra of the events for which at time $\tau$ we can say whether they are satisfied or not. The following proposition summarizes some elementary properties of stopping times. The proof is a straightforward application of the definitions and it is suggested to do it as an exercise (looking at the actual proof only later).

**Proposition 3.5** Let $\sigma$ and $\tau$ be stopping times. Then

a) $\tau$ is $\mathscr{F}_\tau$-measurable.
b) $\sigma \vee \tau$ and $\sigma \wedge \tau$ are stopping times.
c) If $\sigma \le \tau$, then $\mathscr{F}_\sigma \subset \mathscr{F}_\tau$.
d) $\mathscr{F}_{\sigma\wedge\tau} = \mathscr{F}_\sigma \cap \mathscr{F}_\tau$.

*Proof*

a) By Exercise 1.19, we just have to prove that, for every $s \ge 0$, $\{\tau \le s\} \in \mathscr{F}_\tau$. It is obvious that $\{\tau \le s\} \in \mathscr{F}_s \subset \mathscr{F}_\infty$. We then have to check that, for every $t$,

$\{\tau \le s\} \cap \{\tau \le t\} \in \mathscr{F}_t$. Now, if $t \le s$, we have $\{\tau \le s\} \cap \{\tau \le t\} = \{\tau \le t\} \in \mathscr{F}_t$. If, conversely, $t > s$, $\{\tau \le s\} \cap \{\tau \le t\} = \{\tau \le s\} \in \mathscr{F}_s \subset \mathscr{F}_t$.

b) We have

$$\{\sigma \wedge \tau \le t\} = \{\sigma \le t\} \cup \{\tau \le t\} \in \mathscr{F}_t \ .$$

and therefore $\sigma \wedge \tau$ is a stopping time. The argument for $\sigma \vee \tau$ is similar:

$$\{\sigma \vee \tau \le t\} = \{\sigma \le t\} \cap \{\tau \le t\} \in \mathscr{F}_t \ .$$

c) If $A \in \mathscr{F}_\sigma$ then, for every $t$, $A \cap \{\sigma \le t\} \in \mathscr{F}_t$. As $\{\tau \le t\} \subset \{\sigma \le t\}$,

$$A \cap \{\tau \le t\} = \underbrace{A \cap \{\sigma \le t\}}_{\in \mathscr{F}_t} \cap \underbrace{\{\tau \le t\}}_{\in \mathscr{F}_t} \in \mathscr{F}_t \ .$$

d) Thanks to c), $\mathscr{F}_{\sigma\wedge\tau}$ is contained both in $\mathscr{F}_\sigma$ and in $\mathscr{F}_\tau$. Let us prove the opposite inclusion. Let $A \in \mathscr{F}_\sigma \cap \mathscr{F}_\tau$; then $A \in \mathscr{F}_\infty$, $A \cap \{\tau \le t\} \in \mathscr{F}_t$ and $A \cap \{\sigma \le t\} \in \mathscr{F}_t$. We have therefore

$$A \cap \{\sigma \wedge \tau \le t\} = A \cap (\{\sigma \le t\} \cup \{\tau \le t\}) = (A \cap \{\sigma \le t\}) \cup (A \cap \{\tau \le t\}) \in \mathscr{F}_t$$

and therefore $A \in \mathscr{F}_{\sigma\wedge\tau}$. □

Note that, in particular, if $t \in \mathbb{R}^+$ then $\sigma \equiv t$ is a (deterministic) stopping time. Therefore if $\tau$ is a stopping time, by Proposition 3.5 b), $\tau \wedge t$ is also a stopping time. It is actually a bounded stopping time, even if $\tau$ is not. We shall use this fact very often when dealing with unbounded stopping times.

**Proposition 3.6** If $\tau$ is an a.s. finite stopping time and $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ is a progressively measurable process, then the r.v. $X_\tau$ is $\mathscr{F}_\tau$-measurable.

*Proof* Let us assume that $X$ takes its values in some measurable space $(E, \mathscr{E})$. We must prove that, for every $\Gamma \in \mathscr{E}$, $\{X_\tau \in \Gamma\} \in \mathscr{F}_\tau$. We know already (Example 2.3) that $X_\tau$ is $\mathscr{F}_\infty$-measurable, so that $\{X_\tau \in \Gamma\} \in \mathscr{F}_\infty$.

Recalling the definition of the $\sigma$-algebra $\mathscr{F}_\tau$, we must now prove that, for every $t$, $\{\tau \le t\} \cap \{X_\tau \in \Gamma\} \in \mathscr{F}_t$. Of course $\{\tau \le t\} \cap \{X_\tau \in \Gamma\} = \{\tau \le t\} \cap \{X_{\tau\wedge t} \in \Gamma\}$.

The r.v. $\tau \wedge t$ is $\mathscr{F}_t$-measurable: the event $\{\tau \wedge t \le s\}$ is equal to $\Omega$ if $s \ge t$ and to $\{\tau \le s\}$ if $s < t$ and belongs to $\mathscr{F}_t$ in both cases.

The r.v. $X_{\tau\wedge t}$ then turns out to be $\mathscr{F}_t$-measurable as the composition of the maps $\omega \mapsto (\omega, \tau\wedge t(\omega))$, which is measurable from $(\Omega, \mathscr{F}_t)$ to $(\Omega \times [0,t], \mathscr{F}_t \otimes \mathscr{B}([0,t]))$, and $(\omega, u) \mapsto X_u(\omega)$, from $(\Omega \times [0,t], \mathscr{F}_t \otimes \mathscr{B}([0,t]))$ to $(E, \mathscr{B}(E))$, which is

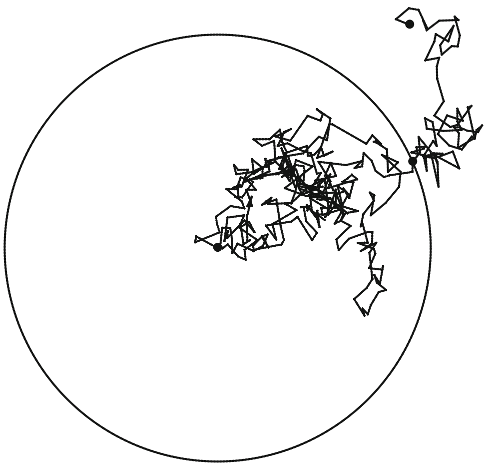

**Fig. 3.3** The exit time of a two-dimensional Brownian motion $B$ from the unit ball. The three black small circles denote the origin, the position at time 1 and the exit position $B_\tau$

measurable thanks to the assumption of progressive measurability. Hence $\{\tau \leq t\} \cap \{X_\tau \in \Gamma\} \in \mathscr{F}_t$ as the intersection of $\mathscr{F}_t$-measurable events.

□

If $A \in \mathscr{B}(E)$, let

$$\tau_A = \inf\{t \geq 0; X_t \notin A\} .$$

$\tau_A$ is called the *exit time* from $A$ (Fig. 3.3). In this definition, as well as in other similar situations, we shall always understand, unless otherwise indicated, that the infimum of the empty set is equal to $+\infty$. Therefore $\tau_A = +\infty$ if $X_t \in A$ for every $t \geq 0$. Similarly the r.v.

$$\rho_A = \inf\{t \geq 0; X_t \in A\}$$

is the *entrance time* in $A$. It is clear that it coincides with the exit time from $A^c$. Are exit times stopping times? Intuition suggests a positive answer, but we shall see that some assumptions are required.

**Proposition 3.7** Let $E$ be a metric space and $X$ an $E$-valued continuous process.

a) If $A$ is an open set, then $\tau_A$ is a stopping time.
b) If $F$ is a closed set and the filtration $(\mathscr{F}_t)_t$ is right continuous, then $\tau_F$ is a stopping time.

*Proof*

a) If $A$ is an open set, we have

$$\{\tau_A > t\} = \bigcup_{n=1}^{\infty} \bigcap_{\substack{r\in\mathbb{Q} \\ r<t}} \{d(X_r, A^c) > \tfrac{1}{n}\} \in \mathscr{F}_t \;.$$

Indeed, if $\omega$ belongs to the set on the right-hand side, then for some $n$ we have $d(X_r(\omega), A^c) > \frac{1}{n}$ for every $r \in \mathbb{Q} \cap [0, t[$ and, the paths being continuous, we have $d(X_s(\omega), A^c) \geq \frac{1}{n}$ for every $s \leq t$. Therefore $X_s(\omega) \in A$ for every $s \leq t$ and $\tau_A(\omega) > t$.

The opposite inclusion follows from the fact that if $\tau_A(\omega) > t$ then $X_s(\omega) \in A$ for every $s \leq t$ and hence $d(X_s(\omega), A^c) > \frac{1}{n}$ for some $n$ and for every $s \leq t$ (the image of $[0, t]$ through $s \mapsto X_s(\omega)$ is a compact subset of $E$ and its distance from the closed set $A^c$ is therefore strictly positive).

Therefore $\{\tau_A \leq t\} = \{\tau_A > t\}^c \in \mathscr{F}_t$ for every $t$.

b) Similarly, if $F$ is closed,

$$\{\tau_F \geq t\} = \bigcap_{r\in\mathbb{Q}, r<t} \{X_r \in F\} \in \mathscr{F}_t \;,$$

so that $\{\tau_F < t\} \in \mathscr{F}_t$. Therefore

$$\{\tau_F \leq t\} = \bigcap_{\varepsilon>0} \{\tau_F < t + \varepsilon\} \in \mathscr{F}_{t+} = \mathscr{F}_t$$

since we have assumed that the filtration is right-continuous.

□

## 3.6 The stopping theorem

Let $X$ be a Brownian motion. In Proposition 3.2 we have seen that, for every $s \geq 0$, $(X_{t+s} - X_s)_t$ is also a Brownian motion. Moreover, it is immediate that it is independent of $\mathscr{F}_s$. The following result states that these properties remain true if the deterministic time $s$ is replaced by a stopping time $\tau$.

**Theorem 3.3** Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion and $\tau$ an a.s. finite stopping time of the filtration $(\mathscr{F}_t)_t$. Then $Y_t = X_{t+\tau} - X_\tau$ is a Brownian motion independent of $\mathscr{F}_\tau$.

*Proof* Let us assume first that $\tau$ takes only a discrete set of values: $s_1 < s_2 < \dots < s_k < \dots$ Let $C \in \mathscr{F}_\tau$; then, recalling the definition of $\mathscr{F}_\tau$, we know that $C \cap \{\tau \le s_k\} \in \mathscr{F}_{s_k}$. Actually also $C \cap \{\tau = s_k\} \in \mathscr{F}_{s_k}$, as

$$C \cap \{\tau = s_k\} = (C \cap \{\tau \le s_k\}) \setminus \underbrace{(C \cap \{\tau \le s_{k-1}\})}_{\in \mathscr{F}_{s_{k-1}} \subset \mathscr{F}_{s_k}}$$

and both events on the right-hand side belong to $\mathscr{F}_{s_k}$. Then, if $A_1, \dots, A_n \in \mathscr{B}(\mathbb{R}^m)$ and $C \in \mathscr{F}_\tau$, we have

$$\begin{aligned}
&\mathrm{P}(Y_{t_1} \in A_1, \dots, Y_{t_n} \in A_n, C)\\
&= \sum_k \mathrm{P}(X_{t_1+\tau} - X_\tau \in A_1, \dots, X_{t_n+\tau} - X_\tau \in A_n, \tau = s_k, C)\\
&= \sum_k \mathrm{P}(\underbrace{X_{t_1+s_k} - X_{s_k} \in A_1, \dots, X_{t_n+s_k} - X_{s_k} \in A_n}_{\text{independent of } \mathscr{F}_{s_k}}, \underbrace{\tau = s_k, C}_{\mathscr{F}_{s_k}-\text{measurable}})\\
&= \sum_k \mathrm{P}(X_{t_1+s_k} - X_{s_k} \in A_1, \dots, X_{t_n+s_k} - X_{s_k} \in A_n)\mathrm{P}(\tau = s_k, C)\ .
\end{aligned}$$

Now recall that $(X_{t+s_k} - X_{s_k})_t$ is a Brownian motion for every $k$, so that the quantity $\mathrm{P}(X_{t_1+s_k} - X_{s_k} \in A_1, \dots, X_{t_n+s_k} - X_{s_k} \in A_n)$ does not depend on $k$ and is equal to $\mathrm{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n)$. Hence

$$\begin{aligned}
\mathrm{P}(Y_{t_1} \in A_1, \dots, Y_{t_n} \in A_n, C) &= \mathrm{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n) \sum_k \mathrm{P}(\tau = s_k, C)\\
&= \mathrm{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n)\mathrm{P}(C)\ .
\end{aligned}$$

Letting $C = \Omega$ we have that $Y$ is a Brownian motion (it has the same finite-dimensional distributions as $X$); letting $C \in \mathscr{F}_\tau$ we have that $Y$ is independent of $\mathscr{F}_\tau$.

In order to get rid of the assumption that $\tau$ takes a discrete set of values, we use the following result, which will also be useful later.

**Lemma 3.3** Let $\tau$ be a stopping time. Then there exists a non-increasing sequence $(\tau_n)_n$ of stopping times each taking a discrete set of values and such that $\tau_n \searrow \tau$. Moreover, $\{\tau_n = +\infty\} = \{\tau = +\infty\}$ for every $n$.

*Proof* Let

$$\tau_n(\omega) = \begin{cases} 0 & \text{if } \tau(\omega) = 0 \\ \frac{k+1}{2^n} & \text{if } \frac{k}{2^n} < \tau(\omega) \le \frac{k+1}{2^n} \\ +\infty & \text{if } \tau(\omega) = +\infty \, . \end{cases}$$

As $\tau_n - \frac{1}{2^n} \le \tau \le \tau_n$, obviously $\tau_n \searrow \tau$. Moreover,

$$\{\tau_n \le t\} = \bigcup_{k, \frac{k+1}{2^n} \le t} \left\{\tau_n = \tfrac{k+1}{2^n}\right\} = \{\tau(\omega) = 0\} \cup \bigcup_{k, \frac{k+1}{2^n} \le t} \left\{\tfrac{k}{2^n} < \tau \le \tfrac{k+1}{2^n}\right\} \in \mathscr{F}_t$$

and therefore $\tau_n$ is a stopping time. □

*End of the proof* of Theorem 3.3. If $\tau$ is a finite stopping time, let $(\tau_n)_n$ be a sequence of stopping times each taking a discrete set of values as in Lemma 3.3.

In the first part of the proof we have proved that, for every choice of $t_1, \dots, t_k$, the r.v.'s $X_{t_1+\tau_n} - X_{\tau_n}, \dots, X_{t_k+\tau_n} - X_{\tau_n}$ are independent of $\mathscr{F}_{\tau_n}$ and have the same distribution as $X_{t_1}, \dots, X_{t_k}$. Therefore for every bounded continuous function $\Phi : \mathbb{R}^m \to \mathbb{R}$ and $C \in \mathscr{F}_\tau \subset \mathscr{F}_{\tau_n}$ (recall that, as $\tau \le \tau_n$ then $\mathscr{F}_\tau \subset \mathscr{F}_{\tau_n}$, see Proposition 3.5 c)), we have

$$\begin{aligned} &\mathrm{E}\big[\Phi(X_{t_1+\tau_n} - X_{\tau_n}, \dots, X_{t_k+\tau_n} - X_{\tau_n}) 1_C\big] \\ &= \mathrm{E}\big[\Phi(X_{t_1+\tau_n} - X_{\tau_n}, \dots, X_{t_k+\tau_n} - X_{\tau_n})\big] \mathrm{P}(C) \\ &= \mathrm{E}\big[\Phi(X_{t_1}, \dots, X_{t_k})\big] \mathrm{P}(C) \, . \end{aligned} \tag{3.12}$$

Let us take the limit as $n \to \infty$ in (3.12). As the paths are continuous and by Lebesgue's theorem

$$\mathrm{E}\big[\Phi(X_{t_1+\tau_n} - X_{\tau_n}, \dots, X_{t_k+\tau_n} - X_{\tau_n}) 1_C\big] \underset{n\to\infty}{\to} \mathrm{E}\big[\Phi(X_{t_1+\tau} - X_{\tau}, \dots, X_{t_k+\tau} - X_{\tau}) 1_C\big]$$

so that

$$\mathrm{E}\big[\Phi(X_{t_1+\tau} - X_{\tau}, \dots, X_{t_k+\tau} - X_{\tau}) 1_C\big] = \mathrm{E}\big[\Phi(X_{t_1}, \dots, X_{t_k})\big] \mathrm{P}(C) \, . \tag{3.13}$$

Again by first choosing $C = \Omega$ we find that the joint distributions of $(X_{t_1+\tau} - X_\tau, \dots, X_{t_k+\tau} - X_\tau)$ and $(X_{t_1}, \dots, X_{t_k})$ coincide. Therefore the process $(X_{t+\tau} - X_\tau)_t$, having the same joint distributions of a Brownian motion, is a Brownian motion itself. Then taking $C \in \mathscr{F}_\tau$ we find that it is independent of $\mathscr{F}_\tau$. □

Let us describe a first, very useful, application of the stopping theorem. Let $a > 0$ and $\tau_a = \inf\{t; X_t \ge a\}$; $\tau_a$ is a stopping time, being the exit time from the half-line $]-\infty, a[$, and is called the *passage time at a*. By the Iterated Logarithm Law, Corollary 3.2, the Brownian motion takes arbitrarily large values a.s., hence $\tau_a$ is a.s. finite.

It is often of interest to determine the probability that a Brownian motion $X$ will cross the level $a$ before a time $t$. This will be a consequence of the following more general (and interesting) statement. Note that the events $\{\sup_{s\le t} X_s \ge a\}$ and $\{\tau_a \le t\}$ coincide.

**Theorem 3.4 (The joint distribution of a Brownian motion and of its running maximum)** Let $X$ be a Brownian motion. For every $t \ge 0$, let $X_t^* = \sup_{0\le s\le t} X_s$. Then, if $a \ge 0$ and $b \le a$,

$$\mathrm{P}(X_t^* \ge a, X_t \le b) = \mathrm{P}(X_t \ge 2a - b) . \tag{3.14}$$

*Proof* Let $\tau_a = \inf\{t \ge 0; X_t = a\}$. If we write $W_t = X_{t+\tau_a} - X_{\tau_a}$, we have

$$\mathrm{P}(X_t^* \ge a, X_t \le b) = \mathrm{P}(\tau_a \le t, X_t \le b) = \mathrm{P}(\tau_a \le t, W_{t-\tau_a} \le b - a) ,$$

as $W_{t-\tau_a} = X_t - X_{\tau_a} = X_t - a$. We know that $W$ is a Brownian motion independent of $\mathscr{F}_{\tau_a}$, hence of $\tau_a$. As $W$ has same law as $-W$, the pair $(\tau_a, W)$, as an $\mathbb{R}^+ \times \mathscr{C}([0,t],\mathbb{R})$-valued r.v., has the same law as $(\tau_a, -W)$.

Let $A = \{(s,\gamma) \in \mathbb{R} \times \mathscr{C}([0,t],\mathbb{R}), s \le t, \gamma(t-s) \le b-a\}$, so that we have $\mathrm{P}(\tau_a \le t, W_{t-\tau_a} \le b-a) = \mathrm{P}((\tau_a, W) \in A)$. Then, if $a \ge b$,

$$\begin{aligned}\mathrm{P}(X_t^* \ge a, X_t \le b) &= \mathrm{P}(\tau_a \le t, W_{t-\tau_a} \le b-a) = \mathrm{P}((\tau_a, W) \in A)\\ &= \mathrm{P}((\tau_a, -W) \in A) = \mathrm{P}(\tau_a \le t, -X_t + X_{\tau_a} \le b-a)\\ &= \mathrm{P}(\tau_a \le t, X_t \ge 2a-b) = \mathrm{P}(X_t \ge 2a-b)\end{aligned}$$

since if $X_t \ge 2a-b$ then *a fortiori* $X_t \ge a$, so that $\{X_t \ge 2a-b\} \subset \{\tau_a \le t\}$. □

Let us denote by $f$ the joint density of $(X_t, X_t^*)$. From (3.14), for $a > 0$, $b \le a$,

$$\begin{aligned}\int_{-\infty}^{b} dx \int_a^{+\infty} f(x,y)\, dy &= \mathrm{P}(X_t \le b, X^* \ge a) = \mathrm{P}(X_t \ge 2a-b)\\ &= \frac{1}{\sqrt{2\pi}} \int_{(2a-b)/\sqrt{t}}^{+\infty} \mathrm{e}^{-x^2/2}\, dx .\end{aligned}$$

Taking first the derivative with respect to $b$ we have

$$\int_a^{+\infty} f(b,y)\, dy = \frac{1}{\sqrt{2\pi t}}\, \mathrm{e}^{-\frac{1}{2t}(2a-b)^2}$$

and then taking the derivative with respect to $a$ we find

**Corollary 3.3** The joint density of $(X_t, X_t^*)$, i.e. of the running maximum and the position at time $t$ of a Brownian motion, is

$$f(b,a) = \frac{2}{\sqrt{2\pi t^3}}\,(2a-b)\,\mathrm{e}^{-\frac{1}{2t}\,(2a-b)^2}$$

for $a > 0, b \le a$ and $f(b,a) = 0$ otherwise.

Another consequence of Theorem 3.4 is the following, which is an important improvement of Lemma 3.1.

**Corollary 3.4 (The reflection principle)**

$$\mathrm{P}(X_t^* \ge a) = \mathrm{P}(\tau_a \le t) = 2\mathrm{P}(X_t \ge a)\ .$$

*Proof* We have, from (3.14), $\mathrm{P}(X_t^* \ge a, X_t \le a) = \mathrm{P}(X_t \ge a)$. Moreover, $\{X_t^* \ge a, X_t \ge a\} = \{X_t \ge a\}$. Hence

$$\mathrm{P}(X_t^* \ge a) = \mathrm{P}(X_t^* \ge a, X_t \le a) + \mathrm{P}(X_t^* \ge a, X_t \ge a) = 2\mathrm{P}(X_t \ge a)\ .$$

□

*Remark 3.5* Because of the symmetry of the centered Gaussian distributions we have $2\mathrm{P}(X_t \ge a) = \mathrm{P}(|X_t| \ge a)$. Let again $X_t^* = \sup_{0\le s\le t} X_s$. As a consequence of the reflection principle, for every $a > 0$,

$$\mathrm{P}(X_t^* \ge a) = \mathrm{P}(\tau_a \le t) = 2\mathrm{P}(X_t \ge a) = \mathrm{P}(|X_t| \ge a)\ .$$

This means that, for every $t$, the two r.v.'s $X_t^*$ and $|X_t|$ have the same distribution. Of course, the two processes are different: $(X_t^*)_t$ is increasing, whereas $(|X_t|)_t$ is not.

## 3.7 The simulation of Brownian motion

How to simulate the path of a Brownian motion? This is a very simple task, but it already enables us to investigate a number of interesting situations. Very often a simulation is the first step toward the understanding of complex phenomena.

Applications may require the computation of the expectation of a functional of the path of a stochastic process and if closed formulas are not available one must resort to simulation. This means that one must instruct a computer to simulate many paths and the corresponding value of the functional: by the law of large numbers the arithmetic mean of the values obtained in this way is an approximation of the desired expectation.

In this section we see the first aspects of the problem of simulating a stochastic process starting from the case of Brownian motion. For more information about simulation, also concerning more general processes, see Chap. 11. On this subject, because of today's ease of access to numerical computation, a vast literature is available (see, for example, Kloeden and Platen 1992; Fishman 1996; Glasserman 2004).

The first idea (others are possible) for the simulation of a Brownian motion is very simple: its increments being independent Gaussian r.v.'s, the problem is solved as soon as we are able to simulate a sequence of independent Gaussian r.v.'s, which can be done as explained in Sect. 1.8.

If $Z_1, Z_2, \dots$ are independent $m$-dimensional and $N(0, I)$-distributed r.v.'s on some probability space $(\Omega, \mathscr{F}, \mathrm{P})$, let us choose a grid of times $0 < h < 2h < \dots$ where $h > 0$ is a positive number (typically to be taken small). Then the r.v. $\sqrt{h}\, Z_1$ is $N(0, h\, I)$-distributed, i.e. has the same distribution as the increment of a Brownian motion over a time interval of size $h$. Hence the r.v.'s

$$\begin{aligned}
&\overline{B}_h(h) = \sqrt{h}\, Z_1 \\
&\overline{B}_h(2h) = \sqrt{h}\, (Z_1 + Z_2) \\
&\overline{B}_h(kh) = \sqrt{h}\, (Z_1 + \dots + Z_k) \\
&\dots
\end{aligned} \tag{3.15}$$

have the same joint distributions as the positions at times $h, 2h, \dots, kh, \dots$ of a Brownian motion.

If, for $kh \le t \le (k+1)h$, we define $\overline{B}_h(t)$ as a linear interpolation of the positions $\overline{B}_h(kh)$ and $\overline{B}_h((k+1)h)$, this is obviously an approximation of a Brownian motion. More precisely:

**Theorem 3.5** Let $T > 0$, $n > 0$ and $h = \frac{T}{n}$. Let us denote by $\overline{\mathrm{P}}_h$ the law of the process $\overline{B}_h$ ($\overline{\mathrm{P}}_h$ is therefore a probability on the canonical space $\mathscr{C} = \mathscr{C}([0, T], \mathbb{R}^m)$). Then $\overline{\mathrm{P}}_h$ converges weakly to the Wiener measure $\mathrm{P}^W$.

*Proof* Note first that the law $\overline{\mathrm{P}}_h$ neither depends on the choice of the r.v.'s $(Z_n)_n$, nor on the probability space on which they are defined, provided they are $N(0, I)$-distributed and independent.

Let $(\mathscr{C}, \mathscr{M}, \mathrm{P}^W)$ be the canonical space and, as usual, denote by $X_t$ the coordinate applications (recall that $\mathrm{P}^W$ denotes the Wiener measure, as defined p. 53).

The process $X = (\mathscr{C}, \mathscr{M}, (X_t)_t, \mathrm{P}^W)$ is therefore a Brownian motion. The r.v.'s

$$Z_k = \frac{1}{\sqrt{h}}\,(X_{(k+1)h} - X_{kh})$$

are therefore independent and $N(0, I)$-distributed. If we construct the process $\overline{B}_h$ starting from these r.v.'s, we have $\overline{B}_h(kh) = X_{kh}$. The process $\overline{B}_h$, in this case, is nothing else but the piecewise linear interpolation of $X$. As the paths are continuous (and therefore uniformly continuous on every compact interval), $\lim_{h\to 0} \overline{B}_h(\omega, t) = X_t(\omega)$ for every $\omega \in \mathscr{C}$ uniformly on compact sets. This means that the family of $\mathscr{C}$-valued r.v.'s $(\overline{B}_h)_h$ converges to $X$ a.s. in the topology of $\mathscr{C}$. Therefore, as a.s. convergence implies convergence in law, $\overline{B}_h \to X$ in law as $h \to 0$.

Hence, as remarked above, $\overline{\mathrm{P}}_h \to \mathrm{P}^W$ weakly as $h \to 0$, whatever the choice of the r.v.'s $(Z_n)_n$. □

Theorem 3.5 ensures that if $\phi : \mathscr{C} \to \mathbb{R}$ is a bounded continuous function, then

$$\lim_{h\to 0} \mathrm{E}[\phi(\overline{B}_h)] = \mathrm{E}[\phi(B)]\ .$$

*Example 3.2* Let $B$ be a real Brownian motion and let us consider the problem of computing numerically the quantity

$$\mathrm{E}\Big[\int_0^1 \mathrm{e}^{-B_s^2}\, ds\Big]\ . \tag{3.16}$$

It is easy to determine the exact value:

$$\int_0^1 \mathrm{E}[\mathrm{e}^{-B_s^2}]\, ds = \int_0^1 \frac{1}{\sqrt{1+2s}}\, ds = \sqrt{1+2s}\,\Big|_0^1 = \sqrt{3} - 1 \simeq 0.7320508\ .$$

In order to compute by simulation an approximation of this value, note that, using the same argument as in the proof of Theorem 3.5, the integral $\int_0^1 \mathrm{e}^{-B_s^2}\, ds$ is the limit of its Riemann sums, i.e.

$$\lim_{h\to 0} h \sum_{k=1}^n \mathrm{e}^{-\overline{B}_h(kh)^2} = \int_0^1 \mathrm{e}^{-B_s^2}\, ds$$

and, by Lebesgue's theorem,

$$\lim_{h\to 0} \mathrm{E}\Big[h \sum_{k=1}^n \mathrm{e}^{-\overline{B}_h(kh)^2}\Big] = \mathrm{E}\Big[\int_0^1 \mathrm{e}^{-B_s^2}]\, ds\Big]\ .$$

(continued)

*Example 3.2* (continued)
Hence, if $(Z^i_n)_n$ are independent $N(0,1)$-distributed r.v.'s then, by the Law of Large Numbers,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{n} \mathrm{e}^{-h(Z^i_1+\dots+Z^i_{kh})^2} = \mathrm{E}\Big[h \sum_{k=1}^{n} \mathrm{e}^{-\overline{B}_h(kh)^2}\Big] \tag{3.17}$$

so that the left-hand side above is close to the quantity (3.16) of interest.

The results for $N = 640{,}000$ and various values of the discretization parameter $h$ are given in Table 3.1. It is apparent that, even for relatively large values of $h$, the simulation gives quite accurate results.

**Table 3.1** The estimated values of Example 3.2

| $h$ | Value | Error |
|---|---|---|
| $\frac{1}{100}$ | 0.7295657 | 0.34% |
| $\frac{1}{200}$ | 0.7311367 | 0.12% |
| $\frac{1}{400}$ | 0.7315369 | 0.07% |
| $\frac{1}{800}$ | 0.7320501 | $10^{-5}$% |

Of course it would be very important to know how close to the true value $\mathrm{E}[\phi(B)]$ the approximation $\mathrm{E}[\phi(\overline{B}_h)]$ is. In other words, it would be very important to determine the speed of convergence, as $h \to 0$, of the estimator obtained by the simulated process. We shall address this question in a more general setting in Chap. 11.

Note also that in Example 3.2 we obtained an estimator using only the values of the simulated process at the discretization times $kh$, $k = 0, \dots, n$, and that it was not important how the simulated process was defined between the discretization times. This is almost always the case in applications.

*Example 3.3* Let $B$ be a real Brownian motion and let us consider the problem of computing numerically the quantity

$$\mathrm{P}\Big(\sup_{0\le s\le t} B_t \ge 1\Big)\,. \tag{3.18}$$

(continued)

*Example 3.3* (continued)
We could try to reproduce the simulation algorithm above, i.e. consider the approximation

$$\overline{X}_h := \frac{1}{N}\sum_{i=1}^{N} 1_{\{\sup_{0\le k\le n} \overline{B}^i_{kh}\ge 1\}} \;.$$

However, there is a difficulty: the functional $\mathscr{C} \ni \gamma \to 1_{\{\sup_{0\le s\le T}\gamma_s\ge 1\}}$ is not continuous because of the presence of the indicator function, so that Theorem 3.5 is not immediately sufficient in order to guarantee that $\lim_{h\to 0}\mathrm{E}[\overline{X}_h] = \mathrm{P}\big(\sup_{0\le t\le 1} B_s \ge 1\big)$.

One should, however, recall that the convergence holds if the set $A = \{\sup_{0\le s\le t} B_t \ge 1\}$ is such that $\mathrm{P}^W(\partial A) = 0$, as indicated on p. 11. This is the case here, as

$$\partial A = \Big\{\sup_{0\le s\le t} B_t = 1\Big\}$$

and the r.v. $B^* = \sup_{0\le s\le t} B_t$ has a density, which is an immediate consequence of Remark 3.5. Hence $\mathrm{P}^W(\partial A) = \mathrm{P}(B^* = 1) = 0$.

Table 3.2 reproduces the results of the computation of the probability (3.18) for different values of $h$.

Here again $N = 640{,}000$. It is apparent that the relative error is much larger than in Example 3.2 and, more noteworthy, that the error decreases very slowly as the discretization step $h$ becomes smaller. This is certainly related to the lack of regularity of the functional to be computed.

**Table 3.2** The outcomes of the numerical simulation and their relative errors. Thanks to the reflection principle, Corollary 3.4, the true value is $2\mathrm{P}(B_1 \ge 1) = 2(1 - \Phi(1)) = 0.3173$, $\Phi$ denoting the partition function of an $N(0,1)$-distributed r.v.

| $h$ | Value | Error |
|---|---|---|
| $\frac{1}{100}$ | 0.2899 | 8.64% |
| $\frac{1}{200}$ | 0.2975 | 6.23% |
| $\frac{1}{400}$ | 0.3031 | 4.45% |
| $\frac{1}{800}$ | 0.3070 | 3.22% |
| $\frac{1}{1600}$ | 0.3099 | 2.32% |

## Exercises

**3.1** (p. 455) Let $B$ be a Brownian motion and let $s \le t$.

a) Compute $\mathrm{E}[B_s B_t^2]$.
b) Compute $\mathrm{E}[B_s^2 B_t^2]$.
c) Show that

$$\mathrm{E}[B_s \mathrm{e}^{B_s}] = s\mathrm{e}^{s/2} .$$

d) Compute $\mathrm{E}[B_s \mathrm{e}^{B_t}]$.

Recall: $\mathrm{E}[X^4] = 3$ if $X \sim N(0, 1)$.

**3.2** (p. 455) Let $B$ be a Brownian motion. Compute

a) $\lim_{t\to+\infty} \mathrm{E}[1_{\{B_t \le a\}}]$.
b) $\lim_{t\to+\infty} \mathrm{E}[B_t 1_{\{B_t \le a\}}]$.

**3.3** (p. 456) Let $B$ be a Brownian motion. Compute

$$\lim_{t\to+\infty} \sqrt{t}\, \mathrm{E}[B_t^2 \mathrm{e}^{-B_t^2}] .$$

**3.4** (p. 456) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a $d$-dimensional Brownian motion. We prove here that the increments $(B_t - B_s)_{t\ge s}$ form a family that is independent of $\mathscr{F}_s$, as announced in Remark 3.2. This fact is almost obvious if $\mathscr{F}_s = \sigma(B_u, u \le s)$, i.e. if $B$ is a natural Brownian motion (why?). It requires some dexterity otherwise. Let $s > 0$.

a) If $s \le t_1 < \dots < t_m$ and $\Gamma_1, \dots, \Gamma_m \in \mathscr{B}(\mathbb{R}^d)$ prove that, for every $A \in \mathscr{F}_s$,

$$\begin{aligned} &\mathrm{P}\big(B_{t_m} - B_{t_{m-1}} \in \Gamma_m, \dots, B_{t_1} - B_s \in \Gamma_1, A\big) \\ &= \mathrm{P}\big(B_{t_m} - B_{t_{m-1}} \in \Gamma_m, \dots, B_{t_1} - B_s \in \Gamma_1\big) \cdot \mathrm{P}(A) . \end{aligned}$$

b) Show that the two $\sigma$-algebras

$$\sigma(B_{t_m} - B_s, \dots, B_{t_1} - B_s) \quad \text{and} \quad \sigma(B_{t_m} - B_{t_{m-1}}, B_{t_{m-1}} - B_{t_{m-2}}, \dots, B_{t_1} - B_s)$$

are equal.
c) Prove that the $\sigma$-algebra $\sigma(B_t - B_s, t \ge s)$ is independent of $\mathscr{F}_s$.

**3.5** (p. 457) Let $B$ be a $(\mathscr{F}_t)_t$-Brownian motion, let $\mathscr{G}$ a $\sigma$-algebra independent of $\mathscr{F}_\infty = \bigvee_t \mathscr{F}_t$ and $\widetilde{\mathscr{F}}_s = \mathscr{F}_s \vee \mathscr{G}$. The object of this exercise is to prove that $B$ remains a Brownian motion with respect to the larger filtration $(\widetilde{\mathscr{F}}_t)_t$.

a) Let $\mathscr{C}$ be the class of the events of the form $A \cap G$ with $A \in \mathscr{F}_s$ and $G \in \mathscr{G}$. Prove that $\mathscr{C}$ is stable with respect to finite intersections and that it generates the $\sigma$-algebra $\widetilde{\mathscr{F}}_s$.
b) Prove that $B$ is also a $(\widetilde{\mathscr{F}}_t)_t$-Brownian motion.

**3.6** (p. 457) Let $B$ be a Brownian motion.

a) Show that if $0 \le s \le t$, then the joint law of $(B_s, B_t)$ is

$$f_{s,t}(x,y) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{1}{2s}x^2} \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{1}{2(t-s)}(y-x)^2} \cdot \tag{3.19}$$

b) Show that, for every $s > 0$,

$$P(B_s < 0, B_{2s} > 0) = \frac{1}{8} \cdot$$

**3.7** (p. 459)

a) Let $B$ be a Brownian motion and let $X_t = e^{-t} B_{e^{2t}}$.
a1) Show that $(X_t)_{-\infty<t<+\infty}$ is a Gaussian process and compute its covariance function $K(s,t) = \mathrm{Cov}(X_s, X_t)$.
a2) Prove that $X$ is a stationary process, i.e. such that for every $t_1 < t_2 < \dots < t_m$ and $h > 0$, the r.v.'s $(X_{t_1}, \dots, X_{t_m})$ and $(X_{t_1+h}, \dots, X_{t_m+h})$ have the same distribution.
a3) Prove that the kernel $K(s,t) = e^{-|t-s|}$ is positive definite (i.e. that it satisfies (2.1)).
b) Let $(X_t)_{-\infty<t<+\infty}$ be a (not necessarily continuous) centered Gaussian process with covariance function $K(s,t) = e^{-|t-s|}$.
b1) Prove that $W_u = \sqrt{u}\, X_{\frac{1}{2}\log u}$ is a natural Brownian motion.
b2) Prove that $X$ has a continuous version.

**3.8** (p. 459) Sometimes in applications a process appears that is called a *$\rho$-correlated Brownian motion*. In dimension 2 it is a Gaussian process $(X_1(t), X_2(t))_t$ such that $X_1$ and $X_2$ are real Brownian motions, whereas $\mathrm{Cov}(X_1(t), X_2(s)) = \rho(s \wedge t)$, where $-1 \le \rho \le 1$ (Figs. 3.4 and 3.5).

a) Let $B = (B_1(t), B_2(t))$ be a two-dimensional Brownian motion. Then if $X_2 = B_2$ and $X_1(t) = \sqrt{1-\rho^2}\, B_1 + \rho B_2(t)$, then $X = (X_1(t), X_2(t))$ is a $\rho$-correlated Brownian motion.
b) Conversely, if $X$ is a $\rho$-correlated Brownian motion and $|\rho| < 1$, let $B_2(t) = X_2(t)$ and

$$B_1(t) = \frac{1}{\sqrt{1-\rho^2}} X_1(t) - \frac{\rho}{\sqrt{1-\rho^2}} X_2(t)\,.$$

Prove that $B = (B_1(t), B_2(t))$ is a two-dimensional Brownian motion such that $X_2(t) = B_2(t)$ and $X_1(t) = \sqrt{1-\rho^2}\, B_1 + \rho B_2(t)$.

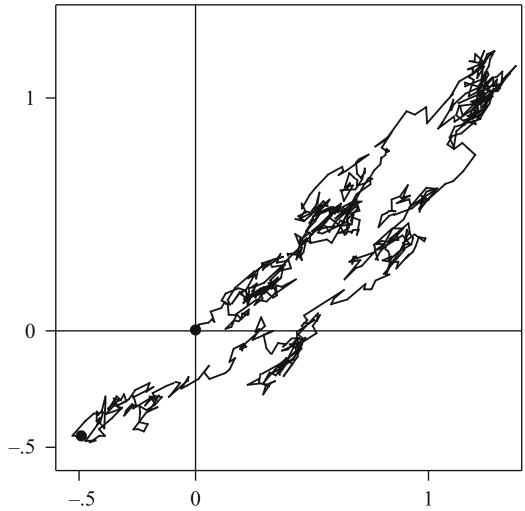


**Fig. 3.4** Example of a path of a two-dimensional $\rho$-correlated Brownian motion for $0 \le t \le 1$ and $\rho = 0.7$ (a black small circle denotes the origin and the position at time 1). Of course, for $\rho = 1$ the paths concentrate on the main diagonal

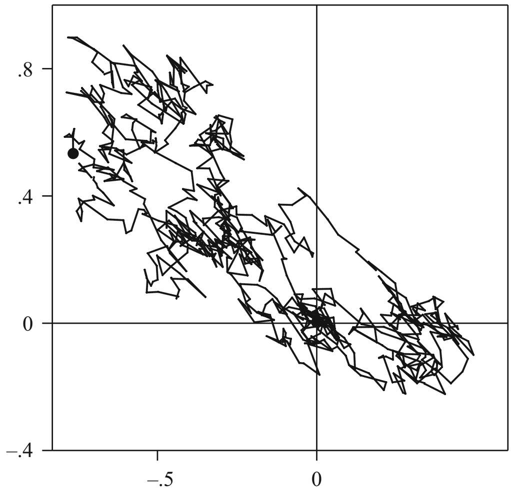


**Fig. 3.5** Example of a path of a two-dimensional $\rho$-correlated Brownian motion for $0 \le t \le 1$ and $\rho = -0.6$. Now the correlation is negative. (a black small circle denotes the origin and the position at time 1)

**3.9** (p. 460) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion. Determine the matrices $A \in M(m)$ such that $X_t = AB_t$ is also an $m$-dimensional Brownian motion.

**3.10** (p. 460) Let $B$ be an $m$-dimensional Brownian motion and $A \subset \mathbb{R}^m$ a Borel set having finite and strictly positive Lebesgue measure. Let

$$S_A(\omega) = \{t \in \mathbb{R}^+; B_t(\omega) \in A\} .$$

$S_A(\omega)$ is the set of the times $t$ such that $B_t(\omega) \in A$. Let us denote by $\lambda$ the Lebesgue measure of $\mathbb{R}$; therefore the quantity $\mathrm{E}[\lambda(S_A)]$ is the mean time spent by $B$ on the set $A$ ($S_A$ is called the *occupation time* of $A$). Prove that

$$\mathrm{E}[\lambda(S_A)] = \begin{cases} +\infty & \text{if } m \le 2 \\ \dfrac{1}{2\pi^{m/2}} \Gamma(\frac{m}{2} - 1) \displaystyle\int_A |x|^{2-m}\, dx & \text{if } m > 2 . \end{cases}$$

Recall the definition of the Gamma function: $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} \mathrm{e}^{-x}\, dx$, for $\alpha > 0$.

Note that $\mathrm{E}[\lambda(S_A)] = \int_0^{+\infty} \mathrm{P}(X_t \in A)\, dt$ by Fubini's theorem.

**3.11** (p. 461) Let $X$ be a continuous Gaussian process.

a) Let us denote by $G$ the family of the compactly supported finite measures $\gamma$ on $\mathbb{R}^+$ (i.e. such that $\gamma([a, b]^c) = 0$ for some $a, b \in \mathbb{R}^+$). Let

$$X_\gamma = \int X_s\, d\gamma(s) . \tag{3.20}$$

Show that $(X_\gamma)_{\gamma \in G}$ is a Gaussian family.

b) Let $\mu$ be a measure on $\mathbb{R}^+$ such that $\mu([a, b]) < +\infty$ for every finite interval $[a, b]$ (i.e. a *Borel measure*) and let

$$Y_t = \int_0^t X_s\, \mu(ds) .$$

b1) Prove that $Y$ is a Gaussian process.

b2) Let us assume that $X$ is a Brownian motion. Using the relation

$$\mu(]r, t])^2 = 2 \int_r^t \mu(]r, u])\, d\mu(u) \tag{3.21}$$

prove that $Y_t$ has centered Gaussian law with variance $\sigma_t^2 = \int_0^t \mu(]s, t])^2\, ds$. Compute $\mathrm{Cov}(Y_t, Y_s)$.

a) The integral is the limit of the Riemann sums $\sum_{i \ge 0} X_{i/n} \gamma(]i\frac{t}{n}, (i+1)\frac{t}{n}])$ and one can apply Proposition 1.9;

b2) use the relations (Fubini)

$$\mathrm{E}\Big(\int_0^t X_u\, d\mu(u) \cdot \int_0^s X_v\, d\mu(v)\Big) = \int_0^t d\mu(u) \int_0^s \mathrm{E}(X_u X_v)\, d\mu(v) ,$$

$$\int_0^s v\, d\mu(v) = \int_0^s d\mu(v) \int_0^v du = \int_0^s du \int_u^s d\mu(v) = \int_0^s \mu(]u, s])\, du .$$

**3.12** (p. 463) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and let

$$Z_t = B_t - \int_0^t \frac{B_u}{u}\, du\ .$$

a) Note that the integral converges for almost every $\omega$ and show that $(Z_t)_t$ is a Brownian motion with respect to its natural filtration $(\mathscr{G}_t)_t$.
b) Show that $(Z_t)_t$ is a process adapted to the filtration $(\mathscr{F}_t)_t$, but it is not a Brownian motion with respect to this filtration.
c) Prove that, for every $t$, $B_t$ is independent of $\mathscr{G}_t$.

a) Use in an appropriate way Exercise 3.11.

**3.13** (p. 464) Let $B$ be a Brownian motion. Prove that, for every $a > 0$ and $T > 0$,

$$\mathrm{P}(B_t \le a\sqrt{t} \text{ for every } t \le T) = 0\ .$$

**3.14** (p. 465) Let $B$ be a Brownian motion and $b \in \mathbb{R}$, $\sigma > 0$. Let $X_t = \mathrm{e}^{bt+\sigma B_t}$.

a) Investigate the existence and finiteness of the a.s. limit

$$\lim_{t\to+\infty} X_t$$

according to the possible values of $b, \sigma$.
b) Investigate the existence and finiteness of

$$\lim_{t\to+\infty} \mathrm{E}[X_t] \tag{3.22}$$

according to the possible values of $b, \sigma$.

**3.15** (p. 465) Let $B$ be a Brownian motion and $b \in \mathbb{R}$, $\sigma > 0$.

a) For which values of $b, \sigma$, $b \neq 0$, is the integral

$$\int_0^{+\infty} \mathrm{e}^{bu+\sigma B_u}\, du$$

a.s. finite?
b1) Prove that the r.v.

$$\int_0^1 1_{\{B_u>0\}}\, du$$

is $> 0$ a.s.

b2) Prove that

$$\lim_{t\to+\infty}\int_0^t 1_{\{B_u>0\}}\,du = +\infty \qquad \text{a.s.}$$

b3) Deduce that

$$\int_0^{+\infty} e^{\sigma B_u}\,du = +\infty \qquad \text{a.s.}$$

c) For which values of $b, \sigma$ is

$$\mathrm{E}\Big[\int_0^{+\infty} e^{bu+\sigma B_u}\,du\Big] < +\infty\,?$$

**3.16** (p. 466) (Approximating exit times) Let $E$ be a metric space, $D \subset E$ an open set, $X = (X_t)_t$ a *continuous* $E$-valued process, $\tau$ the exit time of $X$ from $D$. Let $(D_n)_n$ be an increasing sequence of open sets with $D_n \subset D$ for every $n > 0$, and such that $d(\partial D_n, \partial D) \le \frac{1}{n}$. Let us denote by $\tau_n$ the exit time of $X$ from $D_n$. Then, as $n \to \infty$, $\tau_n \to \tau$ and, on $\{\tau < +\infty\}$, $X_{\tau_n} \to X_\tau$.

**3.17** (p. 467) Let $B$ be an $m$-dimensional Brownian motion and $D$ an open set containing the origin. For $\rho > 0$ let us denote by $D_\rho$ the set $\rho D$ homothetic to $D$ and by $\tau$ and $\tau_\rho$ the exit times of $B$ from $D$ and $D_\rho$ respectively.

Show that the r.v.'s $\tau_\rho$ and $\rho^2\,\tau$ have the same law. In particular, $\mathrm{E}[\tau_\rho] = \rho^2\mathrm{E}[\tau]$, this relation being true whether the quantities appearing on the two sides of the equation are finite or not.

**3.18** (p. 467) Let $S$ be the unit ball centered at the origin of $\mathbb{R}^m$, $m \ge 1$, and $X$ a continuous $m$-dimensional Brownian motion. Let $\tau = \inf\{t; X_t \notin S\}$ be the exit time of $X$ from $S$.

a) Prove that $\tau < +\infty$ a.s.; $X_\tau$ is therefore a r.v. with values in $\partial S$. Show that the law of $X_\tau$ is the $(m-1)$-dimensional Lebesgue measure of $\partial S$, normalized so that it has total mass 1.
b) Prove that $\tau$ and $X_\tau$ are independent.

Recall that an orthogonal matrix transforms a Brownian motion into a Brownian motion (Exercise 3.9 b)) and use the fact that the only measures on $\partial S$ that are invariant under the action of orthogonal matrices are of the form $c \cdot \lambda$, where $\lambda$ is the $(m-1)$-dimensional Lebesgue measure of $\partial S$ and $c$ some positive constant.

**3.19** (p. 468)

a) Let $f : [0,1] \to \mathbb{R}$ be a continuous function. Prove that

$$\log \|e^f\|_\infty = \sup_{0\le s\le 1} f(s)\ .$$

b) Let $B$ be a Brownian motion.

b1) Prove that, for every $t > 0$,

$$\int_0^t e^{B_s}\,ds \ \overset{\mathscr{L}}{\sim}\ t\int_0^1 e^{\sqrt{t}B_s}\,ds\ .$$

b2) Prove that

$$\frac{1}{\sqrt{t}}\log\int_0^t e^{B_s}\,ds \ \underset{t\to+\infty}{\overset{\mathscr{L}}{\to}}\ \sup_{s\le 1} B_s\ .$$

b3) Give an approximation, for $t$ large, of

$$\mathrm{P}\Big(\int_0^t e^{B_s}\,ds \le 1\Big) \qquad \mathrm{P}\Big(\int_0^t e^{B_s}\,ds \le e^{.77337\cdot\sqrt{t}}\Big)$$

(0.77337 is, approximately, the quantile of order $\frac{3}{4}$ of the $N(0,1)$ distribution).

Recall the relation among the $L^p$ norms: $\lim_{p\to+\infty}\|f\|_p = \|f\|_\infty$.

**3.20** (p. 469)

a) Let $X$ be a continuous Brownian motion, $a > 0$ and $\tau_a = \inf\{t; X_t \ge a\}$ the passage time of $X$ at $a$. Prove that the law of $\tau_a$ has a density, with respect to Lebesgue measure, given by $f_a(t) = 0$ for $t \le 0$ and (Fig. 3.6)

$$f_a(t) = \frac{a}{(2\pi)^{1/2}\,t^{3/2}}\,e^{-a^2/2t} \tag{3.23}$$

for $t > 0$. In particular, $\tau_a$ does not have finite expectation (and neither does $\sqrt{\tau_a}$). Show that, as $t \to +\infty$,

$$\mathrm{P}(\tau_a > t) \sim \frac{2a}{\sqrt{2\pi}}\,\frac{1}{\sqrt{t}}\ \cdot$$

b) A researcher wants to explore the distribution of $\tau_1$ by simulation. He therefore realizes a computer program and simulates $N$ paths of a Brownian motion and records the passage times at $a = 1$ of each of them, $T_1, \dots, T_N$ say.

Assume $N = 10{,}000$. What is the probability that $T_j \ge 10^8$ for at least one path? And the probability that $T_j \ge 10^{10}$ (1 hundred millions) for at least one path? Why is the program taking so long? This should be done numerically with a computer.

c) Let us recall that a probability $\mu$ is said to be *stable* with exponent $\alpha$, $0 < \alpha \le 2$, if

$$\frac{X_1 + \cdots + X_n}{n^{1/\alpha}} \sim X_1$$

where $X_1, \dots, X_n$ are independent $\mu$-distributed r.v.'s (for instance, a centered normal law is stable with exponent 2).

Prove that the law of $\tau_a$ is stable with exponent $\alpha = \frac{1}{2}$.

a) Note that $\mathrm{P}(\tau_a \le t) = \mathrm{P}(\sup_{0\le s\le t} B_s > a)$. c) Use Theorem 3.3 to prove that if $X_1, \dots, X_n$ are independent copies of $\tau_1$, then $X_1 + \cdots + X_n$ has the same distribution as $\tau_{na}$. This property of stability of the law of $\tau_a$ will also be proved in Exercise 5.30 in quite a different manner. Do not try to compute the law of $X_1 + \cdots + X_n$ by convolution!

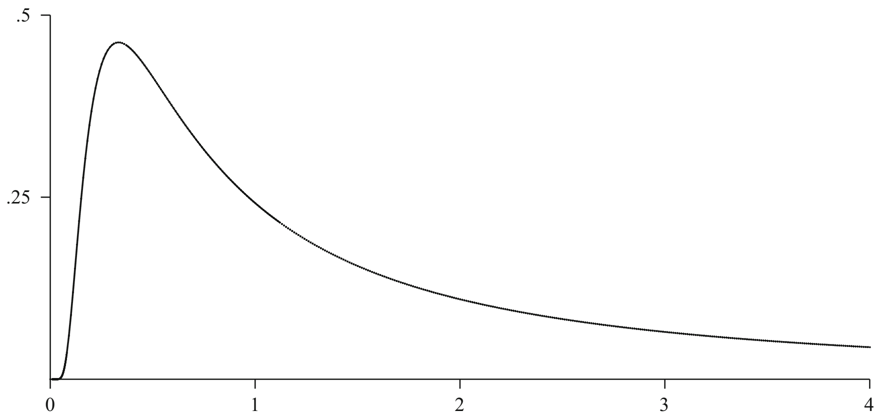


**Fig. 3.6** The graph of the density $f$ of the passage time at $a = 1$. Note that, as $t \to 0+$, $f$ tends to 0 very fast, whereas its decrease for $t \to +\infty$ is much slower, which is also immediate from (3.23)

**3.21** (p. 470)

a) Prove, without explicitly computing the values, that

$$\mathrm{E}\Big[\sup_{0\le s\le t} B_s\Big] = \sqrt{t}\,\mathrm{E}\Big[\sup_{s\le 1} B_s\Big]\,. \tag{3.24}$$

b) Compute

$$\mathrm{E}\Big[\sup_{0\le s\le t} B_s\Big]$$

(recall Exercise 1.3).

**3.22** (p. 471)

a) Let $B$ be an $m$-dimensional Brownian motion and let, for $\mathbb{R}^m \ni z \neq 0$, $k > 0$, $H = \{x \in \mathbb{R}^m, \langle z, x\rangle \leq k\}$ be a hyperplane containing the origin in its interior. Let us denote by $\tau$ the exit time of $B$ from $H$. The object of this exercise is to compute the law of $\tau$ and its expectation.

a1) Prove that

$$X_t = \frac{1}{|z|}\,\langle z, B_t\rangle$$

is a Brownian motion and that $\tau$ coincides with the passage time of $X$ at $a = \frac{k}{|z|}$.

a2) Compute the density of $\tau$ and $\mathrm{E}[\tau]$. Is this expectation finite?

b) Let $B = (B_1, B_2)$ be a two-dimensional $\rho$-correlated Brownian motion (see Exercise 3.8) and consider the half-space $H = \{(x, y), x + y < 1\}$.

b1) Find a number $\alpha > 0$ such that

$$X_t = \alpha(B_1(t) + B_2(t))$$

is a natural Brownian motion.

b2) Let $\tau$ be the exit time of $B$ from $H$. Compute the density of $\tau$ and $\mathrm{E}[\tau]$. Is this expectation finite? Compute $\mathrm{P}(\tau \leq 1)$ and determine for which values of $\rho$ this quantity is maximum.

**3.23** (p. 472) (Do Exercise 3.22 first)

a) Let $B$ be a two-dimensional Brownian motion and $z \in \mathbb{R}^2, z \neq 0$. Let $X_t = \langle B_t, z\rangle$. Show that there exists a number $v > 0$ such that $W_t = vX_t$ is a Brownian motion.

b) Let $\tau$ be the first time at which $B$ reaches the straight line $x + y = 1$.

b1) Let $z = (1, 1)$. Show that $\tau$ coincides with the passage time at $a = 1$ of $X_t = \langle B_t, z\rangle$.

b2) Show that the law of $\tau$ has a density and compute it. Compute $\mathrm{P}(\tau \leq 1)$.

c) Let $\tau_1$ and $\tau_2$ be the first times at which $B$ reaches the lines $2x + y = 1$ and $\frac{1}{2}x - y = 1$, respectively.

c1) Compute the densities of $\tau_1$ and $\tau_2$.

c2) Prove that $\tau_1$ and $\tau_2$ are independent.

c3) Let $\sigma = \tau_1 \wedge \tau_2$: $\sigma$ is the exit time of $B$ out of the shaded infinite region in Fig. 3.7. With the help of numerical tables compute $\mathrm{P}(\sigma \leq 1)$.

c4) Compute $\mathrm{P}(\tau_1 \leq \tau_2)$.

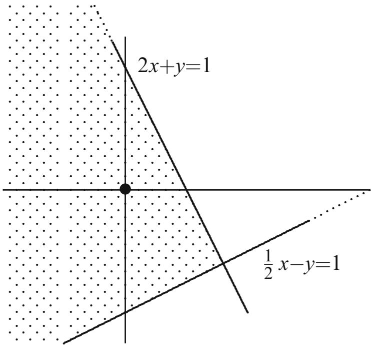


**Fig. 3.7** $\sigma = \tau_1 \wedge \tau_2$ is the exit time out of the shaded area

# Chapter 4
# Conditional Probability

## 4.1 Conditioning

Let $(\Omega, \mathscr{F}, \mathrm{P})$ be a probability space and $A \in \mathscr{F}$ an event having strictly positive probability. Recall that the conditional probability of P with respect to $A$ is the probability $\mathrm{P}_A$ on $(\Omega, \mathscr{F})$, which is defined as

$$\mathrm{P}_A(B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)} \quad \text{for every } B \in \mathscr{F} \ . \tag{4.1}$$

Intuitively the situation is the following: initially we know that every event $B \in \mathscr{F}$ can appear with probability $\mathrm{P}(B)$. If, later, we acquire the information that the event $A$ has occurred or will certainly occur, we replace the law P with $\mathrm{P}_A$, in order to keep into account the new information.

A similar situation is the following. Let $X$ be a real r.v. and $Z$ another r.v. taking values in a countable set $E$ and such that $\mathrm{P}(Z = z) > 0$ for every $z \in E$. For every Borel set $A \subset \mathbb{R}$ and for every $z \in E$ let

$$n(z, A) = \mathrm{P}(X \in A | Z = z) = \frac{\mathrm{P}(X \in A, Z = z)}{\mathrm{P}(Z = z)} \cdot$$

For every $z \in E$, $A \mapsto n(z, A)$ is a probability on $\mathbb{R}$: it is the *conditional law of $X$ given $Z = z$*. This probability has an intuitive meaning not dissimilar from the one just pointed out above: $A \mapsto n(z, A)$ is the law that it is reasonable to consider for the r.v. $X$, should we acquire the information that the event $\{Z = z\}$ certainly occurs. The *conditional expectation of $X$ given $Z = z$* is defined as the mean of this law, if it exists:

$$\mathrm{E}[X | Z = z] = \int x\, n(z, dx) = \frac{1}{\mathrm{P}(Z = z)} \int_{\{Z=z\}} X\, d\mathrm{P} = \frac{\mathrm{E}[X 1_{\{Z=z\}}]}{\mathrm{P}(Z = z)} \cdot$$

© Springer International Publishing AG 2017 
P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_4

This is a very important concept, as we shall see constantly throughout. For this reason we need to extend it to the case of a general r.v. $Z$ (i.e. without the assumption that it takes at most only countably many values). This is the object of this chapter.

Let us see first how it is possible to characterize the function $h(z) = \mathrm{E}[X|Z = z]$ in a way that continues to be meaningful in general. For every $B \subset E$ we have

$$\int_{\{Z\in B\}} h(Z)\, d\mathrm{P} = \sum_{z\in B} \mathrm{E}[X|Z = z]\mathrm{P}(Z = z) = \sum_{z\in B} \mathrm{E}[X1_{\{Z=z\}}]$$
$$= \mathrm{E}[X1_{\{Z\in B\}}] = \int_{\{Z\in B\}} X\, d\mathrm{P} .$$

Therefore the r.v. $h(Z)$, which is of course $\sigma(Z)$-measurable, is such that its integral on any event $B$ of $\sigma(Z)$ coincides with the integral of $X$ on the same $B$. We shall see that this property characterizes the conditional expectation.

In the following sections we use this property in order to extend the notion of conditional expectation to a more general (and interesting) situation. We shall come back to conditional distributions at the end.

## 4.2 Conditional expectations

Let $X$ be a real r.v. and $X = X^+ - X^-$ its decomposition into positive and negative parts and assume $X$ to be *lower semi-integrable* (l.s.i.) i.e. such that $\mathrm{E}[X^-] < +\infty$. See p. 2 for more explanations.

**Definition and Theorem 4.1** Let $X$ be a real l.s.i. r.v. and $\mathscr{D} \subset \mathscr{F}$ a sub-$\sigma$-algebra. The conditional expectation of $X$ with respect to $\mathscr{D}$, denoted $\mathrm{E}[X|\mathscr{D}]$, is the (equivalence class of) r.v.'s $Z$ which are $\mathscr{D}$-measurable and l.s.i. and such that for every $D \in \mathscr{D}$

$$\int_D Z\, d\mathrm{P} = \int_D X\, d\mathrm{P} . \tag{4.2}$$

A r.v. $Z$ with these properties always exists and is unique up to P-equivalence.

*Proof* Existence. Let us assume first that $X \ge 0$; then let us consider on $(\Omega, \mathscr{D})$ the positive measure

$$\mathrm{Q}(B) = \int_B X\, d\mathrm{P} \qquad B \in \mathscr{D} .$$

Obviously $\mathrm{Q} \ll \mathrm{P}$ (see p. 4); thus by the Radon–Nikodym theorem there exists a real $\mathscr{D}$-measurable r.v. $Z$ such that

$$\mathrm{Q}(B) = \int_B Z\, d\mathrm{P}$$

and hence $Z \in \mathrm{E}[X|\mathscr{D}]$. Note that if $X$ is integrable then Q is finite and $Z$ is integrable. For a general r.v. $X$ just decompose $X = X^+ - X^-$ and check that $\mathrm{E}[X|\mathscr{D}] = \mathrm{E}[X^+|\mathscr{D}] - \mathrm{E}[X^-|\mathscr{D}]$ satisfies the conditions of Definition 4.1. This is well defined since, $X^-$ being integrable, $\mathrm{E}[X^-|\mathscr{D}]$ is integrable itself and a.s. finite (so that the form $+\infty - \infty$ cannot appear).

Uniqueness. Let us assume first that $X$ is integrable. If $Z_1, Z_2$ are $\mathscr{D}$-measurable and satisfy (4.2) then the event $B = \{Z_1 > Z_2\}$ belongs to $\mathscr{D}$ and

$$\int_B (Z_1 - Z_2)\, d\mathrm{P} = \int_B X\, d\mathrm{P} - \int_B X\, d\mathrm{P} = 0\ .$$

As the r.v. $Z_1 - Z_2$ is strictly positive on $B$ this implies $\mathrm{P}(B) = 0$ and hence $Z_2 \geq Z_1$ a.s. By symmetry, $Z_1 = Z_2$ a.s.

In general if $X$ is l.s.i. and not necessarily integrable this argument cannot be applied as is, since it is possible that $\int_B X\, d\mathrm{P} = +\infty$. However, $B$ can be approximated by events $B_n$ on which the integral of $X$ is finite; the details are left as an exercise. □

*Remark 4.1* (4.2) states that if $Z = \mathrm{E}[X|\mathscr{D}]$, the relation

$$\mathrm{E}[ZW] = \mathrm{E}[XW] \tag{4.3}$$

holds for every r.v. $W$ of the form $W = 1_A$, $A \in \mathscr{D}$. Of course, then, (4.3) also holds for every linear combination of these indicator functions and, by standard approximation results (see Proposition 1.11), for every positive bounded $\mathscr{D}$-measurable r.v. $W$.

If $Y$ is a r.v. taking values in some measurable space $(E, \mathscr{E})$, sometimes we shall write $\mathrm{E}[X|Y]$ instead of $\mathrm{E}[X|\sigma(Y)]$. We know that every real $\sigma(Y)$-measurable r.v. is of the form $g(Y)$, where $g : E \to \mathbb{R}$ is a measurable function (Doob's criterion, Lemma 1.1). Therefore there exists a measurable function $g : E \to \mathbb{R}$ such that $\mathrm{E}[X|Y] = g(Y)$ a.s. Sometimes we shall use for the function $g$, in an evocative manner, the notation

$$g(y) = \mathrm{E}[X|Y = y]\ .$$

Always keeping in mind that every $\sigma(Y)$-measurable r.v. $W$ is of the form $\psi(Y)$ for a suitable measurable function $\psi$, such a $g$ must satisfy the relation

$$\mathrm{E}[X\psi(Y)] = \mathrm{E}[g(Y)\psi(Y)] \tag{4.4}$$

for every bounded measurable function $\psi$. The next section provides some tools for the computation of $g(y) = \mathrm{E}[X\,|\,Y = y]$.

To be precise (or pedantic…) a conditional expectation is an equivalence class of r.v.'s but, in order to simplify the arguments, we shall often identify the equivalence class $\mathrm{E}[X\,|\,\mathscr{D}]$ with one of its elements $Z$.

*Remark 4.2* We shall often be called to verify statements of the form "a certain r.v. $Z$ is equal to $\mathrm{E}[X\,|\,\mathscr{D}]$". On the basis of Definition 4.1 this is equivalent to show that

a) $Z$ is $\mathscr{D}$-measurable; and
b) $\mathrm{E}[Z1_D] = \mathrm{E}[X1_D]$ for every $D \in \mathscr{D}$.

In fact it will be enough to check b) for every $D$ in a class $\mathscr{C} \subset \mathscr{D}$ generating $\mathscr{D}$ and stable with respect to finite intersections and containing $\Omega$. This is actually a simple application of Theorem 1.4: let $\mathscr{H}$ be the space of the $\mathscr{D}$-measurable bounded r.v.'s $W$ such that

$$\mathrm{E}[WX] = \mathrm{E}[WZ]\,. \tag{4.5}$$

Clearly $\mathscr{H}$ contains the function 1 and the indicator functions of the events of $\mathscr{C}$. Moreover, if $(W_n)_n$ is an increasing sequence of functions of $\mathscr{H}$ all bounded above by the same element $W^* \in \mathscr{H}$ and $W_n \uparrow W$, then we have, for every $n$, $W_1 \le W_n \le W^*$. As both $W_1$ and $W^*$ are bounded (as is every function of $\mathscr{H}$), with two applications of Lebesgue's theorem we have

$$\mathrm{E}[WX] = \lim_{n\to\infty} \mathrm{E}[W_nX] = \lim_{n\to\infty} \mathrm{E}[W_nZ] = \mathrm{E}[WZ].$$

Therefore condition i) of Theorem 1.4 is satisfied and $\mathscr{H}$ contains every $\sigma(\mathscr{C})$-bounded measurable r.v.; therefore (4.5) holds for every $\mathscr{D}$-bounded measurable r.v. $W$, which allows us to conclude the proof.

**Proposition 4.1** Let $X, X_1, X_2$ be *integrable* r.v.'s and $\alpha, \beta \in \mathbb{R}$. Then

a) $\mathrm{E}[\alpha X_1 + \beta X_2\,|\,\mathscr{D}] = \alpha\mathrm{E}[X_1\,|\,\mathscr{D}] + \beta\mathrm{E}[X_2\,|\,\mathscr{D}]$ a.s.
b) If $X \ge 0$ a.s., then $\mathrm{E}[X\,|\,\mathscr{D}] \ge 0$ a.s.
c) $\mathrm{E}[\mathrm{E}[X\,|\,\mathscr{D}]] = \mathrm{E}[X]$.

(continued)

**Proposition 4.1** (continued)
d) If $Z$ is bounded and $\mathscr{D}$-measurable then $\mathrm{E}[ZX\,|\,\mathscr{D}] = Z\mathrm{E}[X\,|\,\mathscr{D}]$ a.s.
e) If $\mathscr{D} \subset \mathscr{D}'$ then $\mathrm{E}[\mathrm{E}[X\,|\,\mathscr{D}']\,|\,\mathscr{D}] = \mathrm{E}[X\,|\,\mathscr{D}]$ a.s.
f) If $X$ is independent of $\mathscr{D}$ then $\mathrm{E}[X\,|\,\mathscr{D}] = \mathrm{E}[X]$ a.s.

*Proof* These are immediate applications of the definition; let us look more carefully at the proofs of the last three statements.

d) As it is immediate that the r.v. $Z\,\mathrm{E}[X\,|\,\mathscr{D}]$ is $\mathscr{D}$-measurable, we must only prove that the r.v. $Z\mathrm{E}[X\,|\,\mathscr{D}]$ satisfies the relation $\mathrm{E}\big[W\,Z\mathrm{E}[X\,|\,\mathscr{D}]\big] = \mathrm{E}[WZX]$ for every r.v. $W$. This is also immediate, as $ZW$ is itself bounded and $\mathscr{D}$-measurable.
e) The r.v. $\mathrm{E}[\mathrm{E}[X\,|\,\mathscr{D}']\,|\,\mathscr{D}]$ is $\mathscr{D}$-measurable; moreover, if $W$ is bounded and $\mathscr{D}$-measurable, then it is also $\mathscr{D}'$-measurable and, using c) and d),

$$\mathrm{E}\big[W\mathrm{E}[\mathrm{E}[X\,|\,\mathscr{D}']\,|\,\mathscr{D}]\big] = \mathrm{E}\big[\mathrm{E}[\mathrm{E}[WX\,|\,\mathscr{D}']\,|\,\mathscr{D}]\big] = \mathrm{E}\big[\mathrm{E}[WX\,|\,\mathscr{D}']\big] = \mathrm{E}[WX]\;,$$

which allows us to conclude the proof.
f) The r.v. $\omega \mapsto \mathrm{E}[X]$ is constant, hence $\mathscr{D}$-measurable. If $W$ is $\mathscr{D}$-measurable then it is independent of $X$ and

$$\mathrm{E}[WX] = \mathrm{E}[W]\mathrm{E}[X] = \mathrm{E}[W\mathrm{E}[X]]\;,$$

hence $\mathrm{E}[X] = \mathrm{E}[X\,|\,\mathscr{D}]$ a.s. □

It is easy to extend Proposition 4.1 to the case of r.v.'s that are only l.s.i. We shall only need to observe that a) holds only if $\alpha, \beta \ge 0$ (otherwise $\alpha X_1 + \beta X_2$ might not be l.s.i. anymore) and that d) holds only if $Z$ is bounded and *positive* (otherwise $ZX$ might not be l.s.i. anymore).

**Proposition 4.2** Let $X, X_n, n = 1, 2, \dots,$ be real l.s.i. r.v.'s. Then

a) (Beppo Levi) if $X_n \nearrow X$ a.s. then $\mathrm{E}[X_n\,|\,\mathscr{D}] \nearrow \mathrm{E}[X\,|\,\mathscr{D}]$ a.s.
b) (Fatou) If $\underline{\lim}_{n\to\infty} X_n = X$ and for every $n$ $X_n \ge Z$ for some integrable r.v. $Z$ then

$$\mathrm{E}[X\,|\,\mathscr{D}] \le \varliminf_{n\to\infty} \mathrm{E}[X_n\,|\,\mathscr{D}] \qquad \text{a.s.}$$

c) (Lebesgue) If $\lim_{n\to\infty} X_n = X$ and for every $n$ $|X_n| \le Z$ for some integrable r.v. $Z$ then

$$\lim_{n\to\infty} \mathrm{E}[X_n\,|\,\mathscr{D}] = \mathrm{E}[X\,|\,\mathscr{D}] \qquad \text{a.s.}$$

(continued)

**Proposition 4.2** (continued)

d) (Jensen's inequality) If $\Phi$ is a convex lower semi-continuous function $\mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ then $\Phi(Y)$ is also l.s.i. and

$$\mathrm{E}[\Phi(Y)\,|\,\mathscr{D}] \geq \Phi\big(\mathrm{E}[Y\,|\,\mathscr{D}]\big) \qquad \text{a.s.}$$

*Proof*

a) As $(X_n)_n$ is a.s. increasing, the same is true for the sequence $(\mathrm{E}[X_n\,|\,\mathscr{D}])_n$ by Proposition 4.1 b); hence the limit $\lim_{n\to\infty} \mathrm{E}[X_n\,|\,\mathscr{D}]$ exists a.s. Let us set $Z := \lim_{n\to\infty} \mathrm{E}[X_n\,|\,\mathscr{D}]$. By Beppo Levi's theorem applied twice we have, for every $D \in \mathscr{D}$,

$$\mathrm{E}[Z1_D] = \mathrm{E}\big[\lim_{n\to\infty} \mathrm{E}[X_n\,|\,\mathscr{D}]1_D\big] = \lim_{n\to\infty} \mathrm{E}[\mathrm{E}[X_n\,|\,\mathscr{D}]1_D] = \lim_{n\to\infty} \mathrm{E}[X_n 1_D] = \mathrm{E}[X1_D]\ ,$$

hence $Z = \mathrm{E}[X\,|\,\mathscr{D}]$ a.s.

b) If $Y_n = \inf_{k\geq n} X_k$ then

$$\lim_{n\to\infty} Y_n = \varliminf_{n\to\infty} X_n = X\ .$$

Moreover, $(Y_n)_n$ is increasing and $Y_n \leq X_n$. We can therefore apply a) and obtain

$$\mathrm{E}[X\,|\,\mathscr{D}] = \lim_{n\to\infty} \mathrm{E}[Y_n\,|\,\mathscr{D}] \leq \varliminf_{n\to\infty} \mathrm{E}[X_n\,|\,\mathscr{D}]\ .$$

c) Immediate consequence of b), applied first to $X_n$ and then to $-X_n$.

d) The proof is identical to the proof of the usual Jensen's inequality: recall that a convex l.s.c. function $\Phi$ is equal to the supremum of all the affine linear functions that are bounded above by $\Phi$. If $f(x) = ax + b$ is an affine function minorizing $\Phi$, then

$$\mathrm{E}[\Phi(X)\,|\,\mathscr{D}] \geq \mathrm{E}[f(X)\,|\,\mathscr{D}] = f\big(\mathrm{E}[X\,|\,\mathscr{D}]\big).$$

Then just take the upper bound in $f$ in the previous inequality among all the affine functions minorizing $\Phi$. □

Sometimes we shall write $\mathrm{P}(A\,|\,\mathscr{D})$ instead of $\mathrm{E}[1_A\,|\,\mathscr{D}]$ and shall speak of the *conditional probability of $A$ given $\mathscr{D}$*.

*Remark 4.3* It is immediate that, if $X = X'$ a.s., then $\mathrm{E}[X\,|\,\mathscr{D}] = \mathrm{E}[X'\,|\,\mathscr{D}]$ a.s. Actually, if $Z$ is a $\mathscr{D}$-measurable r.v. such that $\mathrm{E}[Z1_D] = \mathrm{E}[X1_D]$ for every $D \in \mathscr{D}$, then also $\mathrm{E}[Z1_D] = \mathrm{E}[X'1_D]$ for every $D \in \mathscr{D}$. The conditional expectation is therefore defined on equivalence classes of r.v.'s.

Let us investigate the action of the conditional expectation on $L^p$ spaces. We must not forget that $L^p$ is a space of equivalence classes of r.v.'s, not of r.v.'s. Taking care of this fact, Jensen's inequality (Proposition 4.2 d)), applied to the convex function $x \mapsto |x|^p$ with $p \geq 1$, gives $|\mathrm{E}[X\,|\,\mathscr{D}]|^p \leq \mathrm{E}[|X|^p\,|\,\mathscr{D}]$, so that

$$\mathrm{E}\big[\big|\mathrm{E}[X\,|\,\mathscr{D}]\big|^p\big] \leq \mathrm{E}\big[\mathrm{E}[|X|^p\,|\,\mathscr{D})\big] = \mathrm{E}[|X|^p]\,. \tag{4.6}$$

The conditional expectation is therefore a *continuous* linear operator $L^p \to L^p, p \geq 1$; moreover, it has norm $\leq 1$, i.e. it is a *contraction*. The image of $L^p$ through the operator $X \mapsto \mathrm{E}[X\,|\,\mathscr{D}]$ is the subspace of $L^p$, which we shall denote by $L^p(\mathscr{D})$, that is formed by the equivalence classes of r.v.'s that contain at least one $\mathscr{D}$-measurable representative. In particular, since $p \geq 1$, by (4.6)

$$\mathrm{E}\big[\big|\mathrm{E}[X_n\,|\,\mathscr{D}] - \mathrm{E}[X\,|\,\mathscr{D}]\big|^p\big] = \mathrm{E}\big[\big|\mathrm{E}[X_n - X\,|\,\mathscr{D}]\big|^p\big] \leq \mathrm{E}\big[\big|X_n - X\big|^p\big]\,,$$

we have that $X_n \underset{n\to\infty}{\overset{L^p}{\to}} X$ implies $\mathrm{E}[X_n\,|\,\mathscr{D}] \underset{n\to\infty}{\overset{L^p}{\to}} \mathrm{E}[X\,|\,\mathscr{D}]$.

The case $L^2$ deserves particular attention. If $X \in L^2$, we have for every bounded $\mathscr{D}$-measurable r.v. $W$,

$$\mathrm{E}\big[(X - \mathrm{E}[X\,|\,\mathscr{D}])W\big] = \mathrm{E}[XW] - \mathrm{E}\big[\mathrm{E}[X\,|\,\mathscr{D}]W\big] = 0\,. \tag{4.7}$$

As bounded r.v.'s are dense in $L^2$, the relation $\mathrm{E}[(X - \mathrm{E}[X\,|\,\mathscr{D}])W] = 0$ also holds for every r.v. $W \in L^2(\mathscr{D})$. In other words, $X - \mathrm{E}[X\,|\,\mathscr{D}]$ is orthogonal to $L^2(\mathscr{D})$, i.e. $\mathrm{E}[X\,|\,\mathscr{D}]$ is the orthogonal projection of $X$ on $L^2(\mathscr{D})$. In particular, this implies that

> $\mathrm{E}[X\,|\,\mathscr{D}]$ is the element of $L^2(\mathscr{D})$ that minimizes the $L^2$ distance from $X$.

If for simplicity we set $Z = \mathrm{E}[X\,|\,\mathscr{D}]$, then, for every $W \in L^2(\mathscr{D})$,

$$\begin{aligned}\|X - W\|_2^2 &= \mathrm{E}[(X - W)^2] = \mathrm{E}[(X - Z + Z - W)^2]\\ &= \mathrm{E}[(X - Z)^2] + 2\mathrm{E}[(X - Z)(Z - W)] + \mathrm{E}[(Z - W)^2]\,.\end{aligned}$$

(continued)

*Remark 4.3* (continued)
Note that $\mathrm{E}[(X-Z)(Z-W)] = 0$ thanks to (4.7), since $Z-W$ is $\mathscr{D}$-measurable. Hence the double product above vanishes and

$$\|X - W\|_2^2 = \mathrm{E}[(X-Z)^2] + \mathrm{E}[(Z-W)^2] \geq \mathrm{E}[(X-Z)^2] = \|X - Z\|_2^2\ ,$$

where the inequality is even strict, unless $Z = W$ a.s.

Therefore, in the sense of $L^2$, $\mathrm{E}[X\,|\,\mathscr{D}]$ *is the best approximation of* $X$ *by a* $\mathscr{D}$*-measurable r.v.* (Fig. 4.1).

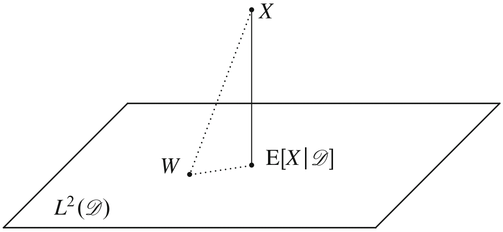


**Fig. 4.1** The $L^2$ distance between $\mathrm{E}[X\,|\,\mathscr{D}]$ and $X$ is the shortest because the angle between the segments $\mathrm{E}[X\,|\,\mathscr{D}] \to W$ and $\mathrm{E}[X\,|\,\mathscr{D}] \to X$ is $90°$

*Example 4.1* If $\mathscr{D} = \{\Omega, \emptyset\}$ is the trivial $\sigma$-algebra, then

$$\mathrm{E}[X\,|\,\mathscr{D}] = \mathrm{E}[X]\ .$$

Actually, the only $\mathscr{D}$-measurable r.v.'s are the constants and, if $c = \mathrm{E}[X\,|\,\mathscr{D}]$, then the constant $c$ is determined by the relationship $c = \mathrm{E}[\mathrm{E}[X\,|\,\mathscr{D}]] = \mathrm{E}[X]$. The notion of expectation thus appears to be a particular case of conditional expectation.

*Example 4.2* Let $A \in \mathscr{F}$ be an event such $\mathrm{P}(A) > 0$ and let $\mathscr{D}$ be the $\sigma$-algebra $\{A, A^c, \Omega, \emptyset\}$. Then $\mathrm{E}[X\,|\,\mathscr{D}]$, being $\mathscr{D}$-measurable, is constant on $A$ and on $A^c$. Its value, $c$ say, on $A$ is determined by the relation

$$\mathrm{E}[X1_A] = \mathrm{E}[1_A \mathrm{E}[X\,|\,\mathscr{D}]] = c\mathrm{P}(A)\ .$$

From this relation and the similar one for $A^c$ we easily derive that

$$\mathrm{E}[X\,|\,\mathscr{D}] = \begin{cases} \frac{1}{\mathrm{P}(A)}\,\mathrm{E}[X1_A] & \text{on } A \\ \frac{1}{\mathrm{P}(A^c)}\,\mathrm{E}[X1_{A^c}] & \text{on } A^c\ . \end{cases}$$

(continued)

*Example 4.2* (continued)
In particular, $\mathrm{E}[X|\mathscr{D}]$ is equal to $\int X\,d\mathrm{P}_A$ on $A$, where $\mathrm{P}_A$ is the conditional probability (4.1), and is equal to $\int X\,d\mathrm{P}_{A^c}$ on $A^c$.

*Example 4.3* Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion. Then if $s \le t$

$$\mathrm{E}[B_t|\mathscr{F}_s] = B_s \qquad \text{a.s.}$$

Indeed

$$\mathrm{E}[B_t|\mathscr{F}_s] = \mathrm{E}[B_t - B_s|\mathscr{F}_s] + \mathrm{E}[B_s|\mathscr{F}_s] = \mathrm{E}[B_t - B_s] + B_s = B_s$$

since $B_t - B_s$ is independent of $\mathscr{F}_s$ and centered, whereas $B_s$ is already $\mathscr{F}_s$-measurable.

The computation of a conditional expectation is an operation that we are led to perform quite often and which, sometimes, is even the objective of the researcher. Let us now make two remarks that can be of help towards this goal.

*Remark 4.4* Sometimes one must compute the conditional expectation of a r.v. $X$ with respect to a $\sigma$-algebra $\mathscr{D}$ that is obtained by adding to a $\sigma$-algebra $\mathscr{D}_0$ the family $\mathscr{N}$ of negligible events of a larger $\sigma$-algebra, $\mathscr{F}$ for example. It is useful to observe that

$$\mathrm{E}[X|\mathscr{D}] = \mathrm{E}[X|\mathscr{D}_0] \qquad \text{a.s.}$$

If we denote $\mathrm{E}[X|\mathscr{D}_0]$ by $Y$, for simplicity, then $Y$ is *a fortiori* $\mathscr{D}$-measurable. Moreover, for every event of the form $A \cap G$ with $A \in \mathscr{D}_0$ and $G \in \mathscr{N}$ or $G = \Omega$,

$$\mathrm{E}[Y 1_{A\cap G}] = \mathrm{E}[X 1_{A\cap G}]\ .$$

Actually, both sides vanish if $G \in \mathscr{N}$ whereas they obviously coincide if $G = \Omega$. As pointed out in Remark 3.1, the events of this form are stable with respect to finite intersections and generate the $\sigma$-algebra $\mathscr{D}$, therefore $Y = \mathrm{E}[X|\mathscr{D}]$ a.s.

Let $\mathscr{D}$ be a $\sigma$-algebra. We have seen two situations in which the computation of $\mathrm{E}[\cdot\,|\,\mathscr{D}]$ is easy: if $X$ is $\mathscr{D}$-measurable then

$$\mathrm{E}[X\,|\,\mathscr{D}] = X$$

whereas if it is independent of $\mathscr{D}$ then

$$\mathrm{E}[X\,|\,\mathscr{D}] = \mathrm{E}[X] \,. \tag{4.8}$$

The following lemma, which is quite (very!) useful, combines these two situations.

**Lemma 4.1 (The freezing lemma)** Let $(\Omega, \mathscr{F}, \mathrm{P})$ be a probability space and $\mathscr{G}$ and $\mathscr{D}$ independent sub-$\sigma$-algebras of $\mathscr{F}$. Let $X$ be a $\mathscr{D}$-measurable r.v. taking values in the measurable space $(E, \mathscr{E})$ and $\psi : E \times \Omega \to \mathbb{R}$ an $\mathscr{E} \otimes \mathscr{G}$-measurable function such that $\omega \mapsto \psi(X(\omega), \omega)$ is integrable. Then

$$\mathrm{E}[\psi(X, \cdot)\,|\,\mathscr{D}] = \Phi(X) \,, \tag{4.9}$$

where $\Phi(x) = \mathrm{E}[\psi(x, \cdot)]$.

*Proof* Let us assume first that $\psi$ is of the form $\psi(x, \omega) = f(x)Z(\omega)$, where $Z$ is $\mathscr{G}$-measurable. In this case, $\Phi(x) = f(x)\mathrm{E}[Z]$ and (4.9) becomes

$$\mathrm{E}[f(X)Z\,|\,\mathscr{D}] = f(X)\mathrm{E}[Z] \,,$$

which is immediately verified. The lemma is therefore proved for all functions $\psi$ of the type described above and, of course, for their linear combinations. One obtains the general case with the help of Theorem 1.5. □

Lemma 4.1 allows us to easily compute the conditional expectation in many instances. It says that you can just freeze one of the variables and compute the expectation of the resulting expression. The following examples show typical applications.

*Example 4.4* Let $B$ be an $m$-dimensional $(\mathscr{F}_t)_t$-Brownian motion and $f : \mathbb{R}^m \to \mathbb{R}$ a bounded Borel function. Let $s < t$. What is the value of $\mathrm{E}[f(B_t)\,|\,\mathscr{F}_s]$?

We can write $f(B_t) = f((B_t - B_s) + B_s)$: we are then in the situation of Lemma 4.1, i.e. the computation of the conditional expectation with respect to $\mathscr{F}_s$ of a function of a r.v. that is already $\mathscr{F}_s$-measurable and of a r.v. that is independent of $\mathscr{F}_s$. If we define $\psi(x, \omega) = f(x + B_t(\omega) - B_s(\omega))$, then $\psi$ is

(continued)

*Example 4.4* (continued)
$\mathscr{B}(\mathbb{R}^m) \otimes \mathscr{G}$-measurable with $\mathscr{G} = \sigma(B_t - B_s)$ and we have

$$\begin{aligned} \mathrm{E}[f(B_t)|\mathscr{F}_s] &= \mathrm{E}[f(B_s + (B_t - B_s))|\mathscr{F}_s] \\ &= \mathrm{E}[\psi(B_s, B_t - B_s)|\mathscr{F}_s] = \Phi(B_s) , \end{aligned} \tag{4.10}$$

where $\Phi(x) = \mathrm{E}[f(x + B_t - B_s)]$. Note that the r.v. $\mathrm{E}[f(B_t)|\mathscr{F}_s]$ just obtained turns out to be a function of $B_s$ and is therefore $\sigma(B_s)$-measurable. Hence

$$\mathrm{E}[f(B_t)|B_s] = \mathrm{E}\big[\mathrm{E}[f(B_t)|\mathscr{F}_s]|B_s\big] = \mathrm{E}[\Phi(B_s)|B_s] = \Phi(B_s) = \mathrm{E}[f(B_t)|\mathscr{F}_s] ,$$

i.e. the conditional expectation knowing the position at time $s$ is the same as the conditional expectation knowing the entire past of the process up to time $s$. We shall see later that this means that the Brownian motion is a Markov process.

It is also possible to make explicit the function $\Phi$: as $x+B_t-B_s \sim N(x, (t-s)I)$,

$$\Phi(x) = \mathrm{E}[f(x + B_t - B_s)] = \frac{1}{[2\pi(t-s)]^{m/2}} \int f(y) \exp\Big[-\frac{|y-x|^2}{2(t-s)}\Big]\, dy .$$

*Example 4.5 (The position of a Brownian motion at a random time)* Let $B$ be an $m$-dimensional Brownian motion and $\zeta$ a positive r.v. *independent* of $B$. How can we compute the law of the r.v. $B_\zeta$? Lemma 4.1 provides a simple way of doing this. We assume, for simplicity, $\zeta > 0$ a.s.

Let $A \subset \mathbb{R}^m$ be a Borel set. We have

$$\mathrm{P}(B_\zeta \in A) = \mathrm{E}[1_A(B_\zeta)] = \mathrm{E}\big[\mathrm{E}[1_A(B_\zeta)|\sigma(\zeta)]\big] .$$

This is a very common trick: instead of computing the expectation directly we first compute a conditional expectation and then its expectation. In the computation of the conditional expectation above we are in the situation of Lemma 4.1 as $B_\zeta$ is a function of the r.v. $\zeta$, which is $\sigma(\zeta)$-measurable, and of the Brownian motion, which is independent of $\sigma(\zeta)$. In order to apply the freezing lemma let $A \in \mathscr{B}(\mathbb{R}^m)$ and let us define $\psi(t, \omega) = 1_A(B_t(\omega))$, so that $1_A(B_\zeta) = \psi(\zeta, \cdot)$. Therefore

$$\mathrm{E}\big[\mathrm{E}[1_A(B_\zeta)|\sigma(\zeta)]\big] = \mathrm{E}[\Phi(\zeta)]$$

(continued)

*Example 4.5* (continued)
where

$$\Phi(t) = \mathrm{E}[\psi(t,\cdot)] = \mathrm{E}[1_A(B_t)] = \frac{1}{(2\pi t)^{m/2}} \int_A \mathrm{e}^{-\frac{|x|^2}{2t}}\, dx\,.$$

By Fubini's theorem and denoting the law of $\zeta$ by $\nu$,

$$\begin{aligned} \mathrm{P}(B_\zeta \in A) = \mathrm{E}[1_A(B_\zeta)] = \mathrm{E}[\Phi(\zeta)] &= \int_0^{+\infty} \Phi(t)\, d\nu(t) \\ &= \int_0^{+\infty} d\nu(t) \int_A \frac{1}{(2\pi t)^{m/2}} \mathrm{e}^{-\frac{|x|^2}{2t}}\, dx \\ &= \int_A dx\, \underbrace{\int_0^{+\infty} \frac{1}{(2\pi t)^{m/2}} \mathrm{e}^{-\frac{|x|^2}{2t}}\, d\nu(t)}_{:=g(x)} \end{aligned} \tag{4.11}$$

so that the density $g$ of $B_\zeta$ is obtained by computing the integral.

One could also compute the characteristic function of $B_\zeta$ in quite a similar way: if $\psi(t,\omega) = \mathrm{e}^{i\langle\theta, B_t(\omega)\rangle}$, we can write

$$\mathrm{E}[\mathrm{e}^{i\langle\theta, B_\zeta\rangle}] = \mathrm{E}\big[\mathrm{E}[\psi(\zeta,\omega)\,|\,\sigma(\zeta)]\big] = \mathrm{E}[\Phi(\zeta)]$$

where now $\Phi(\zeta) = \mathrm{E}[\psi(t,\omega)] = \mathrm{E}[\mathrm{e}^{i\langle\theta, B_t\rangle}] = \mathrm{e}^{-\frac{t}{2}|\theta|^2}$. Therefore

$$\mathrm{E}[\mathrm{e}^{i\langle\theta, B_\zeta\rangle}] = \int_0^{+\infty} \mathrm{e}^{-\frac{t}{2}|\theta|^2}\, d\nu(t)\,. \tag{4.12}$$

In order to determine the law of $B_\zeta$ we can therefore choose whether to compute the density using (4.11) or the characteristic function using (4.12), according to the difficulty of the computation of the corresponding integral. Exercises 4.6 and 4.8 give some important examples of the application of this technique.

*Remark 4.5* Sometimes we are confronted with the computation of something of the kind

$$\mathrm{E}\Big[\int_0^T X_t\, dt \,\Big|\, \mathscr{D}\Big]\,, \tag{4.13}$$

where $X$ is some integrable measurable process and we are tempted to write

$$\mathrm{E}\Big[\int_0^T X_t\, dt \,\Big|\, \mathscr{D}\Big] = \int_0^T \mathrm{E}[X_t\,|\,\mathscr{D}]\, dt\,. \tag{4.14}$$

(continued)

*Remark 4.5* (continued)
Is this correct? It is immediate that if $W$ is a bounded $\mathscr{D}$-measurable r.v. then

$$\mathrm{E}\Big[W\int_0^T X_t\,dt\Big] = \int_0^T \mathrm{E}[WX_t]\,dt = \int_0^T \mathrm{E}\big[W\mathrm{E}[X_t\,|\,\mathscr{D}]\big]\,dt$$
$$= \mathrm{E}\Big[W\int_0^T \mathrm{E}[X_t\,|\,\mathscr{D}]\,dt\Big]\,.$$

There are, however, a couple of issues to be fixed. The first is that the quantity $\mathrm{E}[X_t\,|\,\mathscr{D}]$ is, for every $t$, defined only a.s. and we do not know whether it has a version such that $t \mapsto \mathrm{E}[X_t\,|\,\mathscr{D}]$ is integrable.

The second is that we must also prove that the r.v.

$$Z = \int_0^T \mathrm{E}[X_t\,|\,\mathscr{D}]\,dt \tag{4.15}$$

is actually $\mathscr{D}$-measurable.

In practice, without looking for a general statement, these question are easily handled: very often we shall see that $t \mapsto \mathrm{E}[X_t\,|\,\mathscr{D}]$ has a continuous version so that the integral in (4.15) is well defined. In this case, the r.v. $Z$ of (4.15) is also $\mathscr{D}$-measurable, as it is the limit of Riemann sums of the form

$$\sum \mathrm{E}[X_{t_i}\,|\,\mathscr{D}](t_{i+1}-t_i)\;,$$

which are certainly $\mathscr{D}$-measurable.

*Example 4.6* Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion. Compute, for $s \le t$,

$$\mathrm{E}\Big[\int_0^t B_u\,du\,\Big|\,\mathscr{F}_s\Big]\,.$$

We have $\mathrm{E}[B_u\,|\,\mathscr{F}_s] = B_s$ if $u \ge s$ and $\mathrm{E}[B_u\,|\,\mathscr{F}_s] = B_u$ if $u \le s$ (as $B_u$ is then already $\mathscr{F}_s$-measurable), i.e.

$$\mathrm{E}[B_u\,|\,\mathscr{F}_s] = B_{s\wedge u}\,.$$

This is a continuous process, so that we can apply formula (4.14), which gives

$$\mathrm{E}\Big[\int_0^t B_u\,du\,\Big|\,\mathscr{F}_s\Big] = \int_0^t \mathrm{E}[B_u\,|\,\mathscr{F}_s]\,du = \int_0^t B_{s\wedge u}\,du = \int_0^s B_u\,du + (t-s)B_s\,.$$

## 4.3 Conditional laws

At the beginning of this chapter we spoke of conditional distributions given the value of some discrete r.v. $Z$, but then we investigated the conditional expectations, i.e. the expectations of these conditional distributions. Let us go back now to conditional distributions.

Let $Y, X$ be r.v.'s with values in the measurable spaces $(G, \mathscr{G})$ and $(E, \mathscr{E})$, respectively, and let us denote by $\nu_Y$ the law of $Y$.

A family of probabilities $(n(y, dx))_{y\in G}$ on $(E, \mathscr{E})$ is a *conditional law of $X$ given $Y$* if,

i) For every $A \in \mathscr{E}$, the map $y \mapsto n(y, A)$ is $\mathscr{G}$-measurable.
ii) For every $A \in \mathscr{E}$ and $B \in \mathscr{G}$,

$$\mathrm{P}(X \in A, Y \in B) = \int_B n(y, A)\, \nu_Y(dy)\ . \tag{4.16}$$

Intuitively $n(y, dx)$, which is a probability on $(E, \mathscr{E})$, is the law that it is suitable to consider for the r.v. $X$, keeping into account the information that $Y = y$. (4.16) implies that, if $f : E \to \mathbb{R}$ and $g : G \to \mathbb{R}$ are functions that are linear combinations of indicator functions, then

$$\mathrm{E}[f(X)\psi(Y)] = \int_G \Big(\int_E f(x)\, n(y, dx)\Big)\psi(y)\, \nu_Y(dy)\ . \tag{4.17}$$

Actually, (4.17) coincides with (4.16) if $f = 1_A$ and $\psi = 1_B$. Therefore, by linearity, (4.17) is true if $f$ and $\psi$ are linear combinations of indicator functions. With the usual approximation arguments of measure theory, such as Proposition 1.11 for example, we have that (4.17) holds for every choice of positive functions $f, \psi$ and then, decomposing into the sum of the positive and negative parts, for every choice of functions $f, \psi$ such that $f(X)\psi(Y)$ is integrable.

If we set $g(y) = \int_E f(x)\, n(y, dx)$, then (4.17) can be written as

$$\mathrm{E}[f(X)\psi(Y)] = \int_G g(y)\psi(y)\, \nu_Y(dy) = \mathrm{E}[g(Y)\psi(Y)]\ .$$

A comparison with (4.4) shows that this means that $(n(y, dx))_{y\in G}$ is a conditional distribution of $X$ given $Y$ if and only if, for every bounded measurable function $f$

such that $f(X)$ is integrable,

$$\mathrm{E}[f(X)\,|\,Y] = g(Y) = \int_E f(x)\,n(Y,dx) \qquad \text{a.s.} \tag{4.18}$$

i.e. $\mathrm{E}[f(X)\,|\,Y = y] = g(y)$. In particular, if $X$ is an integrable real r.v., choosing $f(x) = x$ we have

$$\mathrm{E}[X\,|\,Y] = \int_{\mathbb{R}} x\,n(Y,dx) \qquad \text{a.s.} \tag{4.19}$$

i.e. $\mathrm{E}[X\,|\,Y = y] = \int_{\mathbb{R}} x\,n(y,dx)$. We then find the conditional expectation as the mean of the conditional law.

The next example highlights a general situation where the computation of the conditional law is easy.

*Example 4.7* Let $X, Y$ be respectively $\mathbb{R}^d$- and $\mathbb{R}^m$-valued r.v.'s having joint density $h(x,y)$ with respect to the Lebesgue measure of $\mathbb{R}^d \times \mathbb{R}^m$. Let

$$h_Y(y) = \int_{\mathbb{R}^d} h(x,y)\,dx$$

be the marginal density of $Y$ (see Sect. 1.4) and let $Q = \{y; h_Y(y) = 0\}$. Obviously $\mathrm{P}(Y \in Q) = \int_Q h_Y(y)\,dy = 0$. If we define

$$\overline{h}(x;y) = \begin{cases} \dfrac{h(x,y)}{h_Y(y)} & \text{if } y \notin Q \\ \text{any arbitrary density} & \text{if } y \in Q\,, \end{cases} \tag{4.20}$$

we have immediately that $n(y,dx) = \overline{h}(x;y)\,dx$ is a conditional law of $X$ given $Y = y$: if $f$ and $g$ are bounded measurable functions on $\mathbb{R}^d$ and $\mathbb{R}^m$ respectively, then

$$\begin{aligned} \mathrm{E}[f(X)g(Y)] &= \int_{\mathbb{R}^m}\int_{\mathbb{R}^d} f(x)g(y)h(x,y)\,dy\,dx \\ &= \int_{\mathbb{R}^m} g(y)h_Y(y)\,dy \int_{\mathbb{R}^d} f(x)\overline{h}(x;y)\,dx\,. \end{aligned}$$

Therefore, for every measurable function $f$ such that $f(X)$ is integrable,

$$\mathrm{E}[f(X)\,|\,Y = y] = \int_{\mathbb{R}^m} f(x)\overline{h}(x;y)\,dx\,,$$

(continued)

*Example 4.7* (continued)
which allows us to compute explicitly the conditional expectation in many situations.

Note, however, that we have proved that the conditional expectation $\mathrm{E}[X\,|\,Y]$ always exists whereas, until now at least, we know nothing about the existence of a conditional distribution of $X$ given $Y = y$.

## 4.4 Conditional laws of Gaussian vectors

In this section we see a particular case of computation of conditional laws (hence also of conditional expectations) when the r.v. $X$ (whose conditional law we want to compute) and $Y$ (the conditioning r.v.) are jointly Gaussian. In order to do this it is also possible to use the procedure of Example 4.7 (taking the quotient of the joint distribution and the marginal), but the method we are about to present is much more convenient (besides the fact that a joint density of $X$ and $Y$ does not necessarily exist).

Let $X, Y$ be Gaussian vectors with values in $\mathbb{R}^k$ and $\mathbb{R}^p$ respectively. Let us assume that their joint law on $(\mathbb{R}^{k+p}, \mathscr{B}(\mathbb{R}^{k+p}))$ is Gaussian with mean and covariance matrix given respectively by

$$\begin{pmatrix} m_X \\ m_Y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} C_X & C_{XY} \\ C_{YX} & C_Y \end{pmatrix} ,$$

where $C_X, C_Y$ are the covariance matrices of $X$ and $Y$, respectively, and

$$C_{XY} = \mathrm{E}[(X - \mathrm{E}[X])\,(Y - \mathrm{E}[Y])^*] = \; C_{YX}^*$$

is the $k \times p$ matrix of the covariances of the components of $X$ and of those of $Y$; let us assume, moreover, that $C_Y$ is strictly positive definite (and therefore invertible).

Let us look first for a $k\times p$ matrix, $A$, such that $X - AY$ and $Y$ are independent. Let $Z = X - AY$. The pair $(X, Y)$ is Gaussian and the same is true for $(Z, Y)$ which is a linear function of $(X, Y)$. Therefore, for the independence of $Z$ and $Y$, it is sufficient to check that $\mathrm{Cov}(Z_i, Y_j) = 0$ for every $i = 1, \dots, k, j = 1, \dots, p$. In order to keep the notation simple, let us first assume that the means $m_X$ and $m_Y$ vanish so that the covariance coincides with the expectation of the product. The condition of absence of correlation between the components of $Z$ and $Y$ can then be expressed as

$$0 = \mathrm{E}[ZY^*] = \mathrm{E}[(X - AY)Y^*] = \mathrm{E}[XY^*] - A\mathrm{E}[YY^*] = C_{XY} - AC_Y \, .$$

Therefore the required property holds with $A = C_{XY}C_Y^{-1}$. If we remove the hypothesis that the means vanish, it suffices to repeat the same computation with $X - m_X$ and $Y - m_Y$ instead of $X$ and $Y$. Thus we can write

$$X = AY + (X - AY) ,$$

where $X - AY$ and $Y$ are independent. It is rather intuitive now (see Exercise 4.10 for a rigorous, albeit simple, verification) that the conditional law of $X$ given $Y = y$ is the law of $Ay + X - AY$, since, intuitively, the knowledge of the value of $Y$ does not give any information on the value of $X - AY$. As $X - AY$ is Gaussian, this law is characterized by its mean

$$Ay + m_X - Am_Y = m_X + C_{XY}C_Y^{-1}(y - m_Y)$$

and covariance matrix (recall that $A = C_{XY}C_Y^{-1}$)

$$\begin{aligned} C_{X-AY} &= \mathrm{E}[((X - m_X) - A(Y - m_Y))((X - m_X) - A(Y - m_Y))^*] \\ &= C_X - C_{XY}A^* - AC_{YX} + AC_YA^* \\ &= C_X - C_{XY}C_Y^{-1}C_{XY}^* - C_{XY}C_Y^{-1}C_{XY}^* + C_{XY}C_Y^{-1}C_YC_Y^{-1}C_{XY}^* \\ &= C_X - C_{XY}C_Y^{-1}C_{XY}^* , \end{aligned}$$

where we took advantage of the fact that $C_Y$ is symmetric and of the relation $C_{YX} = C_{XY}^*$. In conclusion

the conditional distribution of $X$ given $Y = y$ is Gaussian with mean

$$m_X + C_{XY}C_Y^{-1}(y - m_Y) \tag{4.21}$$

and covariance matrix

$$C_X - C_{XY}C_Y^{-1}C_{XY}^*. \tag{4.22}$$

In particular, from (4.21), we obtain the value of the conditional expectation, which is equal to

$$\mathrm{E}[X|Y] = m_X + C_{XY}C_Y^{-1}(Y - m_Y) .$$

If $X$ and $Y$ are both one-dimensional, this formula becomes

$$\mathrm{E}[X|Y] = m_X + \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(Y)}(Y - m_Y) , \tag{4.23}$$

whereas the variance (4.22) of the conditional distribution of $X$ given $Y = y$ is

$$\mathrm{Var}(X) - \frac{\mathrm{Cov}(X, Y)^2}{\mathrm{Var}(Y)} \tag{4.24}$$

and is therefore always smaller that the variance of $X$. Let us point out two important things in the previous computation:

- The conditional laws of a Gaussian vector are Gaussian themselves.
- Only the mean of the conditional law depends on the value $y$ of the conditioning r.v. $Y$. The covariance matrix of the conditional law *does not* depend on the value of $Y$ and can therefore be computed before knowing the observation $Y$.

## 4.5 The augmented Brownian filtration

Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional (continuous) Brownian motion. Let us consider its natural filtration $\mathscr{G}_t = \sigma(B_s, s \le t)$ and the augmented natural filtration $\overline{\mathscr{G}} = \sigma(\mathscr{G}_t, \mathscr{N})$ (see p. 32). We know that $B$ is still a Brownian motion if considered with respect to $(\mathscr{G}_t)_t$ and also with respect to $(\overline{\mathscr{G}}_t)_t$ (Remark 3.1). In this section we show that this filtration is right-continuous and therefore $(\Omega, \mathscr{F}, (\overline{\mathscr{G}}_t)_t, (B_t)_t, \mathrm{P})$ is a standard Brownian motion. More precisely, let

$$\overline{\mathscr{G}}_{s+} = \bigcap_{u>s} \overline{\mathscr{G}}_u, \qquad \overline{\mathscr{G}}_{s-} = \bigvee_{u<s} \overline{\mathscr{G}}_u \,.$$

It is immediate that $\mathscr{G}_{s-} = \mathscr{G}_s$. Indeed, as $B$ is continuous, we have

$$\lim_{u \to s-} B_u = B_s \,.$$

Hence the r.v. $B_s$, being the limit of $\mathscr{G}_{s-}$-measurable r.v.'s, is $\mathscr{G}_{s-}$-measurable itself, which implies $\mathscr{G}_s \subset \mathscr{G}_{s-}$. More precisely, we have the following

**Proposition 4.3** $\overline{\mathscr{G}}_{s-} = \overline{\mathscr{G}}_s = \overline{\mathscr{G}}_{s+}$.

*Proof* Of course, we only need to prove the rightmost equality. We shall assume for simplicity $m = 1$. It is sufficient to show that for every bounded $\overline{\mathscr{G}}_\infty$-measurable r.v. $W$

$$\mathrm{E}[W|\overline{\mathscr{G}}_{s+}] = \mathrm{E}[W|\overline{\mathscr{G}}_s] \quad \text{a.s.} \tag{4.25}$$

This relation, applied to a r.v. $W$ that is already $\overline{\mathscr{G}}_{s+}$-measurable, will imply that it is also $\overline{\mathscr{G}}_s$-measurable and therefore that $\overline{\mathscr{G}}_s \supset \overline{\mathscr{G}}_{s+}$. As the reciprocal inclusion is obvious, the statement will be proved. Note that here we use the fact that the $\sigma$-algebras $\overline{\mathscr{G}}_s$ contain all the negligible events of $\overline{\mathscr{G}}_\infty$: thanks to this fact, if a r.v. is a.s. equal to a r.v. which is $\overline{\mathscr{G}}_s$-measurable, then it is $\overline{\mathscr{G}}_s$-measurable itself.

We have by Lebesgue's theorem, for $t > s$,

$$\begin{aligned}\mathrm{E}[\mathrm{e}^{i\alpha B_t}|\overline{\mathscr{G}}_{s+}] &= \mathrm{e}^{i\alpha B_s}\mathrm{E}[\mathrm{e}^{i\alpha(B_t-B_s)}|\overline{\mathscr{G}}_{s+}] = \mathrm{e}^{i\alpha B_s}\mathrm{E}\Big[\lim_{\varepsilon\to 0}\mathrm{e}^{i\alpha(B_t-B_{s+\varepsilon})}\,\Big|\,\overline{\mathscr{G}}_{s+}\Big]\\ &= \mathrm{e}^{i\alpha B_s}\lim_{\varepsilon\to 0}\mathrm{E}[\mathrm{e}^{i\alpha(B_t-B_{s+\varepsilon})}|\overline{\mathscr{G}}_{s+}]\;,\end{aligned}$$

as $\mathrm{e}^{i\alpha(B_t-B_{s+\varepsilon})} \to \mathrm{e}^{i\alpha(B_t-B_s)}$ as $\varepsilon \to 0$ and these r.v.'s are bounded by 1. As, for $\varepsilon > 0$, $B_t - B_{s+\varepsilon}$ is independent of $\mathscr{G}_{s+\varepsilon}$ and *a fortiori* of $\mathscr{G}_{s+}$,

$$\begin{aligned}\mathrm{E}[\mathrm{e}^{i\alpha B_t}|\overline{\mathscr{G}}_{s+}] &= \mathrm{e}^{i\alpha B_s}\lim_{\varepsilon\to 0}\mathrm{E}[\mathrm{e}^{i\alpha(B_t-B_{s+\varepsilon})}|\overline{\mathscr{G}}_{s+}] = \mathrm{e}^{i\alpha B_s}\lim_{\varepsilon\to 0}\mathrm{E}[\mathrm{e}^{i\alpha(B_t-B_{s+\varepsilon})}]\\ &= \mathrm{e}^{i\alpha B_s}\lim_{\varepsilon\to 0}\mathrm{e}^{-\frac{1}{2}\alpha^2(t-s-\varepsilon)} = \mathrm{e}^{i\alpha B_s}\mathrm{e}^{-\frac{1}{2}\alpha^2(t-s)}\;.\end{aligned}$$

Hence, the right-hand side above being $\mathscr{G}_s$-measurable, we have, for $t \in \mathbb{R}^+$, $\alpha \in \mathbb{R}$,

$$\mathrm{E}[\mathrm{e}^{i\alpha B_t}|\overline{\mathscr{G}}_{s+}] = \mathrm{E}[\mathrm{e}^{i\alpha B_t}|\overline{\mathscr{G}}_s] \qquad \text{a.s.} \tag{4.26}$$

(if $t \le s$ this relation is obvious).

We now prove that (4.26) implies that (4.25) holds for every bounded $\overline{\mathscr{G}}_\infty$-measurable r.v. $W$. Let us denote by $\mathscr{H}$ the vector space of the real r.v.'s $Z$ such that $\mathrm{E}(Z|\overline{\mathscr{G}}_{s+}] = \mathrm{E}[Z|\overline{\mathscr{G}}_s]$ a.s. We have seen that $\mathscr{H}$ contains the linear combinations of r.v.'s of the form $\Re \mathrm{e}^{i\alpha B_t}$ and $\Im \mathrm{e}^{i\alpha B_t}$, $t \in \mathbb{R}^+$, which is a subspace stable under multiplication; it also contains the r.v.'s that vanish a.s., since for such r.v.'s both $\mathrm{E}[Z|\overline{\mathscr{G}}_{s+}]$ and $\mathrm{E}[Z|\overline{\mathscr{G}}_s]$ vanish a.s. and are $\overline{\mathscr{G}}_t$-measurable for every $t$, as the $\sigma$-algebras $\overline{\mathscr{G}}_t$ contain all negligible events.

Hence $\mathscr{H}$ satisfies the hypotheses of Theorem 1.5 and therefore contains every bounded r.v. that is measurable with respect to the smallest $\sigma$-algebra, $\mathscr{A}$, generated by the r.v.'s $\mathrm{e}^{i\alpha B_t}$, $\alpha \in \mathbb{R}$, $t \in \mathbb{R}^+$ and by the negligible events of $\mathscr{F}$. But, as

$$B_t = \lim_{\alpha\to 0}\frac{\mathrm{e}^{i\alpha B_t} - \mathrm{e}^{-i\alpha B_t}}{2i\alpha}\;,$$

the r.v.'s $B_t$ are $\mathscr{A}$-measurable and thus $\mathscr{A} \supset \overline{\mathscr{G}}_t$ for every $t \geq 0$, hence $\mathscr{A} \supset \overline{\mathscr{G}}_\infty$, which concludes the proof. □

**Corollary 4.1** $B = (\Omega, \mathscr{F}, (\overline{\mathscr{G}}_t)_t, (B_t)_t, \mathrm{P})$ is a standard Brownian motion.

## Exercises

**4.1** (p. 474) Let $X$ be a real r.v. defined on a probability space $(\Omega, \mathscr{F}, \mathrm{P})$ and $\mathscr{G} \subset \mathscr{F}$ a sub-$\sigma$-algebra. Let $\mathscr{D} \subset \mathscr{F}$ be another $\sigma$-algebra independent of $X$ and independent of $\mathscr{G}$.

a) Is it true that

$$\mathrm{E}[X|\mathscr{G} \vee \mathscr{D}] = \mathrm{E}[X|\mathscr{G}]\ ? \tag{4.27}$$

b) Show that if $\mathscr{D}$ is independent of $\sigma(X) \vee \mathscr{G}$, then (4.27) is true.

Use the criterion of Remark 4.2.

**4.2** (p. 475) On $(\Omega, \mathscr{F}, \mathrm{P})$ let $\mathscr{G} \subset \mathscr{F}$ be a sub-$\sigma$-algebra and $Y$ a r.v. independent of $\mathscr{G}$. Show that $Y$ cannot be $\mathscr{G}$-measurable unless it is a.s. constant.

**4.3** (p. 476) Let P and Q be probabilities on $(\Omega, \mathscr{F})$ and let us assume that Q has a density, $Z$, with respect to P, i.e. $\mathrm{Q}(A) = \mathrm{E}(1_A Z)$ for every $A \in \mathscr{F}$. Let $X : (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ be a r.v. and let $\mu_\mathrm{P} = X(\mathrm{P})$, $\mu_\mathrm{Q} = X(\mathrm{Q})$ be the laws of $X$ with respect to P and Q, respectively. Prove that $\mu_\mathrm{P} \gg \mu_\mathrm{Q}$ and that $\frac{d\mu_\mathrm{Q}}{d\mu_\mathrm{P}} = f$, where $f(x) = \mathrm{E}[Z|X = x]$ (E denotes the expectation with respect to P).

**4.4** (p. 476) (Conditioning under a change of probability) Let $Z$ be a real *positive* r.v. defined on $(\Omega, \mathscr{F}, \mathrm{P})$ and $\mathscr{G}$ a sub-$\sigma$-algebra of $\mathscr{F}$.

a) Prove that, a.s., $\{Z = 0\} \supset \{\mathrm{E}[Z|\mathscr{G}] = 0\}$ (i.e. by conditioning the set of zeros of a positive r.v. shrinks). Prove that for every r.v. $Y$ such that $YZ$ is integrable,

$$\mathrm{E}[ZY|\mathscr{G}] = \mathrm{E}[ZY|\mathscr{G}]1_{\{\mathrm{E}(Z|\mathscr{G})>0\}} \qquad \text{a.s.} \tag{4.28}$$

b) Let us assume, moreover, that $\mathrm{E}[Z] = 1$. Let Q be the probability on $(\Omega, \mathscr{F})$ having density $Z$ with respect to P and let $\mathrm{E}^\mathrm{Q}$ denote the expectation with respect to Q. Show that $\mathrm{E}[Z|\mathscr{G}] > 0$ Q-a.s. (E still denotes the expectation with respect to P). Show that, if $Y$ is Q-integrable,

$$\mathrm{E}^\mathrm{Q}[Y|\mathscr{G}] = \frac{\mathrm{E}[YZ|\mathscr{G}]}{\mathrm{E}[Z|\mathscr{G}]} \qquad \mathrm{Q}\text{–a.s.} \tag{4.29}$$

a) Try assuming $Y \geq 0$ first.

**4.5** (p. 477) Let $\mathscr{D} \subset \mathscr{F}$ be a sub-$\sigma$-algebra and $X$ an $m$-dimensional r.v. such that for every $\lambda \in \mathbb{R}^m$

$$\mathrm{E}[\mathrm{e}^{i\langle \lambda, X\rangle} \,|\, \mathscr{D}] = \mathrm{E}[\mathrm{e}^{i\langle \lambda, X\rangle}] \qquad \text{a.s.}$$

Then $X$ is independent of $\mathscr{D}$.

**4.6** (p. 477) Let $B$ be an $m$-dimensional Brownian motion and $\zeta$ an exponential r.v. with parameter $\lambda$ and independent of $B$.

a) Compute the characteristic function of $B_\zeta$ (the position of the Brownian motion at the random time $\zeta$).

b1) Let $X$ be a real r.v. with a Laplace density with parameter $\mu$, i.e. with density

$$f_X(x) = \frac{\mu}{2}\,\mathrm{e}^{-\mu|x|}\,.$$

Compute the characteristic function of $X$.

b2) What is the law of $B_\zeta$ for $m = 1$?

**4.7** (p. 478) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and let $\zeta : \Omega \to \mathbb{R}^+$ be a positive r.v. (not necessarily a stopping time of $(\mathscr{F}_t)_t$) independent of $B$.

Prove that $X_t = B_{\zeta+t} - B_\zeta$ is a Brownian motion and specify with respect to which filtration.

**4.8** (p. 478) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a two-dimensional Brownian motion. Let $a > 0$ and let $\tau = \inf\{t; B_2(t) = a\}$ be the passage time of $B_2$ at $a$, which is also the entrance time of $B$ in the line $y = a$. Recall (Sect. 3.6) that $\tau < +\infty$ a.s.

a) Show that the $\sigma$-algebras $\sigma(\tau)$ and $\mathscr{G}_1 = \sigma(B_1(u), u \geq 0)$ are independent.

b) Compute the law of $B_1(\tau)$ (i.e. the law of the abscissa of $B$ at the time it reaches the line $y = a$).

b) Recall Example 4.5. The law of $\tau$ is computed in Exercise 3.20...

**4.9** (p. 479) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion. Compute

$$\mathrm{E}\Big(\int_s^t B_u^2\,du \,\Big|\, \mathscr{F}_s\Big) \quad \text{and} \quad \mathrm{E}\Big(\int_s^t B_u^2\,du | B_s\Big)\,.$$

**4.10** (p. 479) Let $(E, \mathscr{E})$, $(G, \mathscr{G})$ be measurable spaces, $X$ an $E$-valued r.v. such that

$$X = \phi(Y) + Z\,,$$

where $Y$ and $Z$ are independent r.v.'s with values respectively in $E$ and $G$ and where $\phi : G \to E$ is measurable.

Show that the conditional law of $X$ given $Y = y$ is the law of $Z + \phi(y)$.

**4.11** (p. 480)

a) Let $X$ be a signal having a Gaussian $N(0, 1)$ law. An observer has no access to the value of $X$ and only knows an observation $Y = X + W$, where $W$ is a noise, independent of $X$ and $N(0, \sigma^2)$-distributed. What is your estimate of the value $X$ of the signal knowing that $Y = y$?
b) The same observer, in order to improve its estimate of the signal $X$, decides to take two observations $Y_1 = X + W_1$ and $Y_2 = X + W_2$, where $W_1$ and $W_2$ are $N(0, \sigma^2)$-distributed and the three r.v.'s $X, W_1, W_2$ are independent. What is the estimate of $X$ now given $Y_1 = y_1$ and $Y_2 = y_2$? Compare the variance of the conditional law of $X$ given the observation in situations a) and b).

**4.12** (p. 481)

a) Let $B$ be a Brownian motion. What is the conditional law of $(B_{t_1}, \dots, B_{t_m})$ given $B_1 = y$, $0 \le t_1 < \dots < t_m < 1$?
b) What is the conditional law of $(B_{t_1}, \dots, B_{t_m})$ given $B_1 = y, B_v = x$, $0 \le t_1 < \dots < t_m < 1 < v$?

**4.13** (p. 483) Let $B$ be a Brownian motion. What is the joint law of

$$B_1 \quad \text{and} \quad \int_0^1 B_s\, ds\ ?$$

Let us assume we know that $\int_0^1 B_s\, ds = x$. What is the best estimate of the position $B_1$ of the Brownian motion at time 1?

The joint law is Gaussian, see Exercise 3.11.

**4.14** (p. 483) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and $\eta$ an $N(\mu, \rho^2)$-distributed r.v. *independent* of $(B_t)_t$. Let

$$Y_t = \eta t + \sigma B_t$$

and $\mathscr{G}_t = \sigma(Y_s, s \le t)$. Intuitively the meaning of this exercise is the following: starting from the observation of a path $Y_s(\omega), s \le t$, how can the unknown value of $\eta(\omega)$ be estimated? How does this estimate behave as $t \to \infty$? Will it converge to $\eta(\omega)$?

a) Compute $\mathrm{Cov}(\eta, Y_s)$, $\mathrm{Cov}(Y_s, Y_t)$.
b) Show that $(Y_t)_t$ is a Gaussian process.
c) Prove that, for every $t \ge 0$, there exists a number $\lambda$ (depending on $t$) such that $\eta = \lambda Y_t + Z$, with $Z$ independent of $\mathscr{G}_t$.

d) Compute $\mathrm{E}[\eta|\mathscr{G}_t]$ and the variance of $\eta - \mathrm{E}[\eta|\mathscr{G}_t]$. Show that

$$\lim_{t\to+\infty} \mathrm{E}[\eta|\mathscr{G}_t] = \eta \qquad \text{a.s.}$$

**4.15** (p. 484) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a natural Brownian motion and, for $0 \le t \le 1$, let

$$X_t = B_t - tB_1 \ .$$

The process $(X_t)_t$, defined for $0 \le t \le 1$, is called the *Brownian bridge* .

a) Show that $(X_t)_t$ is a centered Gaussian process independent of $B_1$. Compute $\mathrm{E}(X_t X_s)$.
b) Show that, for $s \le t$, the r.v.

$$X_t - \frac{1-t}{1-s} X_s$$

is independent of $X_s$.
c) Compute $\mathrm{E}[X_t|X_s]$ and show that, with $\mathscr{G}_s = \sigma(X_u, u \le s)$, for $s \le t$

$$\mathrm{E}[X_t|\mathscr{G}_s] = \mathrm{E}[X_t|X_s] \ .$$

d) Compute $\mathrm{E}[X_t|\mathscr{F}_s]$. Do the $\sigma$-algebras $\mathscr{F}_s$ and $\mathscr{G}_s$ coincide?
e) Compute the finite-dimensional distributions of $(X_t)_t$ $(0 \le t \le 1)$ and show that they coincide with the finite-dimensional distributions of $(B_t)_t$ conditioned given $B_1 = 0$.

# Chapter 5
# Martingales

Martingales are stochastic processes that enjoy many important, sometimes surprising, properties. When studying a process $X$, it is always a good idea to look for martingales "associated" to $X$, in order to take advantage of these properties.

## 5.1 Definitions and general facts

Let $T \subset \mathbb{R}^+$.

**Definition 5.1** A real-valued process $M = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in T}, (M_t)_{t\in T}, \mathrm{P})$ is a *martingale* (resp. a *supermartingale*, a *submartingale*) if $M_t$ is integrable for every $t \in T$ and

$$\mathrm{E}(M_t|\mathscr{F}_s) = M_s \qquad (\text{resp. } \le M_s, \ge M_s) \tag{5.1}$$

for every $s \le t$.

When the filtration is not specified it is understood to be the natural one.

*Examples 5.1*

a) If $T = \mathbb{N}$ and $(X_k)_k$ is a sequence of independent real centered r.v.'s, and $Y_n = X_1 + \dots + X_n$, then $(Y_n)_n$ is a martingale.

Indeed let $\mathscr{F}_m = \sigma(Y_1, \dots, Y_m)$ and also observe that $\mathscr{F}_m = \sigma(X_1, \dots, X_m)$. If $n > m$, as we can write $Y_n = Y_m + X_{m+1} + \dots + X_n$

(continued)

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_5

*Examples 5.1* (continued)

and, as the r.v.'s $X_{m+1}, \dots, X_n$ are centered and independent of $\mathscr{F}_m$, we have

$$\begin{aligned} E(Y_n|\mathscr{F}_m) &= E(Y_m|\mathscr{F}_m) + E(X_n + \cdots + X_{m+1}|\mathscr{F}_m) \\ &= Y_m + E(X_n + \cdots + X_{m+1}) = Y_m \,. \end{aligned}$$

b) Let $X$ be an integrable r.v. and $(\mathscr{F}_n)_n$ a filtration, then $X_n = E(X|\mathscr{F}_n)$ is a martingale. Indeed if $n > m$, then $\mathscr{F}_m \subset \mathscr{F}_n$ and

$$E(X_n|\mathscr{F}_m) = E\big[E(X|\mathscr{F}_n)\,\big|\,\mathscr{F}_m\big] = E(X|\mathscr{F}_m) = X_m$$

thanks to Proposition 4.1 e).

c) If $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, P)$ is a Brownian motion, then $(B_t)_t$ is a $(\mathscr{F}_t)_t$-martingale, as we saw in Example 4.3.

It is clear that linear combinations of martingales are also martingales and linear combinations with positive coefficients of supermartingales (resp. submartingales) are still supermartingales (resp. submartingales). If $(M_t)_t$ is a supermartingale, then $(-M_t)_t$ is a submartingale and vice versa.

Moreover, if $M$ is a martingale (resp. a submartingale) and $\Phi : \mathbb{R} \to \mathbb{R}$ is a convex (resp. increasing convex) function such that $\Phi(M_t)$ is also integrable for every $t$, then $(\Phi(M_t))_t$ is a submartingale: it is a consequence of Jensen's inequality, Proposition 4.2 d). Indeed, if $s \le t$,

$$E[\Phi(M_t)|\mathscr{F}_s] \ge \Phi\big(E[M_t|\mathscr{F}_s]\big) = \Phi(M_s) \,.$$

In particular, if $(M_t)_t$ is a martingale then $(|M_t|)_t$ is a submartingale (with respect to the same filtration).

We shall say that a martingale (resp. supermartingale, submartingale) $(M_t)_t$ is in $L^p, p \ge 1$, if $M_t \in L^p$ for every $t$. We shall speak of *square integrable* martingales (resp. supermatingales, submartingales) if $p = 2$. If $(M_t)_t$ is a martingale in $L^p$, $p \ge 1$, then $(|M_t|^p)_t$ is a submartingale.

Note that it is not true, in general, that if $M$ is a submartingale, then also $(|M_t|)_t$ and $(M_t^2)_t$ are submartingales (even if $M$ is square integrable): the functions $x \mapsto |x|$ and $x \mapsto x^2$ are convex but *not increasing*. If $M$ is square integrable $(M_t^2)_t$ will, however, be a submartingale under the additional assumption that $M$ is positive: the function $x \mapsto x^2$ is increasing when restricted to $\mathbb{R}^+$.

## 5.2 Discrete time martingales

In this and in the subsequent sections we assume $T = \mathbb{N}$.

A process $(A_n)_n$ is said to be *an increasing predictable process* for a filtration $(\mathscr{F}_n)_n$ if $A_0 = 0$, $A_n \le A_{n+1}$ a.s. and $A_{n+1}$ is $\mathscr{F}_n$-measurable for every $n \ge 0$.

As $\mathscr{F}_n \subset \mathscr{F}_{n+1}$, $(A_n)_n$ is, in particular, adapted but, intuitively, it is even a process such that at time $n$ we know its value at time $n+1$ (which accounts for the term "predictable"). Let $(X_n)_n$ be an $(\mathscr{F}_n)_n$-submartingale and $(A_n)_n$ the process defined recursively as

$$A_0 = 0, \qquad A_{n+1} = A_n + \underbrace{\mathrm{E}(X_{n+1}\,|\,\mathscr{F}_n) - X_n}_{\ge 0} \,. \tag{5.2}$$

$(A_n)_n$ is clearly increasing and, as $A_{n+1}$ is the sum of $\mathscr{F}_n$-measurable r.v.'s, by construction it is an increasing predictable process. As $A_{n+1}$ is $\mathscr{F}_n$-measurable, (5.2) can be rewritten as

$$\mathrm{E}(X_{n+1} - A_{n+1}\,|\,\mathscr{F}_n) = \mathrm{E}(X_{n+1}\,|\,\mathscr{F}_n) - A_{n+1} = X_n - A_n \,,$$

i.e. $M_n = X_n - A_n$ is an $(\mathscr{F}_n)_n$-martingale, that is we have decomposed $X$ into the sum of a martingale and of an increasing predictable process.

If $X_n = M'_n + A'_n$ were another decomposition of $(X_n)_n$ into the sum of a martingale $M'$ and of an increasing predictable process $A'$, we would have

$$A'_{n+1} - A'_n = X_{n+1} - X_n - (M'_{n+1} - M'_n) \,.$$

By conditioning with respect to $\mathscr{F}_n$, we have $A'_{n+1} - A'_n = \mathrm{E}(X_{n+1}\,|\,\mathscr{F}_n) - X_n = A_{n+1} - A_n$; as $A_0 = A'_0 = 0$, by recurrence it follows that $A'_n = A_n$ and $M'_n = M_n$ . We have therefore proved that

**Theorem 5.1** Every submartingale $(X_n)_n$ can be decomposed uniquely into the sum of a martingale $(M_n)_n$ and an increasing predictable process $(A_n)_n$.

This is *Doob's decomposition* and the process $A$ is the *compensator* of $(X_n)_n$. If $(M_n)_n$ is a square integrable martingale, then $(M_n^2)_n$ is a submartingale. Its compensator is called the *associated increasing process* of $(M_n)_n$.

If $(X_n)_n$ is a martingale, the same is true for the stopped process $X^\tau_n = X_{n\wedge\tau}$, where $\tau$ is a stopping time of the filtration $(\mathscr{F}_n)_n$. Indeed as $X^\tau_{n+1} = X^\tau_n$ on $\{\tau \le n\}$,

$$X^\tau_{n+1} - X^\tau_n = (X^\tau_{n+1} - X^\tau_n)1_{\{\tau\le n\}} + (X^\tau_{n+1} - X^\tau_n)1_{\{\tau\ge n+1\}} = (X^\tau_{n+1} - X^\tau_n)1_{\{\tau\ge n+1\}}$$

Hence, as by the definition of a stopping time $\{\tau \ge n+1\} = \{\tau \le n\}^c \in \mathscr{F}_n$, we have

$$\begin{aligned} \mathrm{E}(X^\tau_{n+1} - X^\tau_n\,|\,\mathscr{F}_n) &= \mathrm{E}[(X_{n+1} - X_n)1_{\{\tau\ge n+1\}}\,|\,\mathscr{F}_n] \\ &= 1_{\{\tau\ge n+1\}}\mathrm{E}(X_{n+1} - X_n\,|\,\mathscr{F}_n) = 0\ . \end{aligned} \tag{5.3}$$

Similarly it is proved that if $(X_n)_n$ is a supermartingale (resp. a submartingale) then the stopped process $(X^\tau_n)_n$ is again a supermartingale (resp. a submartingale).

The following is the key result from which many properties of martingales follow.

**Theorem 5.2 (Stopping theorem)** Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_n, (X_n)_n, \mathrm{P})$ be a supermartingale and $\tau_1, \tau_2$ two *a.s. bounded* stopping times of the filtration $(\mathscr{F}_n)_n$ such that $\tau_1 \le \tau_2$ a.s. Then the r.v.'s $X_{\tau_1}, X_{\tau_2}$ are integrable and

$$\mathrm{E}(X_{\tau_2}\,|\,\mathscr{F}_{\tau_1}) \le X_{\tau_1}\ .$$

*Proof* Integrability of $X_{\tau_1}$ and $X_{\tau_2}$ is immediate, as, for $i = 1, 2$ and denoting by $k$ a number that majorizes $\tau_2$, $|X_{\tau_i}| \le \sum_{j=1}^k |X_j|$. Let us assume, at first, that $\tau_2 \equiv k \in \mathbb{N}$ and let $A \in \mathscr{F}_{\tau_1}$. As $A \cap \{\tau_1 = j\} \in \mathscr{F}_j$, we have for $j \le k$

$$\mathrm{E}[X_{\tau_1}1_{A\cap\{\tau_1=j\}}] = \mathrm{E}[X_j 1_{A\cap\{\tau_1=j\}}] \ge \mathrm{E}[X_k 1_{A\cap\{\tau_1=j\}}]$$

and, taking the sum over $j$, $0 \le j \le k$,

$$\mathrm{E}[X_{\tau_1}1_A] = \sum_{j=0}^k \mathrm{E}[X_j 1_{A\cap\{\tau_1=j\}}] \ge \sum_{j=0}^k \mathrm{E}[X_k 1_{A\cap\{\tau_1=j\}}] = \mathrm{E}[X_{\tau_1}1_A]\ .$$

We have therefore proved the statement if $\tau_2$ is a constant stopping time. Let us now remove this hypothesis and assume $\tau_2 \le k$. If we apply the result proved in the first part of the proof to the stopped martingale $(X^{\tau_2}_n)_n$ and to the stopping times $\tau_1$ and $k$ we obtain

$$\mathrm{E}[X_{\tau_1}1_A] = \mathrm{E}[X^{\tau_2}_{\tau_1}1_A] \ge \mathrm{E}[X^{\tau_2}_k 1_A] = \mathrm{E}[X_{\tau_2}1_A]\ ,$$

which allows us to conclude the proof.

□

Thanks to Theorem 5.2 applied to $X$ and $-X$ we have, of course,

**Corollary 5.1** If $X$ is a martingale and $\tau_1, \tau_2$ are bounded stopping times such that $\tau_1 \le \tau_2$, then

$$\mathrm{E}(X_{\tau_2} \,|\, \mathscr{F}_{\tau_1}) = X_{\tau_1} \text{ a.s.}$$

Let us point out that the assumption of boundedness of the stopping times in the two previous statements is essential and that it is easy to find counterexamples showing that the stopping theorem does not hold under the weaker assumption that $\tau_1$ and $\tau_2$ are only finite.

*Remark 5.1* In particular, if $X$ is a martingale, Corollary 5.1 applied to $\tau_2 = \tau$ and $\tau_1 = 0$ gives

$$\mathrm{E}[X_\tau] = \mathrm{E}[\mathrm{E}(X_\tau \,|\, \mathscr{F}_0)] = \mathrm{E}[X_0]$$

for every bounded stopping time $\tau$. Hence the quantity $\mathrm{E}(X_\tau)$ is constant as $\tau$ ranges among the bounded stopping times. In fact this condition is also sufficient for $X$ to be a martingale (Exercise 5.6).

**Theorem 5.3 (Maximal inequalities)** Let $X$ be a supermartingale and $\lambda > 0$; then

$$\lambda \mathrm{P}\Big(\sup_{0 \le i \le k} X_i \ge \lambda\Big) \le \mathrm{E}(X_0) + \mathrm{E}(X_k^-)\,, \tag{5.4}$$

$$\lambda \mathrm{P}\Big(\inf_{0 \le i \le k} X_i \le -\lambda\Big) \le -\mathrm{E}[X_k 1_{\{\inf_{0 \le i \le k} X_i \le -\lambda\}}] \le \mathrm{E}(|X_k|)\,. \tag{5.5}$$

*Proof* Let

$$\tau(\omega) = \inf\{n; n \le k, X_n(\omega) \ge \lambda\}$$

with the understanding that $\tau = k$ if $\{\ \} = \emptyset$. $\tau$ is a finite stopping time and we have $X_\tau \geq \lambda$ on the event $\{\sup_{0\leq i\leq k} X_i \geq \lambda\}$. By Theorem 5.2 applied to $\tau_2 = \tau$, $\tau_1 = 0$,

$$\begin{aligned} \mathrm{E}(X_0) \geq \mathrm{E}(X_\tau) &= \mathrm{E}[X_\tau 1_{\{\sup_{0\leq i\leq k} X_i \geq \lambda\}}] + \mathrm{E}[X_k 1_{\{\sup_{0\leq i\leq k} X_i < \lambda\}}] \\ &\geq \lambda \mathrm{P}\Big(\sup_{1\leq i\leq k} X_i \geq \lambda\Big) - \mathrm{E}(X_k^-)\ , \end{aligned}$$

from which we obtain (5.4). If conversely

$$\tau(\omega) = \inf\{n; n \leq k, X_n(\omega) \leq -\lambda\}$$

and $\tau(\omega) = k$ if $\{\ \} = \emptyset$, then $\tau$ is still a bounded stopping time and, again by Theorem 5.2,

$$\begin{aligned} \mathrm{E}(X_k) \leq \mathrm{E}(X_\tau) = \mathrm{E}[X_\tau 1_{\{\inf_{0\leq i\leq k} X_i \leq -\lambda\}}] + \mathrm{E}[X_k 1_{\{\inf_{0\leq i\leq k} X_i > -\lambda\}}] \leq \\ \leq -\lambda \mathrm{P}\Big(\inf_{0\leq i\leq k} X_i \leq -\lambda\Big) + \mathrm{E}[X_k 1_{\{\inf_{0\leq i\leq k} X_n > -\lambda\}}] \end{aligned}$$

from which we obtain (5.5).

□

## 5.3 Discrete time martingales: a.s. convergence

One of the reasons why martingales are important is the result of this section; it guarantees, under rather weak (and easy to check) hypotheses, that a martingale converges a.s.

Let $a < b$ be real numbers. We say that $(X_n(\omega))_n$ makes an *upcrossing* of the interval $[a, b]$ in the time interval $[i, j]$ if $X_i(\omega) < a$, $X_j(\omega) > b$ and $X_m(\omega) \leq b$ for $m = i + 1, \dots, j - 1$. Let

$$\gamma_{a,b}^k(\omega) = \text{ number of upcrossings of } (X_n(\omega))_{n\leq k} \text{ over the interval } [a, b]\ .$$

The proof of the theorem of convergence that we have advertised is a bit technical, but the basic idea is rather simple: in order to prove that a sequence is convergent one first needs to prove that it does not oscillate too much. For this reason the key estimate is the following, which states that a supermartingale cannot make too many upcrossings.

**Proposition 5.1** If $X$ is a supermartingale, then

$$(b-a)\mathrm{E}(\gamma^k_{a,b}) \le \mathrm{E}[(X_k-a)^-] .$$

*Proof* Let us define the following sequence of stopping times

$$\begin{aligned}
\tau_1(\omega) &= \inf\{i; i \le k, X_i(\omega) < a\}\\
\tau_2(\omega) &= \inf\{i; \tau_1(\omega) < i \le k, X_i(\omega) > b\}\\
&\dots\\
\tau_{2m-1}(\omega) &= \inf\{i; \tau_{2m-2}(\omega) < i \le k, X_i(\omega) < a\}\\
\tau_{2m}(\omega) &= \inf\{i; \tau_{2m-1}(\omega) < i \le k, X_i(\omega) > b\}
\end{aligned}$$

with the understanding that $\tau_i = k$ if $\{\,\} = \emptyset$. Let

$$\begin{aligned}
\Omega_{2m} &= \{\tau_{2m} \le k, X_{\tau_{2m}} > b\} = \{\gamma^k_{a,b} \ge m\} ,\\
\Omega_{2m-1} &= \{\gamma^k_{a,b} \ge m-1, X_{\tau_{2m-1}} < a\} .
\end{aligned}$$

It is immediate that $\Omega_i \in \mathscr{F}_{\tau_i}$, as $\tau_i$ and $X_{\tau_i}$ are $\mathscr{F}_{\tau_i}$-measurable. Moreover, $X_{\tau_{2m-1}} < a$ on $\Omega_{2m-1}$ and $X_{\tau_{2m}} > b$ on $\Omega_{2m}$. By Theorem 5.2, applied to the bounded stopping times $\tau_{2m-1}$ and $\tau_{2m}$,

$$\begin{aligned}
0 \ge \int_{\Omega_{2m-1}} (X_{\tau_{2m-1}} - a)\, d\mathrm{P} &\ge \int_{\Omega_{2m-1}} (X_{\tau_{2m}} - a)\, d\mathrm{P}\\
&\ge (b-a)\,\mathrm{P}(\Omega_{2m}) + \int_{\Omega_{2m-1}\setminus\Omega_{2m}} (X_k - a)\, d\mathrm{P} ,
\end{aligned}$$

where we take advantage of the fact that $\tau_{2m} = k$ on $\Omega_{2m-1} \setminus \Omega_{2m}$ and that $X_{\tau_{2m}} > b$ on $\Omega_{2m}$. Therefore

$$(b-a)\mathrm{P}(\Omega_{2m}) = (b-a)\mathrm{P}(\gamma^k_{a,b} \ge m) \le \int_{\Omega_{2m-1}\setminus\Omega_{2m}} (X_k - a)^- d\mathrm{P} . \tag{5.6}$$

Let us recall that (Exercise 1.3)

$$\mathrm{E}(\gamma^k_{a,b}) = \sum_{m=1}^{\infty} \mathrm{P}(\gamma^k_{a,b} \ge m) .$$

As the events $\Omega_{2m-1} \setminus \Omega_{2m}$ are pairwise disjoint, taking the sum in $m$ in (5.6) we have

$$(b-a)\mathrm{E}(\gamma^k_{a,b}) = (b-a)\sum_{m=1}^{\infty} \mathrm{P}(\gamma^k_{a,b} \geq m) \leq \mathrm{E}[(X_k-a)^-]\ .$$

□

**Theorem 5.4** Let $X$ be a supermartingale such that

$$\sup_{n\geq 0} \mathrm{E}(X_n^-) < +\infty\ . \tag{5.7}$$

Then it converges a.s. to a finite limit.

*Proof* For fixed $a < b$ let us denote by $\gamma_{a,b}(\omega)$ the number of upcrossings of the path $(X_n(\omega))_n$ over the interval $[a,b]$. As $(X_n-a)^- \leq a^+ + X_n^-$, by Proposition 5.1,

$$\begin{aligned}\mathrm{E}(\gamma_{a,b}) = \lim_{k\to\infty} \mathrm{E}(\gamma^k_{a,b}) &\leq \frac{1}{b-a} \sup_{n\geq 0} \mathrm{E}[(X_n-a)^-] \\ &\leq \frac{1}{b-a}\Big(a^+ + \sup_{n\geq 0} \mathrm{E}(X_n^-)\Big) < +\infty\ .\end{aligned} \tag{5.8}$$

In particular, $\gamma_{a,b} < +\infty$ a.s., i.e. $\gamma_{a,b} < +\infty$ outside a negligible event $N_{a,b}$; considering the union of these negligible events $N_{a,b}$ for every possible $a, b \in \mathbb{Q}$ with $a < b$, we can assume that outside a negligible event $N$ we have $\gamma_{a,b} < +\infty$ for every $a, b \in \mathbb{R}$.

Let us prove that if $\omega \notin N$ then the sequence $(X_n(\omega))_n$ necessarily has a limit. Indeed, if this was not the case, let $a = \underline{\lim}_{n\to\infty} X_n(\omega) < \overline{\lim}_{n\to\infty} X_n(\omega) = b$. This implies that $(X_n(\omega))_n$ is close to both $a$ and $b$ infinitely many times. Therefore if $\alpha, \beta$ are such that $a < \alpha < \beta < b$ we would have necessarily $\gamma_{\alpha,\beta}(\omega) = +\infty$, which is not possible outside $N$.

The limit is, moreover, finite. In fact from (5.8)

$$\lim_{b\to+\infty} \mathrm{E}(\gamma_{a,b}) = 0$$

but $\gamma_{a,b}(\omega)$ is a non-increasing function of $b$ and therefore

$$\lim_{b\to+\infty} \gamma_{a,b}(\omega) = 0 \qquad \text{a.s.}$$

As $\gamma_{a,b}$ takes only integer values, $\gamma_{a,b}(\omega) = 0$ for large $b$ and $(X_n(\omega))_n$ is therefore bounded above a.s. Similarly one sees that it is bounded below.

□

In particular,

**Corollary 5.2** Let $X$ be a positive supermartingale. Then the limit

$$\lim_{n\to\infty} X_n$$

exists a.s. and is finite.

*Example 5.1* Let $(Z_n)_n$ be a sequence of i.i.d. r.v.'s taking the values $\pm 1$ with probability $\frac{1}{2}$ and let $X_0 = 0$ and $X_n = Z_1 + \cdots + Z_n$ for $n \geq 1$. Let $a, b$ be positive integers and let $\tau = \inf\{n; X_n \geq b \text{ or } X_n \leq -a\}$, the exit time of $X$ from the interval $]-a, b[$. Is $\tau < +\infty$ with probability 1? In this case we can define the r.v. $X_\tau$, which is the position of $X$ when it leaves the interval $]-a, b[$. Of course, $X_\tau$ can only take the values $-a$ or $b$. What is the value of $\mathrm{P}(X_\tau = b)$?

We know (as in Example 5.1 a)) that $X$ is a martingale. Also $(X_{n\wedge\tau})_n$ is a martingale, which is moreover bounded as it can take only values that are $\geq -a$ and $\leq b$.

By Theorem 5.4 the limit $\lim_{n\to\infty} X_{n\wedge\tau}$ exists and is finite. This implies that $\tau < +\infty$ a.s.: as at every iteration $X$ makes steps of size 1 to the right or to the left, on $\tau = +\infty$ we have $|X_{(n+1)\wedge\tau} - X_{n\wedge\tau}| = 1$, so that $(X_{n\wedge\tau})_n$ cannot be a Cauchy sequence.

Therefore necessarily $\tau < +\infty$ and the r.v. $X_\tau$ is well defined. In order to compute $\mathrm{P}(X_\tau = b)$, let us assume for a moment that we can apply Theorem 5.2, the stopping theorem, to the stopping times $\tau_2 = \tau$ and $\tau_1 = 0$ (we cannot because $\tau$ is finite but not bounded), then we would have

$$0 = \mathrm{E}[X_0] = \mathrm{E}[X_\tau] \,. \tag{5.9}$$

From this relation, as $\mathrm{P}(X_\tau = -a) = 1 - \mathrm{P}(X_\tau = b)$, we obtain

$$0 = \mathrm{E}[X_\tau] = b\mathrm{P}(X_\tau = b) - a\mathrm{P}(X_\tau = -a) = b\mathrm{P}(X_\tau = b) - a(1 - \mathrm{P}(X_\tau = b))$$

i.e.

$$\mathrm{P}(X_\tau = b) = \frac{a}{a+b} \cdot$$

(continued)

*Example 5.1* (continued)
The problem is therefore solved if (5.9) is satisfied. Let us prove it: for every $n$ the stopping time $\tau \wedge n$ is bounded, therefore the stopping theorem gives

$$0 = \mathrm{E}[X_{\tau\wedge n}] \ .$$

Now observe that $\lim_{n\to\infty} X_{\tau\wedge n} = X_\tau$ and that, as $-a \le X_{\tau\wedge n} \le b$, the sequence $(X_{\tau\wedge n})_n$ is bounded, so that we can apply Lebesgue's theorem and obtain (5.9).

This example shows a typical application of the stopping theorem in order to obtain the distribution of a process stopped at some stopping time, and also how to apply the stopping theorem to stopping times that are not bounded: just apply the stopping theorem to the stopping times $\tau \wedge n$, that are bounded, and then pass to the limit using Lebesgue's theorem as in this example or some other statement, such as Beppo Levi's theorem, in order to pass to the limit.

*Remark 5.2* Condition (5.7) for a supermartingale is equivalent to requiring that $X$ is bounded in $L^1$. Indeed, obviously,

$$X_k^- \le |X_k|$$

so that, if $\sup_k \mathrm{E}[|X_k|] < +\infty$, then also (5.7) holds. Conversely, note that

$$|X_k| = X_k + 2X_k^- \ ,$$

hence

$$\mathrm{E}[|X_k|] = \mathrm{E}[X_k] + 2\mathrm{E}[X_k^-] \le \mathrm{E}[X_0] + 2\mathrm{E}[X_k^-] \ ,$$

so that (5.7) implies boundedness in $L^1$.

## 5.4 Doob's inequality; $L^p$ convergence, the $p > 1$ case

We say that a martingale $M$ is *bounded in* $L^p$ if

$$\sup_n \mathrm{E}\big[|M_n|^p\big] < +\infty \ .$$

**Theorem 5.5 (Doob's inequality)** Let $M = (\Omega, \mathscr{F}, (\mathscr{F}_n)_n, (M_n)_n, \mathrm{P})$ be a martingale bounded in $L^p$ with $p > 1$. Then if $M^* = \sup_n |M_n|$, $M^* \in L^p$ and

$$\|M^*\|_p \le q \sup_n \|M_n\|_p \ , \tag{5.10}$$

where $q = \frac{p}{p-1}$ is the exponent conjugate to $p$.

Theorem 5.5 is a consequence of the following result, applied to the positive submartingale $(|M_n|)_n$.

**Lemma 5.1** If $X$ is a positive submartingale, then for every $p > 1$ and $n \in \mathbb{N}$

$$\mathrm{E}\Big(\max_{0\le i\le n} X_i^p\Big) \le \Big(\frac{p}{p-1}\Big)^p \mathrm{E}(X_n^p) \ .$$

*Proof* Let $Y = \max_{1\le i\le n} X_i$ and, for $\lambda > 0$,

$$\tau_\lambda(\omega) = \inf\{i; 0 \le i \le n, X_i(\omega) > \lambda\}$$

and $\tau_\lambda(\omega) = n + 1$ if $\{\ \} = \emptyset$. Then $\sum_{k=1}^n 1_{\{\tau_\lambda=k\}}) = 1_{\{Y>\lambda\}}$ and, for every $\alpha > 0$,

$$\begin{aligned} Y^p &= p \int_0^Y \lambda^{p-1}\, d\lambda = p \int_0^{+\infty} \lambda^{p-1} 1_{\{Y>\lambda\}}\, d\lambda \\ &= p \int_0^{+\infty} \lambda^{p-1} \sum_{k=1}^n 1_{\{\tau_\lambda=k\}}\, d\lambda \ . \end{aligned} \tag{5.11}$$

Moreover, if $k \le n$, as $X_k \ge \lambda$ on $\{\tau_\lambda = k\}$, we have $\lambda 1_{\{\tau_\lambda=k\}} \le X_k 1_{\{\tau_\lambda=k\}}$ and

$$\lambda^{p-1} \sum_{k=1}^n 1_{\{\tau_\lambda=k\}} \le \lambda^{p-2} \sum_{k=1}^n X_k 1_{\{\tau_\lambda=k\}} \ . \tag{5.12}$$

As $X$ is a submartingale and $1_{\{\tau_\lambda=k\}}$ is $\mathscr{F}_k$-measurable,

$$\mathrm{E}(X_n 1_{\{\tau_\lambda=k\}} | \mathscr{F}_k) = 1_{\{\tau_\lambda=k\}} \mathrm{E}(X_n | \mathscr{F}_k) \ge X_k 1_{\{\tau_\lambda=k\}}$$

and, combining this relation with (5.12), we find

$$\mathrm{E}\Big[\lambda^{p-1} \sum_{k=1}^n 1_{\{\tau_\lambda=k\}}\Big] \le \mathrm{E}\Big[\lambda^{p-2} \sum_{k=1}^n 1_{\{\tau_\lambda=k\}} X_k\Big] \le \mathrm{E}\Big[\lambda^{\alpha-2} X_n \sum_{k=1}^n 1_{\{\tau_\lambda=k\}}\Big] \ .$$

Thanks to this inequality, if we take the expectation of both members in (5.11),

$$\frac{1}{p}\,\mathrm{E}[Y^p] \le \mathrm{E}\Big[X_n \underbrace{\int_0^{+\infty} \lambda^{p-2} \sum_{k=1}^n 1_{\{\tau_\lambda = k\}}\, d\lambda}_{=\frac{1}{p-1}\,Y^{p-1}}\Big] = \frac{1}{p-1}\,\mathrm{E}[X_n Y^{p-1}]\,.$$

Now by Hölder's inequality (recall that we assume $p > 1$)

$$\mathrm{E}[Y^p] \le \frac{p}{p-1}\,\mathrm{E}[Y^p]^{(p-1)/p}\,\mathrm{E}[X_n^p]^{1/p}\,.$$

As we know that $\mathrm{E}(Y^p) < +\infty$ ($Y$ is the maximum of a finite number of r.v.'s of $L^p$), we can divide both terms by $\mathrm{E}[Y^p]^{(p-1)/p}$, which gives

$$\mathrm{E}[Y^p]^{1/p} \le \frac{p}{p-1}\,\mathrm{E}[X_n^p]^{1/p}\,,$$

from which the statement follows. □

*Proof of Theorem 5.5* Lemma 5.1 applied to the submartingale $X_i = |M_i|$ gives the inequality

$$\Big\| \max_{0\le i\le n} |M_i| \Big\|_p \le q\|M_n\|_p = q \max_{k\le n} \|M_k\|_p\,,$$

where we used the fact that $k \mapsto \|M_k\|_p$ is increasing as $(|M_n|^p)_n$ is a submartingale. Doob's inequality now follows by taking the limit as $n \to \infty$ and using Beppo Levi's theorem. □

Thanks to Doob's inequality (5.10), for $p > 1$, the behavior of a martingale bounded in $L^p$ is very nice and simple.

**Theorem 5.6** If $p > 1$ a martingale is bounded in $L^p$ if and only if it converges a.s. and in $L^p$.

*Proof* It is immediate that a martingale $M$ that converges in $L^p$ is also bounded in $L^p$.

Conversely, if a martingale $M$ is bounded in $L^p$ with $p > 1$, then

$$\sup_{n\ge 0} M_n^- \le \sup_{n\ge 0} |M_n| = M^*\,.$$

Since by Doob's inequality $M^* \in L^p$, we have $\sup_n \mathrm{E}(M_n^-) \le \mathrm{E}(M^*) < +\infty$, so that the condition (5.7) of Theorem 5.4 is satisfied and $M$ converges a.s. Moreover, its limit $M_\infty$ belongs to $L^p$, as obviously $|M_\infty| \le M^*$.

Let us prove that the convergence takes place in $L^p$, i.e. that

$$\lim_{n\to\infty} \mathrm{E}\big[|M_n - M_\infty|^p\big] = 0\ .$$

We already know that $|M_n - M_\infty|^p \to_{n\to\infty} 0$ so that we only need to find a bound in order to apply Lebesgue's theorem. Thanks to the inequality $|x+y|^p \le 2^{p-1}(|x|^p + |y|^p)$, which holds for every $x, y \in \mathbb{R}$,

$$|M_n - M_\infty|^p \le 2^{p-1}(|M_n|^p + |M_\infty|^p) \le 2^p M^{*p}$$

and as the r.v. $M^{*p}$ is integrable by Doob's maximal inequality, Lebesgue's theorem can be applied, giving

$$\lim_{n\to\infty} \mathrm{E}[|M_n - M_\infty|^p] = 0\ .$$

□

We see in the next section that for $L^1$-convergence of martingales things are very different.

## 5.5 Uniform integrability and convergence in $L^1$

The notion of uniform integrability is the key tool for the investigation of $L^1$ convergence of martingales.

**Definition 5.2** A family of $m$-dimensional r.v.'s $\mathscr{H}$ is said to be *uniformly integrable* if

$$\lim_{c\to+\infty} \sup_{Y\in\mathscr{H}} \mathrm{E}\big[|Y|1_{\{|Y|>c\}}\big] = 0\ . \tag{5.13}$$

The set formed by a single integrable r.v. is the simplest example of a uniformly integrable family. We have $\lim_{c\to+\infty} |Y|1_{\{|Y|>c\}} = 0$ a.s. and, as $|Y|1_{\{|Y|>c\}} \le |Y|$, by Lebesgue's theorem,

$$\lim_{c\to+\infty} \mathrm{E}\big[|Y|1_{\{|Y|>c\}}\big] = 0\ .$$

Similarly $\mathscr{H}$ turns out to be uniformly integrable if there exists an integrable real r.v. $Z$ such that $Z \geq |Y|$ for every $Y \in \mathscr{H}$. Actually, in this case, for every $Y \in \mathscr{H}$, $|Y|1_{\{|Y|>c\}} \leq |Z|1_{\{|Z|>c\}}$ so that, for every $Y \in \mathscr{H}$,

$$\mathrm{E}\big[|Y|1_{\{|Y|>c\}}\big] \leq \mathrm{E}\big[Z1_{\{Z>c\}}\big] ,$$

from which (5.13) follows. Hence the following theorem is a extension of Lebesgue's theorem. For a proof, as well as a proof of the following Proposition 5.3, see J. Neveu's book (Neveu 1964, §II-5).

**Theorem 5.7** Let $(Y_n)_n$ be a sequence of r.v.'s converging a.s. to $Y$. In order for $Y$ to be integrable and for the convergence to take place in $L^1$ it is necessary and sufficient that $(Y_n)_n$ is uniformly integrable.

In any case a uniformly integrable family $\mathscr{H}$ is bounded in $L^1$: let $c > 0$ be such that

$$\sup_{Y\in\mathscr{H}} \mathrm{E}\big[|Y|1_{\{|Y|>c\}}\big] \leq 1 ,$$

then we have, for every $Y \in \mathscr{H}$,

$$\mathrm{E}(|Y|) = \mathrm{E}\big[|Y|1_{\{|Y|>c\}}\big] + \mathrm{E}\big[|Y|1_{\{|Y|\leq c\}}\big] \leq 1 + c .$$

The next two characterizations of uniform integrability are going to be useful.

**Proposition 5.2** A family $\mathscr{H}$ is uniformly integrable if and only if the following conditions are satisfied.

a) $\sup_{Y\in\mathscr{H}} \mathrm{E}[|Y|] < +\infty$.
b) For every $\varepsilon > 0$ there exists a $\delta_\varepsilon > 0$ such that, for every $Y \in \mathscr{H}$,

$$\mathrm{E}[|Y|1_A] \leq \varepsilon$$

for every $A \in \mathscr{F}$ such that $\mathrm{P}(A) \leq \delta_\varepsilon$.

The proof is left as an exercise or, again, see J. Neveu's book (Neveu 1964, Proposition II-5-2).

**Proposition 5.3** A family $\mathscr{H} \subset L^1$ is uniformly integrable if and only if there exists a positive increasing convex function $g : \mathbb{R}^+ \to \mathbb{R}$ such that $\lim_{t\to+\infty} \frac{g(t)}{t} = +\infty$ and

$$\sup_{Y\in\mathscr{H}} \mathrm{E}[g(|Y|)] < +\infty \, .$$

In particular, if $\mathscr{H}$ is bounded in $L^p$ for some $p > 1$ then it is uniformly integrable: just apply Proposition 5.3 with $g(t) = t^p$.

**Proposition 5.4** Let $Y$ be an integrable r.v. Then the family $\{\mathrm{E}(Y|\mathscr{G})\}_{\mathscr{G}}$, for $\mathscr{G}$ in the class of all sub-$\sigma$-algebras of $\mathscr{F}$, is uniformly integrable.

*Proof* As the family formed by the single r.v. $Y$ is uniformly integrable, by Proposition 5.3 there exists a positive increasing convex function $g : \mathbb{R}^+ \to \mathbb{R}$ such that $\lim_{t\to+\infty} \frac{g(t)}{t} = +\infty$ and $\mathrm{E}[g(|Y|)] < +\infty$. By Jensen's inequality, Proposition 4.2 d),

$$\mathrm{E}\big[g(|\mathrm{E}(Y|\mathscr{G})|)\big] \le \mathrm{E}\big[\mathrm{E}(g(|Y|)\,|\mathscr{G})\big] = \mathrm{E}[g(|Y|)] < +\infty$$

and, again by Proposition 5.3, $\{\mathrm{E}(Y|\mathscr{G})\}_{\mathscr{G}}$ is uniformly integrable.

□

In particular, if $(\mathscr{F}_n)_n$ is a filtration on $(\Omega, \mathscr{F}, \mathrm{P})$ and $Y \in L^1$, $(\mathrm{E}(Y|\mathscr{F}_n))_n$ is a uniformly integrable martingale (recall Example 5.1 b)).

**Theorem 5.8** Let $(M_n)_n$ be a martingale. Then the following properties are equivalent

a) $(M_n)_n$ converges in $L^1$;
b) $(M_n)_n$ is uniformly integrable;
c) $(M_n)_n$ is of the form $M_n = \mathrm{E}(Y|\mathscr{F}_n)$ for some $Y \in L^1(\Omega, \mathscr{F}, \mathrm{P})$.

If any of these conditions is satisfied then $(M_n)_n$ also converges a.s.

*Proof* If $(M_n)_n$ is a uniformly integrable martingale then it is bounded in $L^1$ and therefore the condition (5.7), $\sup_k \mathrm{E}(M_k^-) < +\infty$, is satisfied and $M$ converges a.s. to some r.v. $Y$. By Theorem 5.7, $Y \in L^1$ and the convergence takes place in $L^1$. Therefore (Remark 4.3), for every $m$,

$$M_m = \mathrm{E}(M_n|\mathscr{F}_m) \underset{n\to\infty}{\to} \mathrm{E}(Y|\mathscr{F}_m)$$

in $L^1$. This proves that b)⇒a) and a)⇒c). c)⇒b) is Proposition 5.4.

□

The following proposition identifies the limit of a uniformly integrable martingale.

**Proposition 5.5** Let $Y$ be an integrable r.v. on $L^1(\Omega, \mathscr{F}, \mathrm{P})$, $(\mathscr{F}_n)_n$ a filtration on $(\Omega, \mathscr{F})$ and $\mathscr{F}_\infty = \bigvee_n \mathscr{F}_n$, the $\sigma$-algebra generated by the $\sigma$-algebras $(\mathscr{F}_n)_n$. Then

$$\lim_{n\to\infty} \mathrm{E}(Y|\mathscr{F}_n) = \mathrm{E}(Y|\mathscr{F}_\infty) \qquad \text{a.s. and in } L^1 .$$

*Proof* Let $Z = \lim_{n\to\infty} \mathrm{E}(Y|\mathscr{F}_n)$. Then $Z$ is $\mathscr{F}_\infty$-measurable. Let us check that for every $A \in \mathscr{F}_\infty$

$$\mathrm{E}(Z1_A) = \mathrm{E}(Y1_A) . \tag{5.14}$$

By Remark 4.2 it is sufficient to prove this relation for $A$ in a class of events $\mathscr{C}$, generating $\mathscr{F}_\infty$, stable with respect to finite intersections and containing $\Omega$. If $A \in \mathscr{F}_m$ for some $m$ then

$$\mathrm{E}(Z1_A) = \lim_{n\to\infty} \mathrm{E}(\mathrm{E}(Y|\mathscr{F}_n)1_A) = \mathrm{E}(Y1_A)$$

because for $n \geq m$ $A$ is $\mathscr{F}_n$-measurable and $\mathrm{E}(\mathrm{E}(Y|\mathscr{F}_n)1_A) = \mathrm{E}(\mathrm{E}(1_A Y|\mathscr{F}_n)) = \mathrm{E}(Y1_A)$. (5.14) is therefore proved for $A \in \mathscr{C} = \bigcup_n \mathscr{F}_n$. As this class is stable with respect to finite intersections, generates $\mathscr{F}_\infty$ and contains $\Omega$, the proof is complete.

□

## 5.6 Continuous time martingales

We now extend the results of the previous sections to the continuous case, i.e. when the time set is $\mathbb{R}^+$ or an interval of $\mathbb{R}^+$, which is an assumption that we make from now on.

The main argument is that if $(M_t)_t$ is a supermartingale of the filtration $(\mathscr{F}_t)_t$, then, for every $t_0 < t_1 < \cdots < t_n$, $(M_{t_k})_{k=0,\dots,n}$ is a (discrete time) supermartingale of the filtration $(\mathscr{F}_{t_k})_{k=0,\dots,n}$ to which the results of the previous sections apply.

**Theorem 5.9** Let $M = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (M_t)_t, \mathrm{P})$ be a right (or left) continuous supermartingale. Then for every $T > 0$, $\lambda > 0$,

$$\lambda \mathrm{P}\Big(\sup_{0 \le t \le T} M_t \ge \lambda\Big) \le \mathrm{E}(M_0) + \mathrm{E}(M_T^-) , \tag{5.15}$$

$$\lambda \mathrm{P}\Big(\inf_{0 \le t \le T} M_t \le -\lambda\Big) \le -\mathrm{E}\big[M_T 1_{\{\inf_{0 \le t \le T} M_t \le -\lambda\}}\big] \le \mathrm{E}[|M_T|] . \tag{5.16}$$

*Proof* Let us prove (5.15). Let $0 = t_0 < t_1 < \cdots < t_n = T$. Then by (5.4) applied to the supermartingale $(M_{t_k})_{k=0,\dots,n}$

$$\lambda \mathrm{P}\Big(\sup_{0 \le k \le n} M_{t_k} \ge \lambda\Big) \le \mathrm{E}(M_0) + \mathrm{E}(M_T^-) .$$

Note that the right-hand side does not depend on the choice of $t_1, \dots, t_n$. Letting $\{t_0, \dots, t_n\}$ increase to $\mathbb{Q} \cap [0, T]$ we have that $\sup_{0 \le k \le n} M_{t_k}$ increases to $\sup_{\mathbb{Q} \cap [0,T]} M_t$ and Beppo Levi's theorem gives

$$\lambda \mathrm{P}\Big(\sup_{\mathbb{Q} \cap [0,T]} M_t \ge \lambda\Big) \le \mathrm{E}(M_0) + \mathrm{E}(M_T^-) .$$

The statement now follows because the paths are right (or left) continuous, so that

$$\sup_{\mathbb{Q} \cap [0,T]} M_t = \sup_{[0,T]} M_t .$$

□

The following statements can be proved with similar arguments.

**Theorem 5.10** Let $X$ be a right-continuous supermartingale. If $[a, b[ \subset \mathbb{R}^+$ ($b = +\infty$ possibly) and $\sup_{[a,b[} \mathrm{E}(X_t^-) < +\infty$, then $\lim_{t \to b-} X_t$ exists a.s.

**Theorem 5.11** A right-continuous martingale $M$ is uniformly integrable if and only if it is of the form $M_t = \mathrm{E}(Y | \mathscr{F}_t)$, where $Y \in L^1$ and if and only if it converges a.s. and in $L^1$.

*Example 5.2* Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a real Brownian motion and, for $\theta \in \mathbb{R}$, let

$$X_t = \mathrm{e}^{\theta B_t - \frac{\theta^2}{2} t} .$$

Then $X$ is a martingale.

Indeed, with the old trick of separating increment and actual position,

$$\begin{aligned} \mathrm{E}[X_t | \mathscr{F}_s] &= \mathrm{e}^{-\frac{\theta^2}{2} t} \mathrm{E}[\mathrm{e}^{\theta(B_t - B_s + B_s)} | \mathscr{F}_s] = \mathrm{e}^{-\frac{\theta^2}{2} t} \mathrm{e}^{\theta B_s} \mathrm{E}[\mathrm{e}^{\theta(B_t - B_s)} | \mathscr{F}_s] \\ &= \mathrm{e}^{-\frac{\theta^2}{2} t} \mathrm{e}^{\theta B_s} \mathrm{e}^{\frac{\theta^2}{2}(t-s)} = X_s . \end{aligned}$$

This is the prototype of an important class of martingales that we shall investigate later. Now observe that $X$ is a positive martingale and hence, by Theorem 5.10, the limit $\lim_{t \to +\infty} X_t$ exists a.s. and is finite. What is its value? Is this martingale uniformly integrable? These questions are the object of Exercise 5.9.

**Theorem 5.12 (Doob's inequality)** Let $M$ be a right-continuous martingale bounded in $L^p$, $p > 1$, and let $M^* = \sup_{t \geq 0} |M_t|$. Then $M^* \in L^p$ and

$$\|M^*\|_p \leq q \sup_{t \geq 0} \|M_t\|_p ,$$

where $q = \frac{p}{p-1}$.

**Theorem 5.13 (Stopping theorem)** Let $M$ be a right-continuous martingale (resp. supermartingale, submartingale) and $\tau_1, \tau_2$ two stopping times with $\tau_1 \leq \tau_2$. Then if $\tau_2$ is bounded a.s.

$$\mathrm{E}(M_{\tau_2} | \mathscr{F}_{\tau_1}) = M_{\tau_1} \qquad (\text{resp. } \leq, \geq) . \tag{5.17}$$

*Proof* Let us assume first that $M$ is a martingale. Let $b > 0$ and let $\tau$ be a stopping time taking only finitely many values $t_1 < t_2 < \cdots < t_m$ and bounded above by $b$. Then by Theorem 5.2, applied to the discrete time martingale $(M_{t_k})_{k=0,\dots,m}$ with respect to the filtration $(\mathscr{F}_{t_k})_{k=0,\dots,m}$, we have $\mathrm{E}[M_b | \mathscr{F}_\tau] = M_\tau$ and by Proposition 5.4 the r.v.'s $(M_\tau)_\tau$, for $\tau$ ranging over the set of stopping times taking only finitely many values bounded above by $b$, is a uniformly integrable family.

Let now $(\tau_n^1)_n$, $(\tau_n^2)_n$ be sequences of stopping times taking only finitely many values and decreasing to $\tau_1$ and $\tau_2$, respectively, and such that $\tau_n^2 \ge \tau_n^1$ for every $n$ as in Lemma 3.3. By Theorem 5.2, applied to the martingale $(M_{t_k})_{k=0,\dots,m}$, where $t_0 < t_1 < \dots < t_m$ are the possible values of the two stopping times $\tau_1$, $\tau_2$,

$$\mathrm{E}[M_{\tau_n^2} \,|\, \mathscr{F}_{\tau_n^1}] = M_{\tau_n^1} \ .$$

Conditioning both sides with respect to $\mathscr{F}_{\tau_1}$, as $\mathscr{F}_{\tau_1} \subset \mathscr{F}_{\tau_n^1}$,

$$\mathrm{E}[M_{\tau_n^2} \,|\, \mathscr{F}_{\tau_1}] = \mathrm{E}[M_{\tau_n^1} \,|\, \mathscr{F}_{\tau_1}] \ . \tag{5.18}$$

Since $n \to \infty$, $M_{\tau_n^1} \to M_{\tau_1}$ and $M_{\tau_n^2} \to M_{\tau_2}$ (the paths are right-continuous) and the convergence also takes place in $L^1$, $(M_{\tau_n^1})_n$, $(M_{\tau_n^2})_n$ being uniformly integrable. Therefore one can take the limit in (5.18) leading to (5.17).

Assume now that $M$ is a supermartingale. A repetition of the argument as above brings us to (5.18) with $=$ replaced by $\le$, i.e.

$$\mathrm{E}[M_{\tau_n^2} \,|\, \mathscr{F}_{\tau_1}] \le \mathrm{E}[M_{\tau_n^1} \,|\, \mathscr{F}_{\tau_1}] \ .$$

Now the fact that $(M_{\tau_n^1})_n$ and $(M_{\tau_n^2})_n$ are uniformly integrable is trickier and follows from the next Lemma 5.2 applied to $\mathscr{H}_n = \mathscr{F}_{\tau_n^i}$ and $Z_n = M_{\tau_n^i}$, observing that the stopping theorem for discrete time supermartingales ensures the relation $M_{\tau_{n+1}^i} \ge \mathrm{E}[M_{\tau_n^i} \,|\, \mathscr{F}_{\tau_{n+1}^i}]$. The proof for the case of a submartingale is similar.

□

**Lemma 5.2** Let $(Z_n)_n$ be a sequence of integrable r.v.'s, $(\mathscr{H}_n)_n$ a *decreasing* sequence of $\sigma$-algebras, i.e. such that $\mathscr{H}_{n+1} \subset \mathscr{H}_n$ for every $n = 0, 1, \dots$, with the property that $Z_n$ is $\mathscr{H}_n$-measurable and $Z_{n+1} \ge \mathrm{E}[Z_n \,|\, \mathscr{H}_{n+1}]$. Then, if $\sup_n \mathrm{E}[Z_n] = \ell < +\infty$, $(Z_n)_n$ is uniformly integrable.

*Proof* Let $\varepsilon > 0$. We must prove that, at least for large $n$, there exists a $c > 0$ such that $\mathrm{E}[|Z_n| 1_{\{|Z_n| \ge c\}}] \le \varepsilon$. As the sequence of the expectations $(\mathrm{E}[Z_n])_n$ is increasing to $\ell$ there exists a $k$ such that $\mathrm{E}[Z_k] \ge \ell - \varepsilon$. If $n \ge k$ we obtain for every $A \in \mathscr{H}_n$

$$\mathrm{E}[Z_k 1_A] = \mathrm{E}[\mathrm{E}(Z_k 1_A \,|\, \mathscr{H}_n)] = \mathrm{E}[1_A \mathrm{E}(Z_k \,|\, \mathscr{H}_n)] \le \mathrm{E}[Z_n 1_A] \ ,$$

hence

$$\mathrm{E}[Z_k 1_{\{Z_n < c\}}] \le \mathrm{E}[Z_n 1_{\{Z_n < c\}}] \ ,$$
$$\mathrm{E}[Z_k 1_{\{Z_n \le -c\}}] \le \mathrm{E}[Z_n 1_{\{Z_n \le -c\}}] \ .$$

Now, for every $n \geq k$,

$$
\begin{aligned}
\mathrm{E}[|Z_n|1_{\{|Z_n|\geq c\}}] &= \mathrm{E}[Z_n 1_{\{Z_n\geq c\}}] - \mathrm{E}[Z_n 1_{\{Z_n\leq -c\}}] \\
&= \mathrm{E}[Z_n] - \mathrm{E}[Z_n 1_{\{Z_n<c\}}] - \mathrm{E}[Z_n 1_{\{Z_n\leq -c\}}] \\
&\leq \ell - \mathrm{E}[Z_k 1_{\{Z_n<c\}}] - \mathrm{E}[Z_k 1_{\{Z_n\leq -c\}}] \\
&\leq \varepsilon + \mathrm{E}[Z_k] - \mathrm{E}[Z_k 1_{\{Z_n<c\}}] - \mathrm{E}[Z_k 1_{\{Z_n\leq -c\}}] \\
&= \varepsilon + \mathrm{E}[Z_k 1_{\{Z_n\geq c\}}] - \mathrm{E}[Z_k 1_{\{Z_n\leq -c\}}] \\
&= \varepsilon + \mathrm{E}[|Z_k|1_{\{|Z_n|\geq c\}}]\ .
\end{aligned}
\tag{5.19}
$$

Moreover, for every $n$,

$$
\mathrm{P}(|Z_n| \geq c) \leq \frac{1}{c}\,\mathrm{E}[|Z_n|] = \frac{1}{c}\,(2\mathrm{E}[Z_n^+] - \mathrm{E}[Z_n]) \leq \frac{1}{c}\,(2\ell - \mathrm{E}[Z_0])\ . \tag{5.20}
$$

Let now $\delta$ be such that $\mathrm{E}[Z_k 1_A] \leq \varepsilon$ for every $A \in \mathscr{F}$ satisfying $\mathrm{P}(A) \leq \delta$, as guaranteed by Proposition 5.2 (the single r.v. $Z_k$ forms a uniformly integrable family) and let $c$ be large enough so that $\mathrm{P}(|Z_n| \geq c) \leq \delta$, thanks to (5.20). Then by (5.19), for every $n \geq k$,

$$
\mathrm{E}[|Z_n|1_{\{|Z_n|\geq c\}}] \leq \varepsilon + \mathrm{E}[|Z_k|1_{\{|Z_n|\geq c\}}] \leq 2\varepsilon\ ,
$$

which is the relation we are looking for.

□

**Proposition 5.6** Let $M = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (M_t)_t, \mathrm{P})$ be a right-continuous martingale and $\tau$ a stopping time. Then $(M_{t\wedge\tau})_t$ is also an $(\mathscr{F}_t)_t$-martingale.

*Proof* Thanks to Exercise 5.6 we need only prove that, for every bounded stopping time $\sigma$, $\mathrm{E}(M_{\sigma\wedge\tau}) = \mathrm{E}(M_0)$. As $\sigma \wedge \tau$ is also a bounded stopping time, this follows from Theorem 5.13 applied to the two bounded stopping times $\tau_1 = 0$ and $\tau_2 = \sigma \wedge \tau$.

□

*Example 5.3* Let $B$ be a Brownian motion, $a, b > 0$ and denote by $\tau$ the exit time of $B$ from the interval $[-a, b]$. It is immediate that $\tau < +\infty$ a.s., thanks to the Iterated Logarithm Law. By continuity, the r.v. $B_\tau$, i.e. the position of $B$ at the time at which it exits from $[-a, b]$, can only take the values $-a$ and $b$. What is the value of $\mathrm{P}(B_\tau = -a)$?

(continued)

*Example 5.3* (continued)

As $B$ is a martingale, by the stopping theorem, Theorem 5.13, applied to the stopping times $\tau_1 = 0$, $\tau_2 = t \wedge \tau$, we have for every $t > 0$

$$0 = \mathrm{E}(B_0) = \mathrm{E}(B_{t\wedge\tau}) \ .$$

Letting $t \to +\infty$ $B_{t\wedge\tau} \to B_\tau$ and, by Lebesgue's theorem (the inequalities $-a \le B_{t\wedge\tau} \le b$ are obvious),

$$\begin{aligned} 0 = \mathrm{E}(B_\tau) &= -a\mathrm{P}(B_\tau = -a) + b\mathrm{P}(B_\tau = b) \\ &= -a\mathrm{P}(B_\tau = -a) + b(1 - \mathrm{P}(B_\tau = -a)) = b - (a+b)\mathrm{P}(B_\tau = -a) \end{aligned}$$

from which

$$\mathrm{P}(B_\tau = -a) = \frac{b}{a+b} \ , \qquad \mathrm{P}(B_\tau = b) = \frac{a}{a+b} \ \cdot$$

As in Example 5.1, we have solved a problem concerning the distribution of a process stopped at some stopping time with an application of the stopping theorem (Theorem 5.13 in this case) to some martingale.

Here, as in Example 5.1, the process $B$ itself is a martingale; in general, the stopping theorem will be applied to a martingale which is a function of the process or associated to the process in some way, as developed in the exercises.

The hypothesis of right continuity in all these statements is not going to be too constraining, as almost every martingale we are going to deal with will be continuous. In a couple of situations, however, we shall need the following result, which gives some simple hypotheses guaranteeing the existence of a right-continuous *modification*. See Revuz and Yor (1999) or Karatzas and Shreve (1991) for a proof.

**Theorem 5.14** Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (M_t)_t, \mathrm{P})$ be a supermartingale. Let us assume that the standard hypotheses are satisfied: the filtration $(\mathscr{F}_t)_t$ is right-continuous and is augmented with the negligible events of $\mathscr{F}_\infty$. Then $(M_t)_t$ admits a right-continuous modification if and only if the map $t \mapsto \mathrm{E}(M_t)$ is continuous.

We have seen in Proposition 3.4 that the paths of a Brownian motion do not have finite variation a.s. The following two statements show that this is a general property of every continuous martingale that is square integrable.

**Theorem 5.15** Let $M$ be a square integrable continuous martingale with respect to a filtration $(\mathscr{F}_t)_t$. Then if it has finite variation, it is a.s. constant.

*Proof* Possibly switching to the martingale $M'_t = M_t - M_0$, we can assume $M_0 = 0$. Let $\pi = 0 < t_0 < \dots < t_n = t$ denote a subdivision of the interval $[0,t]$ and let $|\pi| = \max|t_{i+1} - t_i|$. As was the case when investigating the variation of the paths of a Brownian motion, we shall first look at the quadratic variation of the paths. Let us assume first that $M$ has *bounded* variation on the interval $[0,T]$, i.e. that

$$V_T(\omega) = \sup_\pi \sum_{i=0}^{n-1} \big|M_{t_{i+1}}(\omega) - M_{t_i}(\omega)\big| \le K$$

for some $K > 0$ and for every $\omega$. Note that $\mathrm{E}[M_{t_{i+1}} M_{t_i}] = \mathrm{E}[\mathrm{E}(M_{t_{i+1}} \,|\, \mathscr{F}_{t_i}) M_{t_i}] = \mathrm{E}[M_{t_i}^2]$, so that

$$\mathrm{E}[(M_{t_{i+1}} - M_{t_i})^2] = \mathrm{E}[M_{t_{i+1}}^2 - 2M_{t_{i+1}}M_{t_i} + M_{t_i}^2] = \mathrm{E}[M_{t_{i+1}}^2 - M_{t_i}^2]\ .$$

Therefore

$$\mathrm{E}[M_t^2] = \mathrm{E}\Big[\sum_{i=0}^{n-1}(M_{t_{i+1}}^2 - M_{t_i}^2)\Big] = \mathrm{E}\Big[\sum_{i=0}^{n-1}(M_{t_{i+1}} - M_{t_i})^2\Big]\ .$$

As

$$(M_{t_{i+1}} - M_{t_i})^2 = \big|M_{t_{i+1}} - M_{t_i}\big| \cdot \big|M_{t_{i+1}} - M_{t_i}\big| \le \big|M_{t_{i+1}} - M_{t_i}\big| \cdot \max_{i \le n-1} \big|M_{t_{i+1}} - M_{t_i}\big|$$

we have

$$\begin{aligned}\mathrm{E}[M_t^2] \le \mathrm{E}\Big[\max_{0 \le i \le n-1} \big|M_{t_{i+1}} - M_{t_i}\big| \sum_{i=0}^{n-1} \big|M_{t_{i+1}} - M_{t_i}\big|\Big] &\le \mathrm{E}\big[V_T \max_{0 \le i \le n-1} \big|M_{t_{i+1}} - M_{t_i}\big|\big] \\ &\le K\mathrm{E}\big[\max_\pi \big|M_{t_{i+1}} - M_{t_i}\big|\big]\ .\end{aligned}$$

The paths of $M$ being continuous, if $(\pi_n)_n$ is a sequence of subdivisions such that $|\pi_n| \to 0$ then $\max_{\pi_n} |M_{t_{i+1}} - M_{t_i}| \to 0$ as $n \to \infty$ and also $\max_{\pi_n} |M_{t_{i+1}} - M_{t_i}| \le V_t \le K$. Therefore by Lebesgue's theorem

$$\mathrm{E}[M_t^2] = 0\ ,$$

which gives $M_t = 0$ a.s. The time $t$ being arbitrary, we have $\mathrm{P}(M_q = 0 \text{ for every } q \in \mathbb{Q}) = 1$, so that, as $M$ is assumed to be continuous, it is equal to 0 for every $t$ a.s.

In order to get rid of the assumption $V_T \le K$, let $V_t$ be the variation of $M$ up to time $t$, i.e.

$$V_t = \sup_{\pi} \sum_{i=0}^{n-1} |M_{t_{i+1}} - M_{t_i}|$$

for $\pi$ ranging among the partitions of the interval $[0, t]$. The process $(V_t)_t$ is adapted to the filtration $(\mathscr{F}_t)_t$, as $V_t$ is the limit of $\mathscr{F}_t$-measurable r.v.'s. Hence, for every $K > 0$, $\tau_K = \inf(s \ge 0; V_s \ge K)$ is a stopping time. Note that $(V_t)_t$ is continuous, which is a general property of the variation of a continuous function.

Hence $(V_{t\wedge\tau_K})_t$ is bounded by $K$ and is the variation of the stopped martingale $M_t^{\tau_K} = M_{\tau_K \wedge t}$, which therefore has bounded variation. By the first part of the proof $M^{\tau_K}$ is identically zero and this entails that $M_t \equiv 0$ on $\{V_T \le K\}$. Now just let $K \to +\infty$.

□

As a consequence,

> a square integrable stochastic process whose paths are differentiable or even Lipschitz continuous cannot be a martingale.

An important question is whether Doob's decomposition (Theorem 5.1) also holds for time continuous submartingales. We shall just mention the following result (see Revuz and Yor 1999, Chapter IV or Karatzas and Shreve 1991, Chapter 1 for proofs and more details).

**Theorem 5.16** Let $M$ be a square integrable continuous martingale with respect to a filtration $(\mathscr{F}_t)_t$ augmented with the negligible events of $\mathscr{F}_\infty$. Then there exists a unique increasing continuous process $A$ such that $A_0 = 0$ and $(M_t^2 - A_t)_t$ is a $(\mathscr{F}_t)_t$-martingale. If $\pi = \{0 = t_0 < t_1 < \dots < t_m = t\}$ is a partition of the interval $[0, t]$ then, in probability,

$$A_t = \lim_{|\pi|\to 0} \sum_{k=0}^{m-1} |M_{t_{k+1}} - M_{t_k}|^2 \,. \tag{5.21}$$

We shall still call the process $(A_t)_t$ of Theorem 5.16 the *associated increasing process* to the square integrable martingale $M$ and sometimes we shall use the notation $\langle M \rangle$. Thus $\langle M \rangle$ is a process such that $t \mapsto M_t^2 - \langle M \rangle_t$ is a martingale and Theorem 5.16 states that, if $M$ is a square integrable continuous martingale, such a process exists and is unique.

Note that (5.21) implies that the increasing process associated to a square integrable continuous martingale does not depend on the filtration: if $(M_t)_t$ is a martingale with respect to two filtrations, then the corresponding increasing associated processes are the same.

*Remark 5.3* Under the hypotheses of Theorem 5.16, if $\tau$ is a stopping time and we consider the stopped martingale $M_t^\tau = M_{t\wedge\tau}$ (which is also continuous and square integrable), is it true, as it may appear natural, that its associated increasing process can be obtained simply by stopping $\langle M\rangle$?

The answer is yes: it is clear that $(M_t^\tau)^2 - \langle M\rangle_{t\wedge\tau} = M_{t\wedge\tau}^2 - \langle M\rangle_{t\wedge\tau}$ is a martingale, being obtained by stopping the martingale $(M_t^2 - \langle M\rangle_t)_t$. By the uniqueness of the associated increasing process we have therefore, a.s.,

$$\langle M\rangle_{t\wedge\tau} = \langle M^\tau\rangle_t. \tag{5.22}$$

We are not going to prove Theorem 5.16, because we will be able to explicitly compute the associated increasing process of the martingales we are going to deal with.

*Example 5.4* If $(B_t)_t$ is a Brownian motion of the filtration $(\mathscr{F}_t)_t$, then we know (Example 5.1 c)) that it is also a martingale of that filtration. Therefore Proposition 3.4, compared with (5.21), can be rephrased by saying that the increasing process $A$ associated to a Brownian motion is $A_t = t$. This is consistent with the fact that $(B_t^2 - t)_t$ is a martingale, which is easy to verify (see also Exercise 5.10).

There are various characterizations of Brownian motion in terms of martingales. The following one will be very useful later.

**Theorem 5.17** Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be an $m$-dimensional continuous process such that $X_0 = 0$ a.s. and that, for every $\lambda \in \mathbb{R}^m$,

$$Y_t^\lambda = \mathrm{e}^{i\langle\lambda, X_t\rangle + \frac{1}{2}|\lambda|^2 t}$$

is an $(\mathscr{F}_t)_t$-martingale. Then $X$ is an $(\mathscr{F}_t)_t$-Brownian motion.

*Proof* The martingale relation

$$\mathrm{E}\big[\mathrm{e}^{i\langle\lambda,X_t\rangle+\frac{1}{2}|\lambda|^2 t}\,\big|\,\mathscr{F}_s\big]=\mathrm{E}(Y_t^\lambda\,|\,\mathscr{F}_s)=Y_s^\lambda=\mathrm{e}^{i\langle\lambda,X_s\rangle+\frac{1}{2}|\lambda|^2 s}$$

implies that, for every $\lambda\in\mathbb{R}^m$,

$$\mathrm{E}(\mathrm{e}^{i\langle\lambda,X_t-X_s\rangle}\,|\,\mathscr{F}_s)=\mathrm{e}^{-\frac{1}{2}|\lambda|^2(t-s)}\ . \tag{5.23}$$

Taking the expectation in (5.23) we find that the increment $X_t-X_s$ is $N(0,(t-s)I)$-distributed. Moreover, (Exercise 4.5), (5.23) implies that the r.v. $X_t-X_s$ is independent of $\mathscr{F}_s$. Therefore a), b) and c) of Definition 3.2 are satisfied.

□

## 5.7 Complements: the Laplace transform

One of the natural application of martingales appears in the computation of the Laplace transform of stopping times or of processes stopped at some stopping time (see Exercises 5.30, 5.31, 5.32, 5.33).

Let $\mu$ be a $\sigma$-finite measure on $\mathbb{R}^m$. Its *complex Laplace transform* is the function $\psi_\mu:\mathbb{C}^m\to\mathbb{C}^m$ defined as

$$\psi_\mu(z)=\int_{\mathbb{R}^m}\mathrm{e}^{\langle z,x\rangle}\,d\mu(x) \tag{5.24}$$

for every $z\in\mathbb{C}^m$ such that

$$\int_{\mathbb{R}^m}|\mathrm{e}^{\langle z,x\rangle}|\,d\mu(x)<+\infty\ . \tag{5.25}$$

The *domain* of the Laplace transform, denoted $\mathscr{D}_\mu$, is the set of the $z\in\mathbb{C}^m$ for which (5.25) is satisfied (and therefore for which the Laplace transform is defined). Recall that $\mathrm{e}^{\langle z,x\rangle}=\mathrm{e}^{\Re\langle z,x\rangle}(\cos(\Im\langle z,x\rangle)+i\sin(\Im\langle z,x\rangle)$ so that $|\mathrm{e}^{\langle z,x\rangle}|=\mathrm{e}^{\Re\langle z,x\rangle}$.

If $\mu$ has density $f$ with respect to Lebesgue measure then, of course,

$$\psi_\mu(z)=\int_{\mathbb{R}^m}\mathrm{e}^{\langle z,x\rangle}f(x)\,dx$$

and we shall refer to it indifferently as the Laplace transform of $\mu$ or of $f$. If $X$ is an $m$-dimensional r.v., the Laplace transform of its law $\mu_X$ is, of course,

$$\psi_{\mu_X}(z)=\mathrm{E}[\mathrm{e}^{\langle z,X\rangle}]\ ,$$

thanks to the integration rule with respect to an image law, Proposition 1.1.

In this section we recall some useful properties of the Laplace transform. We shall skip the proofs, in order to enhance instead consequences and applications. We shall assume $m = 1$, for simplicity. The passage to the case $m \geq 2$ will be rather obvious.

•**1**) There exist $x_1, x_2 \in \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}, x_1 \leq 0 \leq x_2$, (the *convergence abscissas*) such that, if $x_1 < \Re z < x_2$,

$$\int_{\mathbb{R}} |\, e^{zt}|\, d\mu(t) = \int_{\mathbb{R}} e^{t\Re z}\, d\mu(t) < +\infty$$

i.e. $z \in \mathscr{D}_\mu$, whereas

$$\int_{\mathbb{R}} |\, e^{zt}|\, d\mu(t) = +\infty \tag{5.26}$$

if $\Re z < x_1$ or $x_2 < \Re z$. In other words, the domain $\mathscr{D}_\mu$ contains an open strip (the *convergence strip*) of the form $x_1 < \Re z < x_2$, which can be empty (if $x_1 = x_2$).

Let $x_1 < x_2$ be real numbers such that

$$\int_{\mathbb{R}} e^{tx_i}\, d\mu(t) < +\infty, \quad i = 1, 2$$

and let $z \in \mathbb{C}$ such that $x_1 \leq \Re z \leq x_2$. Then $t\Re z \leq tx_2$ if $t \geq 0$ and $t\Re z \leq tx_1$ if $t \leq 0$. In any case, hence $e^{t\Re z} \leq e^{tx_1} + e^{tx_2}$ and (5.26) holds.

A typical situation appears when the measure $\mu$ has a support that is contained in $\mathbb{R}^+$. In this case $|e^{zt}| \leq 1$ if $\Re z < 0$; therefore the Laplace transform is defined at least on the whole half plane $\Re z < 0$ and in this case $x_1 = -\infty$.

Of course if $\mu$ is a finite measure, $\mathscr{D}_\mu$ always contains the imaginary axis $\Re z = 0$. Even in this situation it can happen that $x_1 = x_2 = 0$. Moreover, if $\mu$ is a probability, then its characteristic function $\widehat{\mu}$ is related to its Laplace transform by the relation $\widehat{\mu}(t) = \psi_\mu(it)$.

*Example 5.5*

a) Let $\mu(dx) = 1_{]0,+\infty[}\, \lambda e^{-\lambda x}$ be an exponential distribution. We have

$$\lambda \int_0^\infty e^{tz} e^{-\lambda t}\, dt = \lambda \int_0^\infty e^{-(\lambda - z)t}\, dt$$

so that the integral converges if and only if $\Re z < \lambda$. Therefore the convergence abscissas are $x_1 = -\infty$, $x_2 = \lambda$. For $\Re z < \lambda$ we have

$$\psi_\mu(z) = \lambda \int_0^\infty e^{-(\lambda - z)t}\, dt = \frac{\lambda}{\lambda - z}\,\cdot \tag{5.27}$$

(continued)

*Example 5.5* (continued)

b) If

$$\mu(dt) = \frac{dt}{\pi(1+t^2)}$$

(Cauchy law) then the two convergence abscissas are both equal to 0. Indeed

$$\int_{-\infty}^{+\infty} \frac{e^{t\Re z}}{\pi(1+t^2)}\, dt = +\infty$$

unless $\Re z = 0$.

**•2**) If the convergence abscissas are different (and therefore the convergence strip is non-empty) then $\psi_\mu$ is a holomorphic function in the convergence strip.

This property can be easily proved in many ways. For instance, one can prove that in the convergence strip we can take the derivative under the integral sign in (5.24). Then if we write $z = x + iy$ we have, using the fact that $z \mapsto e^{zt}$ is analytic and satisfies the Cauchy-Riemann equations itself,

$$\frac{\partial}{\partial x}\Re\psi_\mu(z) - \frac{\partial}{\partial y}\Im\psi_\mu(z) = \int_{-\infty}^{+\infty} \Big(\frac{\partial}{\partial x}\Re e^{zt} - \frac{\partial}{\partial y}\Im e^{zt}\Big)\, d\mu(t) = 0\,,$$
$$\frac{\partial}{\partial y}\Re\psi_\mu(z) + \frac{\partial}{\partial x}\Im\psi_\mu(z) = \int_{-\infty}^{+\infty} \Big(\frac{\partial}{\partial y}\Re e^{zt} + \frac{\partial}{\partial x}\Im e^{zt}\Big)\, d\mu(t) = 0\,,$$

so that the Cauchy–Riemann equations are satisfied by $\psi_\mu$. In particular, inside the convergence strip the Laplace transform is differentiable infinitely many times.

*Example 5.6*

a) If $\mu \sim N(0,1)$, we have already computed its Laplace transform for $x \in \mathbb{R}$ in Exercise 1.6, where we found $\psi_\mu(x) = e^{x^2/2}$. As the convergence abscissas are $x_1 = -\infty, x_2 = +\infty$, $\psi_\mu$ is analytic on the complex plane and by the uniqueness of the analytic extension this relation gives the value of $\psi_\mu$ for every $z \in \mathbb{C}$:

$$\psi_\mu(z) = e^{z^2/2}\,. \tag{5.28}$$

(continued)

*Example 5.6* (continued)

The same result might have been derived immediately from the value of the characteristic function

$$\psi_\mu(it) = \widehat{\mu}(t) = e^{-t^2/2} \tag{5.29}$$

and again using the argument of analytic extension. Note also that the analyticity of the Laplace transform might have been used in order to compute the characteristic function: first compute the Laplace transform for $x \in \mathbb{R}$ as in Exercise 1.6, then for every $z \in \mathbb{C}$ by analytic extension which gives the value of the characteristic function by restriction to the imaginary axis. This is probably the most elegant way of computing the characteristic function of a Gaussian law.

b) What is the characteristic function of a $\Gamma(\alpha, \lambda)$ distribution? Recall that, for $\alpha, \lambda > 0$, this is a distribution having density

$$f(x) = \frac{\lambda}{\Gamma(\alpha)}\, x^{\alpha-1} e^{-\lambda x}$$

for $x > 0$ and $f(x) = 0$ on the negative half-line, where $\Gamma$ is Euler's Gamma function, $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x}\, dx$. As developed above, we can compute the Laplace transform on the real axis and derive its value on the imaginary axis by analytic continuation. If $z \in \mathbb{R}$ and $z < \lambda$. We have

$$\begin{aligned}\psi(z) = \int e^{zx} f(x)\, dx &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} x^{\alpha-1} e^{-\lambda x} e^{zx}\, dx \\ = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} x^{\alpha-1} e^{-(\lambda-z)x}\, dx &= \frac{\lambda^\alpha}{\Gamma(\alpha)}\, \frac{\Gamma(\alpha)}{(\lambda-z)^\alpha} = \Big(\frac{\lambda}{\lambda-z}\Big)^\alpha ,\end{aligned}$$

where the integral was computed by recognizing, but for the normalizing constant, the expression of a $\Gamma(\alpha, \lambda - z)$ distribution. If $z \geq \lambda$ the integral diverges, hence $\lambda$ is the convergence abscissa. The characteristic function of $f$ is then obtained by substituting $z = i\theta$ and is equal to

$$\theta \mapsto \Big(\frac{\lambda}{\lambda - i\theta}\Big)^\alpha .$$

The fact that inside the convergence strip $\psi$ is differentiable infinitely many times and that one can differentiate under the integral sign has another consequence: if 0 belongs to the convergence strip, then

$$\psi'(0) = \int x\, d\mu(x) ,$$

i.e. the derivative at 0 of the Laplace transform is equal to the expectation. Iterating this procedure we have that if the origin belongs to the convergence strip then the r.v. having law $\mu$ has finite moments of all orders.

•**3**) Let us prove that if two $\sigma$-finite measures $\mu_1, \mu_2$ have a Laplace transform coinciding on an open non-empty interval $]x_1, x_2[\subset \mathbb{R}$, then $\mu_1 = \mu_2$. Again by the uniqueness of the analytic extension, the two Laplace transforms have a convergence strip that contains the strip $x_1 < \Re z < x_2$ and coincide in this strip. Moreover, if $\lambda$ is such that $x_1 < \lambda < x_2$ and we denote by $\widetilde{\mu}_j, j = 1, 2$, the measures $d\widetilde{\mu}_j(t) = \mathrm{e}^{\lambda t} d\mu_j(t)$, then, denoting by $\psi$ their common Laplace transform,

$$\psi_{\tilde{\mu}_j}(z) = \int_{\mathbb{R}} \mathrm{e}^{zt}\mathrm{e}^{\lambda t}\, d\mu_j(t) = \psi(\lambda + z)\ .$$

The measures $\widetilde{\mu}_j$ are finite, as

$$\widetilde{\mu}_j(\mathbb{R}) = \psi_{\tilde{\mu}_j}(0) = \psi(\lambda) < +\infty\ .$$

Up to a division by $\psi(\lambda)$, we can even assume that $\widetilde{\mu}_1$ and $\widetilde{\mu}_2$ are probabilities. Hence, as

$$\widehat{\widetilde{\mu}}_j(t) = \psi_{\tilde{\mu}_j}(it) = \psi(\lambda + it)\ ,$$

they have the same characteristic function. Therefore, by 6. of Sect. 1.6, $\widetilde{\mu}_1 = \widetilde{\mu}_2$ and therefore also $\mu_1 = \mu_2$.

This proves that the Laplace transform $\psi_\mu$ characterizes the probability $\mu$.

•**4**) As is the case of characteristic functions, it is easy to compute the Laplace transform of the sum of independent r.v.'s. If $X$ and $Y$ are independent then the product formula for the expectation of independent r.v.'s gives for the Laplace transform of their sum

$$\psi_{X+Y}(z) = \mathrm{E}[\mathrm{e}^{z(X+Y)}] = \mathrm{E}[\mathrm{e}^{zX}\mathrm{e}^{zY}] = \mathrm{E}[\mathrm{e}^{zX}]\mathrm{E}[\mathrm{e}^{zY}] = \psi_X(z)\psi_Y(z)$$

*provided* that both expectations are finite, i.e. the Laplace transform of the sum is equal to the product of the Laplace transforms and is defined in the intersection of the two convergence strips.

As for the characteristic functions this formula can be used in order to determine the distribution of $X + Y$.

•**5**) A useful property of the Laplace transform: *when restricted to* $\mathbb{R}^m$ $\psi$ is a convex function. Actually, even $\log \psi$ is convex, from which the convexity of $\psi$ follows, as the exponential function is convex and increasing. Recall Hölder's inequality (1.2): if $f, g$ are positive functions and $0 \le \alpha \le 1$, then

$$\int f(x)^\alpha g(x)^{1-\alpha}\, d\mu(x) \le \Big(\int f(x)\, d\mu(x)\Big)^\alpha \Big(\int g(x)\, d\mu(x)\Big)^{1-\alpha}\ .$$

Therefore

$$\psi(\alpha\theta + (1-\alpha)\lambda) = \int e^{(\alpha\theta+(1-\alpha)\lambda)x}\, d\mu(x) = \int \bigl(e^{\theta x}\bigr)^{\alpha}\bigl(e^{\lambda x}\bigr)^{1-\alpha}\, d\mu(x)$$
$$\leq \Bigl(\int e^{\theta x}\, d\mu(x)\Bigr)^{\alpha}\Bigl(\int e^{\lambda x}\, d\mu(x)\Bigr)^{1-\alpha} = \psi(\theta)^{\alpha}\psi(\lambda)^{1-\alpha}\ ,$$

from which we obtain the convexity relation for $\log\psi$:

$$\log\psi(\alpha\theta + (1-\alpha)\lambda) \leq \alpha\log\psi(\theta) + (1-\alpha)\log\psi(\lambda)\ .$$

•**6**) Let $X$ be a r.v. and let us assume that its Laplace transform $\psi$ is finite at a certain value $\lambda > 0$. Then for every $x > 0$, by Markov's inequality,

$$\psi(\lambda) = E[e^{\lambda X}] \geq e^{\lambda x}P(X \geq x)$$

i.e. we have for the tail of the distribution of $X$

$$P(X \geq x) \leq \psi(\lambda)e^{-\lambda x}\ .$$

The finiteness of the Laplace transform therefore provides a useful information concerning the decrease of the tail of the distribution. This gives to the convergence abscissa some meaning: if the right abscissa, $x_2$, is $> 0$, then, for every $\varepsilon > 0$,

$$P(X \geq x) \leq \text{const}\, e^{-(x_2-\varepsilon)x}$$

with *const* possibly depending on $\varepsilon$.

*Remark 5.4* To fix the ideas, let us consider Exercise 5.33. There we prove that if $\theta < 0$ then

$$E[e^{\theta\tau_a}] = e^{-a(\sqrt{\mu^2-2\theta}-\mu)}\ , \tag{5.30}$$

where $\tau_a$ is the passage time at $a > 0$ of $X_t = B_t + \mu t$, $\mu > 0$. Let us define $g(z) = e^{-a(\sqrt{\mu^2-2z}-\mu)}$. $g$ defines a holomorphic function also for some values of $z$ with $\Re z > 0$, actually on the whole half-plane $\Re z < \frac{\mu^2}{2}$. Can we say that (5.30) also holds for $0 \leq \theta < \frac{\mu^2}{2}$ and that $\frac{\mu^2}{2}$ is the convergence abscissa of the Laplace transform of $\tau_a$?

The answer is yes, as we shall see now with an argument that can be reproduced in other similar situations. This is very useful: once the value of the Laplace transform is computed for some values of its argument, its value

(continued)

*Remark 5.4* (continued)
is immediately known wherever the argument of analytic continuation can be applied.

Let $x_0 < 0$. The power series development of the exponential function at $x_0\tau_a$ gives

$$e^{x\tau_a} = \sum_{k=0}^{\infty} \frac{(x-x_0)^k}{k!}\, \tau_a^k e^{x_0\tau_a} .$$

If we choose $x$ real and such that $x - x_0 > 0$, the series above has positive terms and therefore can be integrated by series, which gives

$$\mathrm{E}[e^{x\tau_a}] = \sum_{k=0}^{\infty} \frac{(x-x_0)^k}{k!}\, \mathrm{E}[\tau_a^k e^{x_0\tau_a}] \tag{5.31}$$

and this equality remains true also if $\mathrm{E}[e^{x\tau_a}] = +\infty$ since the series on the right-hand side has positive terms. As also $g$, being holomorphic, has a power series development

$$g(z) = \sum_{k=0}^{\infty} \frac{(z-x_0)^k}{k!}\, a_k \tag{5.32}$$

and the two functions $g$ and $z \mapsto \mathrm{E}[e^{z\tau_a}]$ coincide on the half-plane $\Re z < 0$, we have $a_k = \mathrm{E}[\tau_a^k e^{x_0\tau_a}]$ and the two series, the one in (5.31) and the other in (5.32), necessarily have the same radius of convergence. Now a classical property of holomorphic functions is that their power series expansion converges in every ball in which they are holomorphic. As $g$ is holomorphic on the half-plane $\Re z < \frac{\mu^2}{2}$, then the series in (5.31) converges for $x < \frac{\mu^2}{2}$, so that $\frac{\mu^2}{2}$ is the convergence abscissa of the Laplace transform of $\tau_a$ and the relation (5.30) also holds on $\Re z < \frac{\mu^2}{2}$.

## Exercises

**5.1** (p. 485) Let $(\Omega, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be a supermartingale and assume, moreover, that $\mathrm{E}(X_t) =$ const. Then $(X_t)_t$ is a martingale. (This is a useful criterion.)

**5.2** (p. 486) (Continuation of Example 5.1) Let $(Z_n)_n$ be a sequence of i.i.d. r.v.'s taking the values $\pm 1$ with probability $\frac{1}{2}$ and let $X_0 = 0$ and $X_n = Z_1 + \cdots + Z_n$. Let

$a, b$ be positive integers and let $\tau_{a,b} = \inf\{n; X_n \geq b \text{ or } X_n \leq -a\}$, the exit time of $X$ from the interval $]-a, b[$.

a) Compute $\lim_{a\to+\infty} \mathrm{P}(X_{\tau_{a,b}} = b)$.
b) Let $\tau_b = \inf\{n; X_n \geq b\}$ be the passage time of $X$ at $b$. Deduce that $\tau_b < +\infty$ a.s.

**5.3** (p. 486) Let $(Y_n)_n$ be a sequence of i.i.d. r.v.'s such that $\mathrm{P}(Y_i = 1) = p$, $\mathrm{P}(Y_i = -1) = q$ with $q > p$. Let $X_n = Y_1 + \cdots + Y_n$.

a) Compute $\lim_{n\to\infty} \frac{1}{n} X_n$ and show that $\lim_{n\to\infty} X_n = -\infty$ a.s.
b1) Show that

$$Z_n = \Big(\frac{q}{p}\Big)^{X_n}$$

is a martingale.
b2) As $Z$ is positive, it converges a.s. Determine the value of $\lim_{n\to\infty} Z_n$.
c) Let $a, b \in \mathbb{N}$ be positive numbers and let $\tau = \inf\{n; X_n = b \text{ or } X_n = -a\}$. What is the value of $\mathrm{E}[Z_{n\wedge\tau}]$? And of $\mathrm{E}[Z_\tau]$?
d) What is the value of $\mathrm{P}(X_\tau = b)$ (i.e. what is the probability for the random walk $(X_n)_n$ to exit from the interval $]-a, b[$ at $b$)?

**5.4** (p. 487) Let $(X_n)_n$ be a sequence of independent r.v.'s on the probability space $(\Omega, \mathscr{F}, \mathrm{P})$ with mean 0 and variance $\sigma^2$ and let $\mathscr{F}_n = \sigma(X_k, k \leq n)$. Let $M_n = X_1 + \cdots + X_n$ and let $(Z_n)_n$ be a square integrable process predictable with respect to $(\mathscr{F}_n)_n$ (i.e. such that $Z_{n+1}$ is $\mathscr{F}_n$-measurable).

a) Show that

$$Y_n = \sum_{k=1}^n Z_k X_k$$

is a square integrable martingale.
b) Show that $\mathrm{E}[Y_n] = 0$ and that

$$\mathrm{E}[Y_n^2] = \sigma^2 \sum_{k=1}^n \mathrm{E}[Z_k^2]\ .$$

What is the associated increasing process of $(M_n)_n$? And of $(Y_n)_n$?
c) Let us assume $Z_k = \frac{1}{k}$. Is the martingale $(Y_n)_n$ uniformly integrable?

**5.5** (p. 488) Let $(Y_n)_{n\geq 1}$ be a sequence of independent r.v.'s such that

$$\begin{aligned} \mathrm{P}(Y_k = 1) &= 2^{-k} \\ \mathrm{P}(Y_k = 0) &= 1 - 2 \cdot 2^{-k} \\ \mathrm{P}(Y_k = -1) &= 2^{-k} \end{aligned}$$

and let $X_n = Y_1 + \cdots + Y_n$, $\mathscr{F}_n = \sigma(X_k, k \leq n)$.

a) Show that $(X_n)_n$ is a martingale of the filtration $(\mathscr{F}_n)_n$.
b) Show that $(X_n)_n$ is square integrable and compute its associated increasing process.
c) Does the martingale $(X_n)_n$ converge a.s.? In $L^1$? In $L^2$?

**5.6** (p. 489)

a) Let $(\mathscr{F}_t)_t$ be a filtration. Let $s \leq t$ and $A \in \mathscr{F}_s$. Prove that the r.v. $\tau$ defined as

$$\tau(\omega) = \begin{cases} s & \text{if } \omega \in A \\ t & \text{if } \omega \in A^c \end{cases}$$

is a stopping time.
b) Prove that an integrable right-continuous process $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ is a martingale if and only if, for every bounded $(\mathscr{F}_t)_t$-stopping time $\tau$, $\mathrm{E}(X_\tau) = \mathrm{E}(X_0)$.

**5.7** (p. 490) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a real Brownian motion. Prove that, for every $K \in \mathbb{R}$,

$$M_t = (\mathrm{e}^{B_t} - K)^+$$

is an $(\mathscr{F}_t)_t$-submartingale.

**5.8** (p. 490)

a) Let $M$ be a *positive* martingale. Prove that, for $s < t$, $\{M_s = 0\} \subset \{M_t = 0\}$ a.s.
b) Let $M = (M_t)_t$ be a right-continuous martingale.
b1) Prove that if $\tau = \inf\{t; M_t = 0\}$, then $M_\tau = 0$ on $\{\tau < +\infty\}$.
b2) Prove that if $M_T > 0$ a.s., then $\mathrm{P}(M_t > 0 \text{ for every } t \leq T) = 1$.

b2) Use the stopping theorem with the stopping times $T$, $T > 0$, and $T \wedge \tau$.

- Concerning b), let us point out that, in general, it is possible for a continuous process $X$ to have $\mathrm{P}(X_t > 0) = 1$ for every $t \leq T$ and $\mathrm{P}(X_t > 0 \text{ for every } t \leq T) = 0$. Even if, at first sight, this seems unlikely because of continuity.

An example is the following, which is an adaptation of Example 2.1: let $\Omega\ = ]0,1[$, $\mathscr{F} = \mathscr{B}(]0,1[)$ and P =Lebesgue measure. Define a process $X$ as $X_t(\omega) = 0$ if $t = \omega$, $X_0(\omega) = X_1(\omega) = 1$, $X_t(\omega)$ =linearly interpolated otherwise. We have, for every $t$, $\{X_t > 0\} = \Omega \setminus \{t\} = ]0,1[\setminus\{t\}$, which is a set of probability 1, but every path vanishes for one value of $t$. This exercise states that this phenomenon cannot occur if, in addition, $X$ is a martingale (Fig. 5.1).

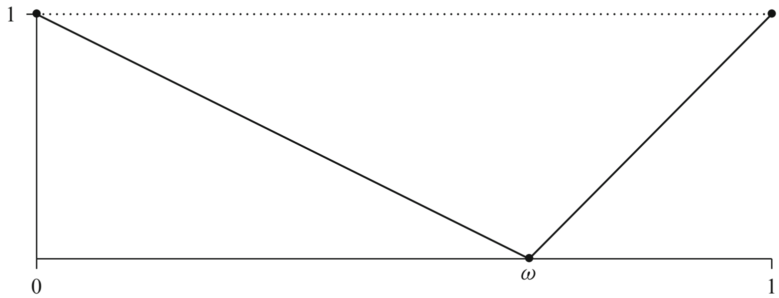


**Fig. 5.1** A typical graph of $t \mapsto X_t(\omega)$

**5.9** (p. 490) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion and let, for $\lambda \in \mathbb{R}^m$,

$$X_t = \mathrm{e}^{\langle \lambda, B_t\rangle - \frac{1}{2}|\lambda|^2 t}\ .$$

a) Prove that $(X_t)_t$ is an $(\mathscr{F}_t)_t$-martingale.
b) As the condition of Theorem 5.12 is satisfied ($X$ is positive) $\lim_{t\to+\infty} X_t$ exists a.s. and is finite. What is its value?
c) Is $X$ uniformly integrable?

**5.10** (p. 491) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a real Brownian motion and let $Y_t = B_t^2 - t$.

a) Prove that $(Y_t)_t$ is an $(\mathscr{F}_t)_t$-martingale. Is it uniformly integrable?
b) Let $\tau$ be the exit time of $B$ from the interval $]-a, b[$. In Example 5.3 we saw that $\tau < +\infty$ a.s. and computed the distribution of $X_\tau$. Can you derive from a) that $\mathrm{E}[B_\tau^2] = \mathrm{E}[\tau]$? What is the value of $\mathrm{E}[\tau]$? Is $\mathrm{E}[\tau]$ finite?
c) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion and let $Y_t = |B_t|^2 - mt$.
c1) Prove that $(Y_t)_t$ is an $(\mathscr{F}_t)_t$-martingale.
c2) Let us denote by $\tau$ the exit time of $B$ from the ball of radius 1 of $\mathbb{R}^m$. Compute $\mathrm{E}[\tau]$.

**5.11** (p. 493) Let $M = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (M_t)_t, \mathrm{P})$ be a square integrable martingale.

a) Show that $\mathrm{E}[(M_t - M_s)^2] = \mathrm{E}[M_t^2 - M_s^2]$.
b) $M$ is said to have *independent increments* if, for every $t > s$, the r.v. $M_t - M_s$ is independent of $\mathscr{F}_s$. Prove that in this case the associated increasing process is $\langle M \rangle_t = \mathrm{E}(M_t^2) - \mathrm{E}(M_0^2) = \mathrm{E}[(M_t - M_0)^2]$ and is therefore deterministic.
c) Show that a Gaussian martingale has necessarily independent increments with respect to its natural filtration.
d) Let us assume that $M = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (M_t)_t, \mathrm{P})$ has independent increments and, moreover, is a Gaussian martingale (i.e. simultaneously a martingale and a Gaussian process). Therefore its associated increasing process is deterministic, thanks to b) above. Show that, for every $\theta \in \mathbb{R}$,

$$Z_t = \mathrm{e}^{\theta M_t - \frac{1}{2}\theta^2 \langle M \rangle_t} \tag{5.33}$$

is an $(\mathscr{F}_t)_t$-martingale.

**5.12** (p. 494)

a) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_n)_n, (X_n)_n, \mathrm{P})$ be a (discrete time) martingale. Let $Y = (Y_n)_n$ be a process equivalent to $X$ (see p. 32 for the definition of equivalent processes). Show that $Y$ is a martingale with respect to its natural filtration.
b) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be a (time continuous) martingale. Let $Y = (Y_t)_t$ be a process equivalent to $X$. Show that $Y$ is a martingale with respect to its natural filtration.

**5.13** (p. 495) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion.

a) Show that

$$Y_t = tB_t - \int_0^t B_u \, du$$

is a martingale. Recall (Exercise 3.11) that we know that $(Y_t)_t$ is a Gaussian process.
b) Prove that, if $t > s$, $\mathrm{E}[(Y_t - Y_s)B_u] = 0$ for every $u \le s$.

**5.14** (p. 495)

a) Let $(\mathscr{F}_n)_n$ be a filtration, $X$ an integrable r.v. and $\tau$ an a.s. finite stopping time of $(\mathscr{F}_n)_n$. Let $X_n = \mathrm{E}(X|\mathscr{F}_n)$; then

$$\mathrm{E}(X|\mathscr{F}_\tau) = X_\tau \ .$$

b) Let $(\mathscr{F}_t)_{t\ge 0}$ be a filtration, $X$ an integrable r.v. and $\tau$ a a.s. finite stopping time. Let $(X_t)_t$ be a right continuous process such that $X_t = \mathrm{E}(X|\mathscr{F}_t)$ a.s. (thanks to Theorem 5.14 the process $(X_t)_t$ thus defined always has a right-continuous modification if $(\mathscr{F}_t)_t$ is the augmented natural filtration). Then

$$\mathrm{E}(X|\mathscr{F}_\tau) = X_\tau \ .$$

b) Use Lemma 3.3.

**5.15** (p. 496) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and $\lambda \in \mathbb{R}$. Prove that

$$M_t = \mathrm{e}^{\lambda t} B_t - \lambda \int_0^t \mathrm{e}^{\lambda u} B_u \, du$$

is an $(\mathscr{F}_t)_t$-martingale with independent increments.

**5.16** (p. 496) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and $\lambda, \nu$ real numbers with $\nu \geq 0$. Let

$$X_t = \mathrm{e}^{\lambda B_t - \nu t} \ .$$

a) Compute, for $\lambda \in \mathbb{R}$, $\nu > 0$, the value of $\lim_{t\to\infty} \mathrm{E}(X_t)$.
b) Show that $(X_t)_t$ is

$$\begin{cases} \text{a supermartingale} & \text{if } \frac{\lambda^2}{2} \leq \nu \\ \text{a martingale} & \text{if } \frac{\lambda^2}{2} = \nu \\ \text{a submartingale} & \text{if } \frac{\lambda^2}{2} \geq \nu \ . \end{cases}$$

c) Assume $\nu > 0$. Show that there exists an $\alpha > 0$ such that $(X_t^\alpha)_t$ is a supermartingale with $\lim_{t\to\infty} \mathrm{E}(X_t^\alpha) = 0$. Deduce that the limit $\lim_{t\to\infty} X_t$ exists a.s. for every $\lambda \in \mathbb{R}$ and compute it.
d) Prove that, if $\frac{\lambda^2}{2} < \nu$, then

$$A_\infty := \int_0^{+\infty} \mathrm{e}^{\lambda B_s - \nu s} \, ds < +\infty \qquad \text{a.s.}$$

and compute $\mathrm{E}(A_\infty)$. Prove that the r.v.'s

$$\int_0^{+\infty} \mathrm{e}^{\lambda B_s - \nu s} \, ds \quad \text{and} \quad \frac{1}{\lambda^2} \int_0^{+\infty} \mathrm{e}^{B_s - \lambda^{-2} \nu s} \, ds$$

have the same law.

**5.17** (p. 498) (The law of the supremum of a Brownian motion with a negative drift)

a) Let $(M_t)_t$ be a continuous positive martingale such that $\lim_{t\to+\infty} M_t = 0$ and $M_0 = 1$ (do you recall any examples of martingales with these properties?). For $x > 1$, let $\tau_x = \inf\{t; M_t \geq x\}$ its passage time at $x$. Show that

$$\lim_{t\to+\infty} M_{t\wedge\tau_x} = x 1_{\{\tau_x<+\infty\}} . \tag{5.34}$$

b) Deduce from (5.34) the value of $\mathrm{P}(\tau_x < +\infty)$.
c) Let $M^* = \sup_{t\geq 0} M_t$. Compute $\mathrm{P}(M^* \geq x)$.
d) Let $B$ be a Brownian motion and, for $\theta > 0$, let $X_t = B_t - \theta t$. Prove that $M_t = \mathrm{e}^{2\theta B_t - 2\theta^2 t}$ is a martingale and show that the r.v. $X^* = \sup_{t\geq 0} X_t$ is a.s. finite and has an exponential law with parameter $2\theta$.

**5.18** (p. 499) The aim of this exercise is the computation of the law of the supremum of a Brownian bridge $X$. As seen in Exercise 4.15, this is a continuous Gaussian process, centered and with covariance function $K_{s,t} = s(1-t)$ for $s \leq t$.

a) Show that there exists a Brownian motion $B$ such that, for $0 \leq t < 1$,

$$X_t = (1-t)B_{\frac{t}{1-t}} . \tag{5.35}$$

b) Prove that, for every $a > 0$,

$$\mathrm{P}\Big(\sup_{0\leq t\leq 1} X_t > a\Big) = \mathrm{P}\Big(\sup_{s>0}(B_s - as) > a\Big)$$

and deduce the partition function and density of $\sup_{0\leq t\leq 1} X_t$.

b) Use Exercise 5.17 d).

**5.19** (p. 499) Let $B$ be a Brownian motion and let $\tau = \inf\{t; B_t \notin ]-x, 1[\}$ be the exit time of $B$ from the interval $]-x, 1[$.

a) Compute $\mathrm{P}(B_\tau = -x)$.
b) We want to compute the distribution of the r.v.

$$Z = -\min_{0\leq t\leq \tau_1} B_t$$

where $\tau_1$ denotes the passage time of $B$ at 1. $Z$ is the minimum level attained by $B$ before passing at 1. Compute

$$\mathrm{P}(Z \geq x) = \mathrm{P}\Big(\min_{0\leq t\leq \tau_1} B_t \leq -x\Big) .$$

Does the r.v. $Z$ have finite mathematical expectation? If so, what is its value?

**5.20** (p. 500) In this exercise we compute the exit distribution from an interval of a Brownian motion with drift. As in Example 5.3, the problem is very simple as soon as we find the right martingale... Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and, for $\mu > 0$, let $X_t = B_t + \mu t$.

a) Prove that
$$M_t = \mathrm{e}^{-2\mu X_t}$$
is an $(\mathscr{F}_t)_t$-martingale.

b) Let $a, b > 0$ and let $\tau$ be the exit time of $(X_t)_t$ from the interval $]-a, b[$.

b1) Show that $\tau < +\infty$ a.s.

b2) What is the value of $\mathrm{P}(X_\tau = b)$? What is the value of the limit of this probability as $\mu \to \infty$?

**5.21** (p. 500) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion.

a) Show that $X_t = B_t^3 - 3tB_t$ is an $(\mathscr{F}_t)_t$-martingale.

b) Let $a, b > 0$ and let us denote by $\tau$ the exit time of $B$ from the interval $]-a, b[$.

Compute the covariance $\mathrm{Cov}(\tau, B_\tau)$.

Recall (Exercise 5.10) that we know that $\tau$ is integrable.

**5.22** (p. 501) (The product of independent martingales) Let $(M_t)_t$, $(N_t)_t$ be martingales on the same probability space $(\Omega, \mathscr{F}, \mathrm{P})$, with respect to the filtrations $(\mathscr{M}_t)_t$, $(\mathscr{N}_t)_t$, respectively. Let us assume, moreover, that the filtrations $(\mathscr{M}_t)_t$ and $(\mathscr{N}_t)_t$ are independent. Then the product $(M_tN_t)_t$ is a martingale of the filtration $\mathscr{H}_t = \mathscr{M}_t \vee \mathscr{N}_t$.

**5.23** (p. 501) Let $(M_t)_t$ be a continuous $(\mathscr{F}_t)_t$-martingale.

a) Prove that if $(M_t^2)_t$ is also a martingale, then $(M_t)_t$ is constant.

b1) Prove that if $p > 1$ and $(|M_t|^p)_t$ is a martingale, then $(|M_t|^{p'})_t$ is a martingale for every $1 \le p' \le p$.

b2) Prove that if $p \ge 2$ and $(|M_t|^p)_t$ is a martingale then $(M_t)_t$ is constant.

**5.24** (p. 502) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion.

a) Prove that, if $i \ne j$, the process $(B_i(t)B_j(t))_t$ is a $(\mathscr{F}_t)_t$-martingale.

b) Prove that if $i_1, \dots, i_d$, $d \le m$ are distinct indices, then $t \mapsto B_{i_1}(t) \dots B_{i_d}(t)$ is a martingale.

c) Let $B$ be an $m \times m$ matrix Brownian motion, i.e. $B_t = (B_{ij}(t))_{ij}$, $i, j = 1, \dots, m$, where the $B_{ij}$'s are independent Brownian motions. Prove that $X_t = \det(B_t)$ is a martingale. Prove that $Y_t = \det(B_tB_t^*)$ is a submartingale.

**5.25** (p. 503) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion.

a) Compute, for $t \le T$,

$$\mathrm{E}[1_{\{B_T>0\}} | \mathscr{F}_t] \; .$$

b) Let us denote by $\Phi$ the partition function of an $N(0, 1)$ distribution. Prove that

$$M_t = \Phi\Big(\frac{B_t}{\sqrt{T-t}}\Big)$$

is an $(\mathscr{F}_t)_t$-martingale for $t < T$. Is it uniformly integrable? Compute the a.s. limit $\lim_{t\to T-} M_t$.

**5.26** (p. 503) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion.

a) Show that the integral

$$X_t = \int_0^t \frac{B_u}{\sqrt{u}}\, du$$

converges a.s. for every $t \ge 0$.

b) Show that $(X_t)_t$ is a Gaussian process and compute its covariance function $K_{s,t} = \mathrm{Cov}(X_s, X_t)$.

c) Is $(X_t)_t$ a martingale?

**5.27** (p. 505) Let P, Q be probabilities on $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t)$. Let us assume that, for every $t > 0$, the restriction $\mathrm{Q}_{|\mathscr{F}_t}$ of Q to $\mathscr{F}_t$ is absolutely continuous with respect to the restriction of P to $\mathscr{F}_t$. Let

$$Z_t = \frac{d\mathrm{Q}_{|\mathscr{F}_t}}{d\mathrm{P}_{|\mathscr{F}_t}} \cdot$$

a) Show that $(Z_t)_t$ is a martingale.

b) Show that $Z_t > 0$ Q-a.s. and that $(Z_t^{-1})_t$ is a Q-supermartingale.

c) Show that if also $\mathrm{Q}_{|\mathscr{F}_t} \gg \mathrm{P}_{|\mathscr{F}_t}$, then $(Z_t^{-1})_t$ is a Q-martingale.

Recall (Exercise 5.8) that the set of the zeros $t \mapsto \{Z_t = 0\}$ is increasing.

**5.28** (p. 505) Let $(X_t)_t$ be a continuous process such that $X_t$ is square integrable for every $t$. Let $\mathscr{G} = \sigma(X_u, u \le s)$ and $\mathscr{G}_n = \sigma(X_{sk/2^n}, k = 1, \dots, 2^n)$. Let $t > s$. Does the conditional expectation of $X_t$ given $\mathscr{G}_n$ converge to that of $X_t$ given $\mathscr{G}$, as $n \to \infty$? In other words, if the process is known at the times $sk/2^n$ for $k = 1, \dots, 2^n$, is it true that if $n$ is large enough then, in order to predict the future position at time

$t$, it is almost as if we had the knowledge of the whole path of the process up to time $s$?

a) Show that the sequence of $\sigma$-algebras $(\mathscr{G}_n)_n$ is increasing and that their union generates $\mathscr{G}$.
b1) Let $Z_n = \mathrm{E}(X_t|\mathscr{G}_n)$. Show that the sequence $(Z_n)_n$ converges a.s. and in $L^2$ to $\mathrm{E}(X_t|\mathscr{G})$.
b2) How would the statement of b1) change if we just assumed $X_t \in L^1$ for every $t$?

**5.29** (p. 506) Let $(B_t)_t$ be an $(\mathscr{F}_t)_t$ Brownian motion and $\tau$ an *integrable* stopping time for the filtration $(\mathscr{F}_t)_t$. We want to prove the Wald equalities: $\mathrm{E}[B_\tau] = 0$, $\mathrm{E}[B_\tau^2] = \mathrm{E}[\tau]$. The situation is similar to Example 4.5 but the arguments are going to be different as $\tau$ here is not in general independent of $(B_t)_t$.

a) Prove that, for every $t \geq 0$, $\mathrm{E}[B_{\tau\wedge t}] = 0$ and $\mathrm{E}[B_{\tau\wedge t}^2] = \mathrm{E}[\tau \wedge t]$.
b) Prove that the martingale $(B_{\tau\wedge t})_t$ is bounded in $L^2$ and that

$$\mathrm{E}\Big[\sup_{t\geq 0} B_{\tau\wedge t}^2\Big] \leq 4\mathrm{E}[\tau]\,. \tag{5.36}$$

c) Prove that $\mathrm{E}[B_\tau] = 0$ and $\mathrm{E}[B_\tau^2] = \mathrm{E}[\tau]$.

**5.30** (p. 507) (The Laplace transform of the passage time of a Brownian motion) Let $B$ be a real Brownian motion, $a > 0$ and $\tau_a = \inf\{t; B_t \geq a\}$ the passage time at $a$. We know from Example 5.2 that $M_t = \mathrm{e}^{\lambda B_t - \frac{1}{2}\lambda^2 t}$ is a martingale.

a) Prove that, for $\lambda \geq 0$, $\mathrm{E}[M_{\tau_a}] = 1$. Why does the proof not work for $\lambda < 0$?
b) Show that the Laplace transform of $\tau_a$ is

$$\psi(\theta) = \mathrm{E}[\mathrm{e}^{\theta\tau_a}] = \mathrm{e}^{-a\sqrt{-2\theta}} \tag{5.37}$$

for $\theta \leq 0$, whereas $\psi(\theta) = +\infty$ for $\theta > 0$.
c) Show that $\tau_a$ has a law that is stable with exponent $\frac{1}{2}$ (this was already proved in a different manner in Exercise 3.20, where the definition of stable law is recalled).

**5.31** (p. 507) Let $B$ be a Brownian motion and, for $a > 0$, let us denote by $\tau$ the exit time of $B$ from the interval $[-a, a]$. In Example 5.3 we remarked that $\tau < +\infty$ a.s. and we computed the distribution of $X_\tau$. Show that, for $\theta > 0$,

$$\mathrm{E}[\mathrm{e}^{-\theta\tau}] = \frac{1}{\cosh(a\sqrt{2\theta})}\,\cdot$$

Find the right martingale... Recall that $B_\tau$ and $\tau$ are independent (Exercise 3.18). Further properties of exit times from bounded intervals are the object of Exercises 5.32, 8.5 and 10.5.

**5.32** (p. 508) Let $B$ be a Brownian motion with respect to a filtration $(\mathscr{F}_t)_t$ and let $\tau = \inf\{t; |B_t| \ge a\}$ be the exit time from $]-a, a[$. In Exercise 5.31 we computed the Laplace transform $\theta \mapsto \mathrm{E}[\mathrm{e}^{\theta\tau}]$ for $\theta \le 0$. Let us investigate this Laplace transform for $\theta > 0$. Is it going to be finite for some values $\theta > 0$?

a) Prove that, for every $\lambda \in \mathbb{R}$, $X_t = \cos(\lambda B_t)\,\mathrm{e}^{\frac{1}{2}\lambda^2 t}$ is an $(\mathscr{F}_t)_t$-martingale.
b) Prove that, if $|\lambda| < \frac{\pi}{2a}$,

$$\mathrm{E}[\mathrm{e}^{\frac{1}{2}\lambda^2(\tau\wedge t)}] \le \frac{1}{\cos(\lambda a)} < +\infty,$$

and that the r.v. $\mathrm{e}^{\frac{1}{2}\lambda^2\tau}$ is integrable. Prove that

$$\mathrm{E}[\mathrm{e}^{\theta\tau}] = \frac{1}{\cos(a\sqrt{2\theta})}, \qquad 0 \le \theta < \frac{\pi^2}{8a^2}\,\cdot \tag{5.38}$$

c) What is the value of $\mathrm{E}(\tau)$? Show that $\tau \in L^p$ for every $p \ge 0$
d) Determine the spectrum of the Laplace operator $\frac{1}{2}\frac{d^2}{dx^2}$ on the interval $]-a, a[$, with the Dirichlet conditions $u(a) = u(-a) = 0$, i.e. determine the numbers $\theta \in \mathbb{R}$ such that there exists a function $u$ not identically zero, twice differentiable in $]-a, a[$, continuous on $[-a, a]$ and such that

$$\begin{cases} \frac{1}{2}u'' = \theta u & \text{on } ]-a, a[ \\ u(-a) = u(a) = 0. \end{cases} \tag{5.39}$$

Which is the largest of these numbers? Do you notice some coincidence with the results of b)?

Recall that if $0 < \rho < \frac{\pi}{2}$, then $\cos x \ge \cos\rho > 0$ for $x \in [-\rho, \rho]$. See Exercise 10.5 concerning the relation between the convergence abscissas of the Laplace transform of the exit time and the spectrum of the generator of the process.

**5.33** (p. 510) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion. Let $\mu \in \mathbb{R}$, $a > 0$. We want to investigate the probability of crossing the level $a > 0$ and the time needed to do this for the process $X_t = B_t + \mu t$. For $\mu = 0$ the reflection principle already answers this question whereas Exercise 5.17 investigates the case $\mu < 0$. We now consider $\mu > 0$. Let $\tau = \inf\{t \ge 0; X_t \ge a\}$.

a) Prove that $\tau < +\infty$ a.s.
b) Prove that, for every $\lambda \in \mathbb{R}$,

$$M_t = \mathrm{e}^{\lambda X_t - (\frac{\lambda^2}{2} + \lambda\mu)t}$$

is an $(\mathscr{F}_t)_t$-martingale. What is the value of

$$\mathrm{E}[M_{t\wedge\tau}] = \mathrm{E}[\mathrm{e}^{\lambda X_{t\wedge\tau}-(\frac{\lambda^2}{2}+\lambda\mu)(t\wedge\tau)}]\ ?$$

c) Let $\lambda \geq 0$. Show that $M_{t\wedge\tau} \leq \mathrm{e}^{\lambda a}$ and that

$$\mathrm{E}[\mathrm{e}^{-(\frac{\lambda^2}{2}+\lambda\mu)\tau}] = \mathrm{e}^{-\lambda a}\ . \tag{5.40}$$

d) Compute, for $\theta > 0$, $\mathrm{E}[\mathrm{e}^{-\theta\tau}]$. What is the value of $\mathrm{E}[\tau]$?

# Chapter 6
# Markov Processes

In this chapter we introduce an important family of stochastic processes. Diffusions, which are the object of our investigation in the subsequent chapters, are instances of Markov processes.

## 6.1 Definitions and general facts

Let $(E, \mathscr{E})$ be a measurable space.

A *Markov transition function* on $(E, \mathscr{E})$ is a function $p(s, t, x, A)$, where $s, t, \in \mathbb{R}^+$, $s \le t$, $x \in E$, $A \in \mathscr{E}$, such that

i) for fixed $s, t, A$, $x \mapsto p(s, t, x, A)$ is $\mathscr{E}$-measurable;
ii) for fixed $s, t, x$, $A \mapsto p(s, t, x, A)$ is a probability on $(E, \mathscr{E})$;
iii) $p$ satisfies the *Chapman–Kolmogorov equation*

$$p(s, t, x, A) = \int_E p(u, t, y, A)\, p(s, u, x, dy) \tag{6.1}$$

for every $s \le u \le t$.

**Definition 6.1** Given on $(E, \mathscr{E})$ a Markov transition function $p$ and a probability $\mu$, an $(E, \mathscr{E})$-valued process $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \ge u}, (X_t)_{t \ge u}, \mathrm{P})$ is said

(continued)

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_6

**Definition 6.1** (continued)
to be a *Markov process* associated to $p$, starting at time $u$ and with initial distribution $\mu$, if

a) $X_u$ has law $\mu$.
b) (The Markov property) For every bounded measurable function $f : E \to \mathbb{R}$ and $t > s \geq u$

$$E(f(X_t)\,|\,\mathscr{F}_s) = \int_E f(x)\,p(s,t,X_s,dx) \qquad \text{a.s.} \tag{6.2}$$

When the filtration $(\mathscr{F}_t)_t$ is not specified it is understood, as usual, that it is the natural filtration.

As we shall see better in the sequel, $p(s,t,x,A)$ represents the probability that the process, being at position $x$ at time $s$, will move to a position in the set $A$ at time $t$. The Chapman–Kolmogorov equation intuitively means that if $s < u < t$, the probability of moving from position $x$ at time $s$ to a position in $A$ at time $t$ is equal to the probability of moving to a position $y$ at the intermediate time $u$ and then from $y$ to $A$, integrated over all possible values of $y$.

*Remark 6.1*

a) Let us have a closer look at the Markov property: if $f = 1_A$, $A \in \mathscr{E}$, (6.2) becomes, for $s \leq t$,

$$P(X_t \in A\,|\,\mathscr{F}_s) = p(s,t,X_s,A)\ .$$

Note that this conditional expectation given $\mathscr{F}_s$ only depends on $X_s$. Therefore

$$\begin{aligned} P(X_t \in A\,|\,X_s) &= E[P(X_t \in A\,|\,\mathscr{F}_s)\,|\,X_s] \\ = E[p(s,t,X_s,A)\,|\,X_s] &= p(s,t,X_s,A) = P(X_t \in A\,|\,\mathscr{F}_s)\ , \end{aligned} \tag{6.3}$$

i.e. the conditional probability of $\{X_t \in A\}$ given $X_s$ or given the whole past $\mathscr{F}_s$ coincide. Intuitively, the knowledge of the whole path of the process up to time $s$ or just of its position at time $s$ give the same information about the future position of the process at a time $t$, $t \geq s$.

(continued)

*Remark 6.1* (continued)

It is important to understand this aspect of the Markov property when modeling a random phenomenon with a Markov process. Consider, for instance, the evolution of financial assets (stocks, bonds, options,…): to assume that, in view of predicting tomorrow's price, the knowledge of the prices of the previous days, knowing today's price, does not give any additional valuable information is certainly not correct. Consider the importance of knowing whether or not we are in a trend of growing prices, and that to determine this we need to know the price evolution of the last few days. To model this kind of situation with a Markovian model may lead to grossly wrong appreciations.

However, we shall see how to derive a non Markovian model from a Markovian one (see Exercises 6.2 and 6.11, for example).

b) Recalling (4.18), the relation

$$\int_E f(x)\, p(s,t,X_s,dx) = \mathrm{E}[f(X_t)\,|\,\mathscr{F}_s] = \mathrm{E}[f(X_t)\,|X_s]$$

says that $p(s,t,x,\cdot)$ is the conditional distribution of $X_t$ given $X_s = x$. This a useful starting point when, given a Markov process, one has to determine the transition function.

c) For every bounded measurable function $f : E \to \mathbb{R}$ and for every $s < t$, the map

$$x \mapsto \int_E f(y)\, p(s,t,x,dy) \tag{6.4}$$

is measurable. Indeed i) of the definition of transition function states that (6.4) is measurable for every $f$ which is an indicator of a set $A \in \mathscr{E}$. Therefore this is true for every linear combination of such indicators and for every positive measurable function, as these are increasing limits of linear combinations of indicators. By subtracting the positive and negative parts, (6.4) is true for every bounded measurable function $f$.

d) Definition 6.1 allows to determine the finite-dimensional distributions of a Markov process. Let $f_0, f_1 : \mathscr{E} \to \mathbb{R}$ be Borel functions and $t > u$ and let us derive the joint distribution of $(X_u, X_t)$. As $X_u$ has law $\mu$, with the trick of performing an "intermediate conditioning",

$$\begin{aligned}\mathrm{E}[f_1(X_t)f_0(X_u)] &= \mathrm{E}\big[f_0(X_u)\mathrm{E}[f_1(X_t)\,|\,\mathscr{F}_u]\big]\\ &= \mathrm{E}\Big[f_0(X_u)\int_E f_1(x_1)\,p(s,t,X_u,dx_1)\Big]\\ &= \int_E f_0(x_0)\mu(dx_0)\int_E f_1(x_1)\,p(s,t,x_0,dx_1)\,.\end{aligned} \tag{6.5}$$

(continued)

*Remark 6.1* (continued)
By induction on $m$ we have, for every $u \le t_1 < \cdots < t_m$,

$$\begin{aligned} &\mathrm{E}[f_0(X_u)f_1(X_{t_1})\dots f_m(X_{t_m})] \\ &= \int_E f_0(x_0)\mu(dx_0)\int_E f_1(x_1)\,p(u,t_1,x_0,dx_1)\dots \\ &\qquad \dots\int_E f_{m-1}(x_{m-1})\,p(t_{m-2},t_{m-1},x_{m-2},dx_{m-1}) \\ &\qquad\qquad \times\int_E f_m(x_m)p(t_{m-1},t_m,x_{m-1},dx_m)\;. \end{aligned} \tag{6.6}$$

The meaning of (6.6) may be difficult to decrypt at first, however it is simple: once the rightmost integral in $dx_m$ has been computed, the result is a function of $x_{m-1}$, which is integrated in the subsequent integral giving, rise to a function of $x_{m-2}$, and so on.

Formula (6.6) states that *a priori* the finite-dimensional distributions of a Markov process only depend on the initial distribution $\mu$, the starting time $u$ and the transition function $p$. In particular, two Markov processes associated to the same transition function and having the same starting time and starting distribution are equivalent.

*Example 6.1* If $E = \mathbb{R}^m$, $\mathscr{E} = \mathscr{B}(\mathbb{R}^m)$, let

$$p(s,t,x,A) = \frac{1}{[2\pi(t-s)]^{m/2}}\int_A \exp\Big[-\frac{|y-x|^2}{2(t-s)}\Big]\,dy\;. \tag{6.7}$$

Then $p$ is a Markov transition function: $p(s,t,x,\cdot)$ is an $N(x,(t-s)I)$ law. The Chapman–Kolmogorov equation is a consequence of the property of the normal laws with respect to the convolution product: if $A \in \mathscr{B}(\mathbb{R}^m)$, then

$$\begin{aligned} &\int p(u,t,y,A)\,p(s,u,x,dy) \\ &= \int \frac{1}{[2\pi(u-s)]^{m/2}}\exp\Big[-\frac{|y-x|^2}{2(u-s)}\Big]\,dy\int_A \frac{1}{[2\pi(t-u)]^{m/2}}\exp\Big[-\frac{|z-y|^2}{2(t-u)}\Big]\,dz \\ &= \int_A dz\int \underbrace{\frac{1}{[2\pi(u-s)]^{m/2}}\exp\Big[-\frac{|y-x|^2}{2(u-s)}\Big]}_{f_1(y)}\underbrace{\frac{1}{[2\pi(t-u)]^{m/2}}\exp\Big[-\frac{|z-y|^2}{2(t-u)}\Big]}_{f_2(z-y)}\,dy \end{aligned}$$

so that the inner integral is the convolution of the densities $f_1$, which is $N(x,(u-s)I)$, and $f_2 \sim N(0,(t-u)I)$. The result, owing to well-known

(continued)

*Example 6.1* (continued)
properties of Gaussian distributions (see Sect. 1.7), is $N(x,(t-s)I)$, i.e. $p(s,t,x,\cdot)$. Therefore we can conclude that

$$\int p(u,t,y,A)\,p(s,u,x,dy) = \frac{1}{[2\pi(t-s)]^{m/2}}\int_A \exp\Big[-\frac{|y-x|^2}{2(t-s)}\Big]\,dy = p(s,t,x,A)\;.$$

In Example 4.4 we computed

$$\begin{aligned}\mathrm{E}[f(B_t)\,|\,\mathscr{F}_s] &= \frac{1}{[2\pi(t-s)]^{m/2}}\int f(y)\exp\Big[-\frac{|y-B_s|^2}{2(t-s)}\Big]\,dy\\ &= \int_E f(x)\,p(s,t,B_s,dx)\;,\end{aligned}$$

i.e. the Markov property, b) of Definition 6.1, is satisfied by Brownian motion, with respect to the Markov transition function given by (6.7). Therefore Brownian motion becomes our first example of Markov process. To be precise, it is a Markov process associated to the transition function (6.7) and starting at 0 at time 0.

Let us now prove the existence of the Markov process associated to a given transition function and initial distribution. The idea is simple: thanks to the previous remark we know what the finite-dimensional distributions are. We must therefore

a) check that these satisfy the coherence Condition 2.1 of Kolmogorov's existence Theorem 2.2 so that a stochastic process with these finite-dimensional distributions does exist;
b) prove that a stochastic process having such finite-dimensional distributions is actually a Markov process associated to the given transition function, which amounts to proving that it enjoys the Markov property (6.3).

Assume $E$ to be a complete separable metric space endowed with its Borel $\sigma$-algebra $\mathscr{B}(E)$ and $p$ a Markov transition function on $(E,\mathscr{B}(E))$. Let $\Omega = E^{\mathbb{R}^+}$; an element $\omega\in\Omega$ is therefore a map $\mathbb{R}^+\to E$. Let us define $X_t:\Omega\to E$ as $X_t(\omega)=\omega(t)$, $t\ge 0$ and then $\mathscr{F}_t^u=\sigma(X_s, u\le s\le t)$, $\mathscr{F}=\mathscr{F}_\infty^u$. Let us consider the system of finite-dimensional distributions defined by (6.6). The Chapman–Kolmogorov equation (6.1) easily implies that this system of finite-dimensional distributions satisfies the Condition 2.1 of coherence.

Therefore there exists a unique probability P on $(\Omega,\mathscr{F})$ such that the probabilities (6.6) are the finite-dimensional distributions of $(\Omega,\mathscr{F},(\mathscr{F}_t^u)_t,(X_t)_t,\mathrm{P})$. We now have to check that this is a Markov process associated to $p$ and with initial (i.e. at time $u$) distribution $\mu$. Part a) of Definition 6.1 being immediate, we have to check the Markov property b), i.e. we must prove that if $D \in \mathscr{F}_s^u$ and $f : E \to \mathbb{R}$ is bounded and measurable, then

$$\mathrm{E}[f(X_t)1_D] = \mathrm{E}\Big[1_D \int_E f(x)\, p(s,t,X_s,dx)\Big]\,. \qquad (6.8)$$

It will, however, be sufficient (Remark 4.2) to prove this relation for a set $D$ of the form

$$D = \{X_{t_0} \in B_0, \dots, X_{t_n} \in B_n\}\ ,$$

where $u = t_0 < t_1 < \dots < t_n = s$, since by definition the events of this form generate $\mathscr{F}_s^u$, are a class that is stable with respect to finite intersections, and $\Omega$ itself is of this form. For this choice of events the verification of (6.8) is now easy as both sides are easily expressed in terms of finite-dimensional distributions: as $1_{\{X_{t_k} \in B_{t_k}\}} = 1_{B_k}(X_{t_k})$, we have $1_D = 1_{B_0}(X_{t_0}) \dots 1_{B_n}(X_{t_n})$ and

$$\begin{aligned}
\mathrm{E}[f(X_t)1_D] &= \mathrm{E}\big[1_{B_0}(X_{t_0}) \dots 1_{B_{n-1}}(X_{t_{n-1}})1_{B_n}(X_s)f(X_t)\big] \\
&= \int 1_{B_0}(y_0)\mu(dy_0) \int 1_{B_1}(y_1)p(u,t_1,y_0,dy_1)\dots \\
&\qquad \dots \int 1_{B_n}(y_n)p(t_{n-1},s,y_{n-1},dy_n) \underbrace{\int f(y)\,p(s,t,y_n,dy)}_{:=\widetilde{f}(y_n)} \\
&= \int 1_{B_0}(y_0)\mu(dy_0) \int 1_{B_1}(y_1)\dots \int 1_{B_n}(y_n)\widetilde{f}(y_n)p(t_{n-1},s,y_{n-1},dy_n) \\
&= \mathrm{E}[1_{B_0}(X_{t_0})\dots 1_{B_{n-1}}(X_{t_{n-1}})1_{B_n}(X_s)\widetilde{f}(X_s)] = \mathrm{E}\big[1_D\widetilde{f}(X_s)\big]\,.
\end{aligned}$$

Hence (6.8) holds for every $D \in \mathscr{F}_s^u$ and as $\widetilde{f}(X_s)$ is $\mathscr{F}_s^u$-measurable we have proved that

$$\mathrm{E}[f(X_t)\,|\,\mathscr{F}_s] = \widetilde{f}(X_s) = \int f(y)\,p(s,t,X_s,dy)\ .$$

Therefore $(\Omega, \mathscr{F}, (\mathscr{F}_t^u)_{t\ge u}, (X_t)_{t\ge u}, \mathrm{P})$ is a process satisfying conditions a) and b) of Definition 6.1. The probability P just constructed depends of course, besides $p$, on $\mu$ and on $u$ and will be denoted $\mathrm{P}^{\mu,u}$. If $\mu = \delta_x$ we shall write $\mathrm{P}^{x,u}$ instead of $\mathrm{P}^{\delta_x,u}$ and we shall denote by $\mathrm{E}^{x,s}$ the expectation computed with respect to $\mathrm{P}^{x,s}$.

Going back to the construction above, we have proved that, if $E$ is a complete separable metric space, then there exist

a) a measurable space $(\Omega, \mathscr{F})$ endowed with a family of filtrations $(\mathscr{F}_t^s)_{t\ge s}$, such that $\mathscr{F}_{t'}^{s'} \subset \mathscr{F}_t^s$ if $s \le s'$, $t' \le t$;

b) a family of r.v.'s $X_t : \Omega \to E$ such that $\omega \mapsto X_t(\omega)$ is $\mathscr{F}_t^s$-measurable for every $s \le t$;
c) a family of probabilities $(\mathrm{P}^{x,s})_{x\in E, s\in\mathbb{R}^+}$ such that $\mathrm{P}^{x,s}$ is a probability on $(\Omega, \mathscr{F}_\infty^s)$ with the properties that, for every $x \in E$ and $s, h > 0$,

$$\mathrm{P}^{x,s}(X_s = x) = 1 \tag{6.9}$$

$$\mathrm{E}^{x,s}[f(X_{t+h})\,|\,\mathscr{F}_t^s] = \int_E f(z)\,p(t, t+h, X_t, dz) \quad \mathrm{P}^{x,s}\text{-a.s.} \tag{6.10}$$

Note that in (6.10) the value of the conditional expectation does not depend on $s$.

A family of processes $(\Omega, \mathscr{F}, (\mathscr{F}_t^s)_{t\ge s}, (X_t)_t, (\mathrm{P}^{x,s})_{x,s})$ satisfying a), b), c) is called a *realization of the Markov process* associated to the given transition function $p$. In some sense, the realization is a unique space $(\Omega, \mathscr{F})$ on which we consider a family of probabilities $\mathrm{P}^{x,s}$ that are the laws of the Markov processes associated to the given transition function, only depending on the starting position and initial time.

Let us try to familiarize ourselves with these notations. Going back to the expression of the finite-dimensional distributions (6.6) let us observe that, if $\mu = \delta_x$ (i.e. if the starting distribution is concentrated at $x$), (6.5) with $f_0 \equiv 1$ gives

$$\mathrm{E}^{x,s}[f(X_t)] = \int_E \delta_x(dx_0) \int_E f(x_1)p(s,t,x_0,dx_1) = \int_E f(x_1)p(s,t,x,dx_1) \tag{6.11}$$

so that, with respect to $\mathrm{P}^{x,s}$, $p(s,t,x,\cdot)$ is the law of $X_t$. If $f = 1_A$ then the previous relation becomes

$$\mathrm{P}^{x,s}(X_t \in A) = p(s,t,x,A)\ . \tag{6.12}$$

Thanks to (6.11) the Markov property (6.10) can also be written as

$$\mathrm{E}^{x,s}[f(X_{t+h})\,|\,\mathscr{F}_t^s] = \int f(z)p(t,t+h,X_t,dz) = \mathrm{E}^{X_t,t}[f(X_{t+h})] \quad \mathrm{P}^{x,s}\text{-a.s.}\,, \tag{6.13}$$

for every $x, s$ or, for $f = 1_A$,

$$\mathrm{P}^{x,s}(X_{t+h} \in A\,|\,\mathscr{F}_t^s) = p(t,t+h,X_t,A) = \mathrm{P}^{X_t,t}(X_{t+h} \in A) \qquad \mathrm{P}^{x,s}\text{-a.s.} \tag{6.14}$$

for every $x, s$. The expression $\mathrm{P}^{X_t,t}(X_{t+h} \in A)$ can initially create some confusion: it is simply the composition of the function $x \mapsto \mathrm{P}^{x,t}(X_{t+h} \in A)$ with the r.v. $X_t$. From now on in this chapter $X = (\Omega, \mathscr{F}, (\mathscr{F}_t^s)_{t\ge s}, (X_t)_{t\ge 0}, (\mathrm{P}^{x,s})_{x,s})$ will denote the realization of a Markov process associated to the transition function $p$.

The Markov property, (6.10) or (6.13), allows us to compute the conditional expectation with respect to $\mathscr{F}_t^s$ of a r.v., $f(X_{t+h})$, that depends on the position of the process at a fixed time $t + h$ posterior to $t$. Sometimes, however, it is necessary to compute the conditional expectation $\mathrm{E}^{x,s}(Y|\mathscr{F}_t^s)$ for a r.v. $Y$ that depends, possibly, on the whole path of the process after time $t$.

The following proposition states that the Markov property can be extended to cover this situation. Note that, as intuition suggests, this conditional expectation also depends only on the position, $X_t$, at time $t$.

Let $\mathscr{G}_t^s = \sigma(X_u, s \le u \le t)$. Intuitively, a $\mathscr{G}_\infty^s$-measurable r.v. is r.v. that depends on the behavior of the paths after time $s$.

**Proposition 6.1** Let $Y$ be a bounded $\mathscr{G}_\infty^t$-measurable r.v. Then

$$\mathrm{E}^{x,s}(Y|\mathscr{F}_t^s) = \mathrm{E}^{X_t,t}(Y) \quad \mathrm{P}^{x,s}\text{-a.s.} \tag{6.15}$$

*Proof* Let us assume first that $Y$ is of the form $f_1(X_{t_1}) \dots f_m(X_{t_m})$, where $t \le t_1 < \dots < t_m$ and $f_1, \dots, f_m$ are bounded measurable functions. If $m = 1$ then (6.15) is the same as (6.13). Let us assume that (6.15) holds for $Y$ as above and for $m-1$. Then, conditioning first with respect to $\mathscr{F}_{t_{m-1}}^s$,

$$\begin{aligned}
&\mathrm{E}^{x,s}[f_1(X_{t_1}) \dots f_m(X_{t_m})|\mathscr{F}_t^s]\\
&= \mathrm{E}^{x,s}\big[f_1(X_{t_1}) \dots f_{m-1}(X_{t_{m-1}})\mathrm{E}^{x,s}[f_m(X_{t_m})|\mathscr{F}_{t_{m-1}}^s]|\mathscr{F}_t^s\big]\\
&= \mathrm{E}^{x,s}\big[f_1(X_{t_1}) \dots f_{m-1}(X_{t_{m-1}})\mathrm{E}^{X_{t_{m-1}},t_{m-1}}[f_m(X_{t_m})] \,\big|\, \mathscr{F}_t^s\big]\\
&= \mathrm{E}^{x,s}[f_1(X_{t_1}) \dots \widetilde{f}_{m-1}(X_{t_{m-1}})|\mathscr{F}_t^s]\ ,
\end{aligned}$$

where $\widetilde{f}_{m-1}(x) = f_{m-1}(x) \cdot \mathrm{E}^{x,t_{m-1}}[f_m(X_{t_m})]$; by the induction hypothesis

$$\mathrm{E}^{x,s}[f_1(X_{t_1}) \dots f_m(X_{t_m})|\mathscr{F}_t^s] = \mathrm{E}^{X_t,t}[f_1(X_{t_1}) \dots \widetilde{f}_{m-1}(X_{t_{m-1}})]\ .$$

However, by (6.13), $\mathrm{E}^{X_{t_{m-1}},t_{m-1}}[f_m(X_{t_m})] = \mathrm{E}^{x,t}[f_m(X_{t_m})|\mathscr{F}_{t_{m-1}}^t]$ $\mathrm{P}^{x,t}$-a.s. for every $x, t$ so that

$$\begin{aligned}
\mathrm{E}^{X_t,t}[f_1(X_{t_1}) \dots \widetilde{f}_{m-1}(X_{t_{m-1}})] &= \mathrm{E}^{X_t,t}[f_1(X_{t_1}) \dots f_{m-1}(X_{t_{m-1}}) \cdot \mathrm{E}^{X_{t_{m-1}},t_{m-1}}[f_m(X_{t_m})]]\\
&= \mathrm{E}^{X_t,t}[f_1(X_{t_1}) \dots f_{m-1}(X_{t_{m-1}}) \cdot \mathrm{E}^{X_t,t}[f_m(X_{t_m})|\mathscr{F}_{t_{m-1}}^t]]\\
&= \mathrm{E}^{X_t,t}[\mathrm{E}^{X_t,t}[f_1(X_{t_1}) \dots f_{m-1}(X_{t_{m-1}})f_m(X_{t_m})|\mathscr{F}_{t_{m-1}}^t]]\\
&= \mathrm{E}^{X_t,t}[f_1(X_{t_1}) \dots f_{m-1}(X_{t_{m-1}})f_m(X_{t_m})]\ .
\end{aligned}$$

We obtain the general case using Theorem 1.4.

□

*Remark 6.2* In some sense the Chapman–Kolmogorov equation follows from the Markov property. More precisely, let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t^s)_{t\ge s}, (X_t)_t, (\mathrm{P}^{x,s})_{x,s})$ be a family of stochastic processes such that $X_s = x$ $\mathrm{P}^{x,s}$-a.s. for every $x, s$ and let $p(s,t,x,A)$ be a function satisfying the conditions i) and ii) of the definition of a transition function on p. 151. Then, if for every $x \in E$ and $s \le u \le t$,

$$\mathrm{P}^{x,s}(X_t \in A\,|\,\mathscr{F}_u^s) = p(u,t,X_u,A) \qquad \mathrm{P}^{x,s}\text{-a.s.}\ ,$$

necessarily $p$ satisfies the Chapman–Kolmogorov equation (6.1) and $X$ is a realization of the Markov process associated to $p$.

Indeed first with $u = s$ in the previous relation we obtain, as $X_s = x$ $\mathrm{P}^{x,s}$-a.s.,

$$\mathrm{P}^{x,s}(X_t \in A\,|\,\mathscr{F}_s^s) = p(s,t,X_s,A) = p(s,t,x,A) \qquad \mathrm{P}^{x,s}\text{-a.s.}$$

i.e. $p(s,t,x,\cdot)$ is the law of $X_t$ with respect to $\mathrm{P}^{x,s}$. Then, if $s \le u \le t$,

$$\begin{aligned} p(s,t,x,A) &= \mathrm{P}^{x,s}(X_t \in A) = \mathrm{E}^{x,s}[\mathrm{P}^{x,s}(X_t \in A\,|\,\mathscr{F}_u^s)] \\ &= \mathrm{E}^{x,s}[p(u,t,X_u,A)] = \int_E p(u,t,y,A)p(s,u,x,dy)\ , \end{aligned}$$

which is the required Chapman–Kolmogorov relation.

The probability of making a given transition in a time span $h$, $p(s,s+h,x,A)$, in general depends on the time $s$. Sometimes the transition function $p$ depends on $s$ and $t$ only as a function of $t-s$, so that $p(s,s+h,x,A)$ does not actually depend on $s$. In this case, $p$ is said to be *time homogeneous* and, recalling the expression of the finite-dimensional distributions,

$$\mathrm{E}^{x,s}[f_1(X_{t_1+s})\dots f_m(X_{t_m+s})] = \mathrm{E}^{x,0}[f_1(X_{t_1})\dots f_m(X_{t_m})]\ .$$

In this case, as the transition $p(s,s+h,x,A)$ does not depend on $s$, in some sense the behavior of the process is the same whatever the initial time. It is now convenient to fix 0 as the initial time and to set $p(t,x,A) = p(0,t,x,A)$, $\mathscr{F}_t = \mathscr{F}_t^0$, $\mathrm{P}^x = \mathrm{P}^{x,0}$ and consider as a realization $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$. The Chapman–Kolmogorov equation becomes, for $0 \le s < t$,

$$p(t,x,A) = \int p(t-s,y,A)p(s,x,dy)$$

and the relations (6.12) and (6.14), for $s \le t$, are now

$$\begin{aligned} &\mathrm{P}^x(X_t \in A) = p(t,x,A) \\ &\mathrm{P}^x(X_t \in A\,|\,\mathscr{F}_s) = p(t-s,X_s,A) = \mathrm{P}^{X_s}(X_{t-s} \in A) \qquad \mathrm{P}^x\text{-a.s.} \end{aligned}$$

The transition function of Brownian motion, see Example 6.1, is time homogeneous.

*Remark 6.3* Transition functions such as that of Brownian motion, Example 6.1, satisfy another kind of invariance: by looking back to the expression for the transition function (6.7), by a change of variable we find that

$$p(t,x,A) = p(t,0,A-x)\ .$$

This implies that, if $(B_t)_t$ is a Brownian motion, then $(B^x_t)_t$, with $B^x_t = x + B_t$, is a Markov process associated to the same transition function and starting at $x$. Indeed, going back to (4.10),

$$\begin{aligned}\mathrm{P}(B^x_t \in A \,|\, \mathscr{F}_s) &= \mathrm{P}(B_t \in A-x \,|\, \mathscr{F}_s) = p(t-s, B_s, A-x)\\ &= p(t-s, B_s + x, A) = p(t-s, B^x_s, A)\ .\end{aligned}$$

Hence the Markov process associated to the transition function of Brownian motion, (6.7), and starting at $x$ has the same law as $t \mapsto B_t + x$.

## 6.2 The Feller and strong Markov properties

A positive r.v. $\tau$ is said to be an *s-stopping time* if $\tau \geq s$ and if it is a stopping time for the filtration $(\mathscr{F}^s_t)_t$, i.e. if $\{\tau \leq t\} \in \mathscr{F}^s_t$ for every $t \geq s$. We define $\mathscr{F}^s_\tau = \{A \in \mathscr{F}^s_\infty; A \cap \{\tau \leq t\} \in \mathscr{F}^s_t \text{ for every } t\}$.

**Definition 6.2** We say that $X$ is *strong Markov* if, for every $x \in E$ and $A \in \mathscr{B}(E)$, for every $s \geq 0$ and for every finite $s$-stopping time $\tau$,

$$\mathrm{P}^{x,s}(X_{t+\tau} \in A \,|\, \mathscr{F}^s_\tau) = p(\tau, t+\tau, X_\tau, A) \quad \mathrm{P}^{x,s}\text{-a.s.} \tag{6.16}$$

It is clear that if $\tau \equiv h \geq s$ then (6.16) boils down to the usual Markov property (6.10). (6.16) is, of course, equivalent to

$$\mathrm{E}^{x,s}[f(X_{t+\tau}) \,|\, \mathscr{F}^s_\tau] = \int_E p(\tau, t+\tau, X_\tau, dy) f(y) \quad \mathrm{P}^{x,s}\text{-a.s.} \tag{6.17}$$

for every bounded Borel function $f$, thanks to the usual arguments of approximating positive Borel functions with increasing sequences of linear combinations of indicator functions.

*Example 6.2* It is immediate to derive from Theorem 3.3 (the stopping theorem for Brownian motion) that a Brownian motion is strong Markov.

Indeed, as the increment $X_{t+\tau} - X_\tau$ is independent of $\mathscr{F}_\tau$ we have, by the freezing Lemma 4.1,

$$\mathrm{E}^x[f(X_{t+\tau})|\mathscr{F}_\tau] = \mathrm{E}^x[f(X_{t+\tau} - X_\tau + X_\tau)|\mathscr{F}_\tau] = \Phi(X_\tau) , \tag{6.18}$$

where

$$\Phi(z) = \mathrm{E}^x[f(X_{t+\tau} - X_\tau + z)] .$$

Now, as $t \mapsto X_{t+\tau} - X_\tau$ is also a Brownian motion, under $\mathrm{P}^x$ $X_{t+\tau} - X_\tau + z$ has law $p(t, z, dy)$, denoting by $p$ the transition function of a Brownian motion. Hence

$$\Phi(z) = \int f(y)\, p(t, z, dy) ,$$

which, going back to (6.18), allows to conclude the proof.

The strong Markov property allows us to compute the conditional expectation, given $\mathscr{F}_\tau$, of a function of the position of the process at a time subsequent to time $\tau$.

It is often very useful to have formulas allowing us to compute the conditional expectation, given $\mathscr{F}_\tau$, of a more complicated function of the path of the process after time $\tau$. This question resembles the one answered by Proposition 6.1 where (the simple) Markov property was extended in a similar way.

Here the situation is more complicated, as it does not seem easy to give a precise meaning to what a "function of the path of the process after time $\tau$" is. In order to do this we are led to the following definitions.

Sometimes the space $\Omega$, on which a Markov process is defined, is a space of *paths*, i.e. $\omega \in \Omega$ is a map $\mathbb{R}^+ \to E$ and the process is defined as $X_t(\omega) = \omega(t)$. This is the case, for instance, for the process constructed in Kolmogorov's theorem, where $\Omega = E^{\mathbb{R}^+}$, or for the canonical space $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ introduced in Sect. 3.2, where $\mathscr{C} = \mathscr{C}(\mathbb{R}^+, \mathbb{R}^m)$. We shall say, in this case, that the Markov process is *canonical*. In this case we can define, for every $t \in \mathbb{R}^+$, the map $\theta_t : \Omega \to \Omega$ through

$$\theta_t\omega(s) = \omega(t + s) .$$

The maps $\theta_t$ are the *translation operators* and we have the relations

$$X_s \circ \theta_t(\omega) = X_s(\theta_t\omega) = X_{t+s}(\omega) .$$

Translation operators are the tool that allows us to give very powerful expressions of the Markov and strong Markov properties.

If $\sigma : \Omega \to \mathbb{R}^+$ is a r.v., we can define the random translation operator $\theta_\sigma$ as

$$(\theta_\sigma \omega)(t) = \omega(t + \sigma(\omega)) .$$

Note that in the time homogeneous case the strong Markov property can be written as

$$\mathrm{E}^x[f(X_{t+\tau})|\mathscr{F}_\tau] = \int_E f(y)p(t, X_\tau, dy)$$

or, with the notations of translation operators,

$$\mathrm{E}^x[f(X_t) \circ \theta_\tau | \mathscr{F}_\tau] = \mathrm{E}^{X_\tau}[f(X_t)] . \tag{6.19}$$

This leads us to the following extension of the strong Markov property for the time homogeneous case. The proof is similar to that of Proposition 6.1.

**Proposition 6.2** Let $X$ be a canonical time homogeneous strong Markov process and $Y$ an $\mathscr{F}_\infty$-measurable r.v., bounded or positive. Then for every finite stopping time $\tau$

$$\mathrm{E}^x[Y \circ \theta_\tau | \mathscr{F}_\tau] = \mathrm{E}^{X_\tau}[Y] . \tag{6.20}$$

*Proof* The proof follows the same pattern as Proposition 6.1, first assuming that $Y$ is of the form $Y = f_1(X_{t_1}) \dots f_m(X_{t_m})$ where $0 \le t_1 < \dots < t_m$ and $f_1, \dots, f_m$ are bounded (or positive) functions. Then (6.20) is immediate if $m = 1$, thanks to (6.19). Assume that (6.20) is satisfied for every $Y$ of this form and for $m - 1$. Then if $Y = f_1(X_{t_1}) \dots f_m(X_{t_m})$ for every $x$ we have $\mathrm{P}^x$-a.s.

$$\begin{aligned}
\mathrm{E}^x[Y \circ \theta_\tau | \mathscr{F}_\tau] &= \mathrm{E}^x[f_1(X_{t_1+\tau}) \dots f_m(X_{t_m+\tau}) | \mathscr{F}_\tau] \\
&= \mathrm{E}^x[f_1(X_{t_1+\tau}) \dots f_{m-1}(X_{t_{m-1}+\tau}) \mathrm{E}^x[f_m(X_{t_m+\tau}) | \mathscr{F}_{t_{m-1}+\tau}] | \mathscr{F}_\tau] \\
&= \mathrm{E}^x[f_1(X_{t_1+\tau}) \dots f_{m-1}(X_{t_{m-1}+\tau}) \mathrm{E}^{X_{t_{m-1}+\tau}}[f_m(X_{t_m})] | \mathscr{F}_\tau] \\
&= \mathrm{E}^x[f_1(X_{t_1+\tau}) \dots \widetilde{f}_{m-1}(X_{t_{m-1}+\tau}) | \mathscr{F}_\tau] ,
\end{aligned}$$

where $\widetilde{f}_{m-1}(x) = f_{m-1}(x)\mathrm{E}^x[f_m(X_{t_m})]$. By the recurrence hypothesis then, going back and using the "simple" Markov property,

$$\begin{aligned}
\mathrm{E}^x[Y \circ \theta_\tau | \mathscr{F}_\tau] &= \mathrm{E}^{X_\tau}[f_1(X_{t_1}) \dots \widetilde{f}_{m-1}(X_{t_{m-1}})] \\
&= \mathrm{E}^{X_\tau}[f_1(X_{t_1}) \dots f_{m-1}(X_{t_{m-1}}) \mathrm{E}^{X_{t_{m-1}}}[f_m(X_{t_m})]]
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{E}^{X_\tau}[f_1(X_{t_1})\dots f_{m-1}(X_{t_{m-1}})\mathrm{E}^{X_\tau}[f_m(X_{t_m})\,|\,\mathscr{F}_{t_{m-1}}]] \\
&= \mathrm{E}^{X_\tau}[\mathrm{E}^{X_\tau}[f_1(X_{t_1})\dots f_{m-1}(X_{t_{m-1}})f_m(X_{t_m})\,|\,\mathscr{F}_{t_{m-1}}]] \\
&= \mathrm{E}^{X_\tau}[f_1(X_{t_1})\dots f_{m-1}(X_{t_{m-1}})f_m(X_{t_m})] \\
&= \mathrm{E}^{X_\tau}[Y] \,.
\end{aligned}
$$

Theorem 1.4 allows us to state that if (6.20) is true for every $Y$ of the form $Y = f_1(X_{t_1})\dots f_m(X_{t_m})$ with $0 \le t_1 \le \dots \le t_m$ and $f_1,\dots,f_m$ bounded or positive functions, then it holds for every bounded (or positive) $\mathscr{F}_\infty$-measurable r.v.

□

*Example 6.3* Let $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ be the canonical realization of an $m$-dimensional Brownian motion. Let $D \subset \mathbb{R}^m$ be a bounded open set, $\tau$ the exit time from $D$ and $\phi : \partial D \to \mathbb{R}$ a bounded measurable function. Let

$$
u(x) = \mathrm{E}^x[\phi(B_\tau)] \,. \tag{6.21}
$$

We see now that $u$ enjoys some important properties.

Let $B_R(x)$ be a ball centered at $x$ with radius $R$ and contained in $D$, $\tau_R$ the exit time from $B_R(x)$. As $\tau > \tau_R$ $\mathrm{P}^x$-a.s (starting at $x$ the process must exit from $B_R(x)$ before leaving $D$), we have

$$
X_\tau = X_\tau \circ \theta_{\tau_R}
$$

as the paths $t \mapsto X_t(\omega)$ and $t \mapsto X_{t+\tau_R(\omega)}(\omega)$ exit $D$ at the same position (not at the same time of course). Hence (6.20) gives

$$
\begin{aligned}
u(x) &= \mathrm{E}^x[\phi(X_\tau)] = \mathrm{E}^x[\phi(X_\tau)\circ\theta_{\tau_R}] \\
&= \mathrm{E}^x[\mathrm{E}^x(\phi(X_\tau)\circ\theta_{\tau_R}\,|\,\mathscr{M}_{\tau_R})] = \mathrm{E}^x[\underbrace{\mathrm{E}^{X_{\tau_R}}[f(X_\tau)]}_{=u(X_{\tau_R})}] \,.
\end{aligned}
$$

Hence

$$
u(x) = \mathrm{E}^x[u(X_{\tau_R})] = \int_{\partial B_R(x)} u(y)\, d\lambda_R(y) \,,
$$

where $\lambda_R$ is the law of $X_{\tau_R}$ under $\mathrm{P}^x$. By Exercise 3.18 $\lambda_R$ is the normalized Lebesgue measure of the spherical surface $\partial B_R(x)$. Therefore the value of $u$ at

(continued)

*Example 6.3* (continued)
$x$ coincides with the mean of $u$ on the surface of the sphere $\partial B_R(x)$ for every $R$ (small enough so that $B_R \subset D$). It is a well-known result in mathematical analysis (see Han and Lin 1997, p. 3, for example) that a bounded measurable function with this property is harmonic, i.e. that it is twice differentiable in $D$ and such that

$$\triangle u = 0 ,$$

where $\triangle$ denotes the Laplace operator

$$\triangle = \sum_{i=1}^{m} \frac{\partial^2}{\partial x_i^2} \cdot$$

Therefore if one wants to look for a solution of the Dirichlet problem

$$\begin{cases} \triangle u = 0 & \text{on } D \\ u_{|\partial D} = \phi \end{cases}$$

the function $u$ of (6.21) is a promising candidate. We shall discuss this kind of problem (i.e., the relation between solutions of PDE problems and stochastic processes) in more detail in Chap. 10.

At this point Brownian motion is our only example of a strong Markov process. We shall snow see that, in fact, a large class of Markov processes enjoy this property.

Let us assume from now on that $E$ is a metric space (we can therefore speak of continuous functions) and that $\mathscr{E} = \mathscr{B}(E)$.

**Definition 6.3** A transition function $p$ is said to enjoy the *Feller property* if, for every fixed $h \geq 0$ and for every bounded continuous function $f : E \to \mathbb{R}$, the map

$$(t, z) \mapsto \int_E f(y) p(t, t + h, z, dy)$$

is continuous. A Markov process $X$ is said to be a *Feller process* if its transition function enjoys Feller's property.

In fact Feller's property is equivalent to saying that if $s_n \to s$ and $x_n \to x$ then

$$p(s_n, s_n + h, x_n, \cdot) \underset{n\to\infty}{\to} p(s, s + h, x, \cdot)$$

weakly. Hence, in a certain sense, that if $x$ is close to $y$ and $s$ close to $t$ then the probability $p(s, s+h, x, \cdot)$ is close to $p(t, t+h, y, \cdot)$.

Of course if the transition function is time homogeneous then

$$\int_E f(y)p(t, t+h, z, dy) = \int_E f(y)p(h, z, dy)$$

and Feller's property requires this to be a continuous function of $z$, for every bounded continuous function $f$ and for every fixed $h > 0$.

*Remark 6.4* Let $p$ be the transition function of Brownian motion, then

$$\int_E f(y)p(h, z, dy) = \frac{1}{(2\pi h)^{m/2}} \int_{\mathbb{R}^m} f(y) e^{-\frac{1}{2h}|z-y|^2}\, dy\ .$$

It is immediate, taking the derivative under the integral sign, that, if $f$ is a bounded measurable function, the left-hand side is actually a $C^\infty$ function of $z$. Therefore Brownian motion is a Feller process.

**Theorem 6.1** Let $X$ be a right-continuous Feller process. Then it is strong Markov.

*Proof* We must prove that, if $\tau$ is a $s$-stopping time,

$$\mathrm{P}^{x,s}(X_{t+\tau} \in A \,|\, \mathscr{F}^s_\tau) = p(\tau, t+\tau, X_\tau, A) \quad \mathrm{P}^{x,s}\text{-a.s.}$$

Note first that, as $X$ is right-continuous, then it is progressively measurable and therefore, by Proposition 3.5, $X_\tau$ is $\mathscr{F}^s_\tau$-measurable. Therefore the right-hand side in the previous equation is certainly $\mathscr{F}^s_\tau$-measurable. We must then prove that if $\Gamma \in \mathscr{F}^s_\tau$ then for every $A \in \mathscr{B}(E)$

$$\mathrm{P}^{x,s}(\{X_{t+\tau} \in A\} \cap \Gamma) = \mathrm{E}^{x,s}(p(\tau, t+\tau, X_\tau, A) 1_\Gamma)\ .$$

Let us assume first that $\tau$ takes at most countably many values $\{t_j\}_j$. Then

$$\begin{aligned} \mathrm{P}^{x,s}(\{X_{t+\tau} \in A\} \cap \Gamma) &= \sum_j \mathrm{P}^{x,s}(\{X_{t+\tau} \in A\} \cap \Gamma \cap \{\tau = t_j\}) \\ &= \sum_j \mathrm{P}^{x,s}(\{X_{t+t_j} \in A\} \cap \Gamma \cap \{\tau = t_j\})\ . \end{aligned} \tag{6.22}$$

As $\Gamma \cap \{\tau = t_j\} \in \mathscr{F}^s_{t_j}$, by the Markov property (6.10),

$$\begin{aligned}\mathrm{P}^{x,s}(\{X_{t+t_j} \in A\} \cap \Gamma \cap \{\tau = t_j\}) &= \mathrm{E}^{x,s}[1_{\{X_{t+t_j}\in A\}} 1_{\Gamma\cap\{\tau=t_j\}}]\\ &= \mathrm{E}^{x,s}\big[\mathrm{E}^{x,s}[1_{\{X_{t+t_j}\in A\}} 1_{\Gamma\cap\{\tau=t_j\}} \,|\, \mathscr{F}^s_{t_j}]\big]\\ = \mathrm{E}^{x,s}\big[1_{\Gamma\cap\{\tau=t_j\}} \mathrm{E}^{x,s}[1_{\{X_{t+t_j}\in A\}} \,|\, \mathscr{F}^s_{t_j}]\big] &= \mathrm{E}^{x,s}\big[1_{\Gamma\cap\{\tau=t_j\}} p(t_j, t+t_j, X_{t_j}, A)\big]\end{aligned}$$

so that, substituting into (6.22)

$$\begin{aligned}\mathrm{P}^{x,s}(\{X_{t+\tau} \in A\} \cap \Gamma) &= \sum_j \mathrm{E}^{x,s}\big[1_{\Gamma\cap\{\tau=t_j\}} p(t_j, t+t_j, X_{t_j}, A)\big]\\ &= \mathrm{E}^{x,s}\big[p(\tau, t+\tau, X_\tau, A) 1_\Gamma\big]\,.\end{aligned}$$

Let now $\tau$ be any finite $s$-stopping time. By Lemma 3.3 there exists a sequence $(\tau_n)_n$ of finite $s$-stopping times, each taking at most countably many values and decreasing to $\tau$. In particular, therefore $\mathscr{F}^s_{\tau_n} \supset \mathscr{F}^s_\tau$. The strong Markov property, already proved for $\tau_n$, and the remark leading to (6.17) guarantee that, for every bounded continuous function $f$ on $E$,

$$\mathrm{E}^{x,s}[f(X_{t+\tau_n}) \,|\, \mathscr{F}^s_{\tau_n}] = \int_E f(y) p(\tau_n, t+\tau_n, X_{\tau_n}, dy)\,.$$

In particular, if $\Gamma \in \mathscr{F}^s_\tau \subset \mathscr{F}^s_{\tau_n}$ then

$$\mathrm{E}^{x,s}[f(X_{t+\tau_n}) 1_\Gamma] = \mathrm{E}^{x,s}\Big[1_\Gamma \int_E f(y) p(\tau_n, t+\tau_n, X_{\tau_n}, dy)\Big]\,.$$

By the right continuity of the paths and the Feller property

$$\int_E f(y) p(\tau_n, t+\tau_n, X_{\tau_n}, dy) \underset{n\to\infty}{\to} \int_E f(y) p(\tau, t+\tau, X_\tau, dy)\,.$$

Since by Lebesgue's theorem we have $\mathrm{E}^{x,s}[f(X_{t+\tau_n})1_\Gamma] \to \mathrm{E}^{x,s}[f(X_{t+\tau})1_\Gamma]$ as $n \to \infty$ (recall that $f$ is bounded and continuous), we have obtained that, for every bounded continuous function $f$,

$$\mathrm{E}^{x,s}[f(X_{t+\tau}) 1_\Gamma] = \mathrm{E}^{x,s}\Big[1_\Gamma \int_E f(y) p(\tau, t+\tau, X_\tau, dy)\Big]\,.$$

By Theorem 1.5 the previous equation holds for every bounded Borel function $f$ and we have proved the statement.

□

Note that the Feller and right continuity assumptions are not needed in the first part of the proof of Theorem 6.1. Therefore the strong Markov property holds for every Markov process when $\tau$ takes at most a countable set of values.

We have already seen (Proposition 4.3) that for a Brownian motion it is always possible to be in a condition where the filtration is right-continuous. In fact, this holds for every right-continuous Feller process. Let $\mathscr{F}^s_{t+} = \bigcap_{\varepsilon>0} \mathscr{F}^s_{t+\varepsilon}$.

**Theorem 6.2** If $X = (\Omega, \mathscr{F}, (\mathscr{F}^s_t)_{t\geq s}, (X_t)_t, (\mathrm{P}^{x,s})_{x,s})$ is a realization of a Feller right-continuous Markov process then $(\Omega, \mathscr{F}, (\mathscr{F}^s_{t+})_{t\geq s}, (X_t)_t, (\mathrm{P}^{x,s})_{x,s})$ is also a realization of a Markov process associated to the same transition function.

*Proof* We must prove the Markov property with respect to $(\mathscr{F}^s_{t+})_t$, i.e. that

$$\mathrm{E}[f(X_{t+h})\,|\,\mathscr{F}^s_{t+}] = \int_E f(y)p(t, t+h, X_t, dy) \tag{6.23}$$

where $f$ can be assumed to be bounded continuous on $E$. A routine application of Theorem 1.5 allows us to extend this relation to every bounded Borel function $f$. For every $\varepsilon > 0$ we have, by the Markov property with respect to $(\mathscr{F}_t)_t$,

$$\mathrm{E}^{x,s}[f(X_{t+h+\varepsilon})\,|\,\mathscr{F}^s_{t+\varepsilon}] = \int_E f(y)p(t+\varepsilon, t+h+\varepsilon, X_{t+\varepsilon}, dy)$$

and, by conditioning both sides with respect to $\mathscr{F}^s_{t+}$, which is contained in $\mathscr{F}^s_{t+\varepsilon}$,

$$\mathrm{E}^{x,s}[f(X_{t+h+\varepsilon})\,|\,\mathscr{F}^s_{t+}] = \mathrm{E}^{x,s}\Big[\int_E f(y)p(t+\varepsilon, t+h+\varepsilon, X_{t+\varepsilon}, dy)\,\Big|\,\mathscr{F}^s_{t+}\Big]\,.$$

As the paths are right-continuous, $f(X_{t+h+\varepsilon}) \to f(X_{t+h})$ a.s. as $\varepsilon \to 0+$. Hence the left-hand side converges to $\mathrm{E}^{x,s}[f(X_{t+h})\,|\,\mathscr{F}^s_{t+}]$ by Lebesgue's theorem for conditional expectations (Proposition 4.2 c)). Thanks to the Feller property, the right-hand side converges to

$$\mathrm{E}^{x,s}\Big[\int_E f(y)p(t, t+h, X_t, dy)\,\Big|\,\mathscr{F}^s_{t+}\Big] = \int_E f(y)p(t, t+h, X_t, dy)\ ,$$

$X_t$ being $\mathscr{F}^s_{t+}$-measurable, which concludes the proof.

□

Recall that a $\sigma$-algebra $\mathscr{G}$ is said to be trivial with respect to a probability P if for every $A \in \mathscr{G}$ the quantity P$(A)$ can only take the values 0 or 1, i.e. if all events in $\mathscr{G}$ are either negligible or almost sure. Recall that we denote by $(\mathscr{G}^s_t)_t$ the natural filtration $\mathscr{G}^s_t = \sigma(X_u, s \leq u \leq t)$. Let $\mathscr{G}^s_{t+} = \bigcap_{\varepsilon>0} \mathscr{G}^s_{t+\varepsilon}$.

**Proposition 6.3 (Blumenthal's 0-1 law)** Let $X$ be a right-continuous Feller process. Then the $\sigma$-algebra $\mathscr{G}^s_{s+}$ is trivial with respect to the probability $\mathrm{P}^{x,s}$ for every $x \in E$, $s \geq 0$.

*Proof* By the previous theorem $X$ is a Markov process with respect to the filtration $(\mathscr{F}^s_{t+})_{t\geq s}$. As $\mathscr{G}^s_{s+} \subset \mathscr{F}^s_{s+}$, if $A \in \mathscr{G}^s_{s+}$, by Proposition 6.1 for $Y = 1_A$ and with $s = t$,

$$1_A = \mathrm{E}^{x,s}(1_A | \mathscr{F}^s_{s+}) = \mathrm{E}^{x,s}(1_A) \quad \text{a.s.}$$

as $X_s = x$ $\mathrm{P}^{x,s}$-a.s. Therefore $\mathrm{E}^{x,s}(1_A) = \mathrm{P}^{x,s}(A)$ can assume only the values 0 or 1. □

*Example 6.4* An application clarifying the otherwise mysterious Blumenthal 0-1 law is the following.

Let $F \subset E$ be a *closed* set and $x \in \partial F$. Let $X$ be a right-continuous Feller process and let $\tau = \inf\{t; t \geq s, X_s \in F^c\}$ be the first exit time from $F$ after time $s$. $X$ being right-continuous, $\tau$ is a stopping time for the filtration $(\mathscr{F}_{t+})_t$, thanks to Proposition 3.7. Then the event $\{\tau = s\}$ belongs to $\mathscr{G}^s_{s+}$:

$$\{\tau = s\} = \{\text{there exists a sequence of times } s_n \searrow s \text{ such that } X_{s_n} \in F^c\}$$

and the event on the right-hand side belongs to $\mathscr{G}^s_{s+\varepsilon}$ for every $\varepsilon > 0$.

Proposition 6.3 states that the probability $\mathrm{P}^{x,s}(\tau = s)$ (i.e. the probability for the process to leave $F$ immediately) can only take the values 0 or 1, i.e. either all paths exit $F$ immediately $\mathrm{P}^{x,s}$-a.s. or no path does (recall that $x \in \partial F$). Figures 6.1 and 6.2 suggest situations where these two possibilities can arise.

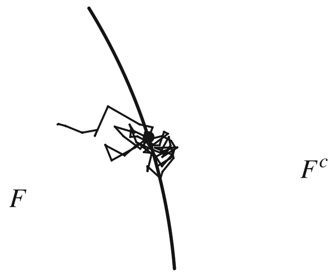


**Fig. 6.1** Typical situation where $\mathrm{P}^{x,s}(\tau = s) = 1$: the boundary $\partial F$ near $x$ is smooth and the oscillations of the path take it immediately outside $F$

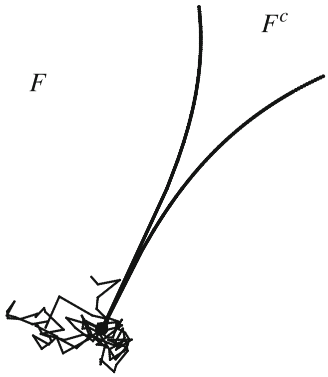


**Fig. 6.2** Typical situation where $\mathrm{P}^{x,s}(\tau = s) = 0$: the set $F^c$ near $x$ is "too thin" to be "caught" immediately by the path

## 6.3 Semigroups, generators, diffusions

Let $p$ be a transition function, which for now we assume to be time homogeneous, on the topological space $E$. To $p$ we can associate the family of linear operators $(T_t)_{t\ge 0}$ defined on $M_b(E)$ (bounded Borel functions on $E$) by

$$T_t f(x) = \int_E f(y)\, p(t, x, dy)$$

or, if $X$ is a realization of the Markov process associated to $p$,

$$T_t f(x) = \mathrm{E}^x[f(X_t)] \ . \tag{6.24}$$

As $\left|\mathrm{E}^x[f(X_t)]\right| \le \|f\|_\infty$, for every $t \ge 0$ the operator $T_t$ is a contraction on $\mathscr{M}_b(E)$, i.e. $\|T_t f\|_\infty \le \|f\|_\infty$. Moreover, we have

$$T_s T_t = T_{s+t} \ ,$$

i.e. the family of operators $(T_t)_t$ is a *semigroup*. Indeed, by the Chapman–Kolmogorov equation,

$$\begin{aligned} T_s T_t f(x) &= \int_E p(s, x, dy) T_t f(y) = \int_E p(s, x, dy) \int_E p(t, y, dz) f(z) \\ &= \int_E p(t+s, x, dz) f(z) = T_{s+t} f(x) \ . \end{aligned}$$

If, moreover, $p$ is Feller, then $T_t$ also operates on $C_b$, i.e. $T_t f \in C_b$ if $f \in C_b$. If $f \in M_b(E)$ let us define, if the limit exists,

$$Af(x) = \lim_{t\to 0+} \frac{1}{t}\,[T_t f(x) - f(x)] \ . \tag{6.25}$$

Let us denote by $\mathscr{D}(A)$ the set of functions $f \in M_b(E)$ for which the limit in (6.25) exists for every $x$. The operator $A$ is defined for $f \in \mathscr{D}(A)$ and is the *infinitesimal generator* of the semigroup $(T_t)_t$ or of the Markov process $X$.

In this section we investigate some properties of the operator $A$ and characterize an important class of Markov processes by imposing some conditions on $A$. We assume from now on that the state space $E$ is an open domain $D \subset \mathbb{R}^m$.
These concepts can be repeated with obvious changes when $p$ is not time homogeneous. In this case we can define the family of operators $(T_{s,t})_{s\le t}$ through

$$T_{s,t}f(x) = \int f(y)\,p(s,t,x,dy) = \mathrm{E}^{x,s}[f(X_t)]$$

and, for $s \le u \le t$, the Chapman–Kolmogorov equation gives

$$T_{s,t} = T_{s,u}T_{u,t}\ .$$

For a time inhomogeneous Markov process, instead of the operator $A$ we are led to consider the family of operators $(A_s)_s$ defined, when the expression is meaningful, by

$$A_s f(x) = \lim_{h\to 0+} \frac{1}{h}\,[T_{s,s+h}f(x) - f(x)]\ .$$

We say that the infinitesimal generator $A$ is *local* when the value of $Af(x)$ depends only on the behavior of $f$ in a neighborhood of $x$, i.e. if, given two functions $f, g$ coinciding in a neighborhood of $x$, if $Af(x)$ is defined, then $Ag(x)$ is also defined and $Af(x) = Ag(x)$.

**Proposition 6.4** Let $B_R(x)$ be the sphere of radius $R$ and centered at $x$ and let us assume that for every $x \in D$ and $R > 0$

$$\lim_{h\to 0+} \frac{1}{h}\,p(t,t+h,x,B_R(x)^c) = 0\ . \tag{6.26}$$

Then $A_t$ is local.

*Proof* Let $f \in \mathscr{D}(A)$ and let $R$ be small enough so that $B_R(x) \subset D$, then

$$\begin{aligned}\frac{1}{h}\,[T_{t,t+h}f(x) - f(x)] &= \frac{1}{h}\Big(\int f(y)\,p(t,t+h,x,dy) - f(x)\Big)\\ &= \frac{1}{h}\int [f(y) - f(x)]\,p(t,t+h,x,dy)\end{aligned}$$

$$= \frac{1}{h}\int_{B_R(x)} [f(y) - f(x)]\,p(t,t+h,x,dy) + \frac{1}{h}\int_{B_R^c} [f(y) - f(x)]\,p(t,t+h,x,dy)\ .$$

As

$$\frac{1}{h}\Big|\int_{B_R(x)^c}[f(y)-f(x)]p(t,t+h,x,dy)\Big| \le \frac{1}{h}\int_{B_R(x)^c}\big|f(y)-f(x)\big|p(t,t+h,x,dy)$$
$$\le \frac{1}{h}\cdot 2\,\|f\|_\infty p(t,t+h,x,B_R(x)^c)$$

we can conclude that the two limits

$$\lim_{h\to 0+}\frac{1}{h}\int[f(y)-f(x)]\,p(t,t+h,x,dy), \quad \lim_{h\to 0+}\frac{1}{h}\int_{B_R(x)}[f(y)-f(x)]\,p(t,t+h,x,dy)$$

either both exist or neither of them exist; moreover, $A_t f(x)$ does not depend on values of $f$ outside of $B_R(x)$, for every $R > 0$.

□

Note that condition (6.26) simply states that the probability of making a transition of length $R$ in a time interval of amplitude $h$ goes to 0 as $h \to 0$ faster than $h$.

The operator $A$ satisfies the *maximum principle* in the following form:

**Proposition 6.5 (The maximum principle)** If $f \in \mathscr{D}(A)$ and $x$ is a point of maximum for $f$, then $Af(x) \le 0$.

*Proof*

$$T_t f(x) = \int f(y)p(t,x,dy) \le \int f(x)p(t,x,dy) = f(x)$$

and therefore $\frac{1}{t}\,[T_t f(x) - f(x)] \le 0$ for every $t > 0$.

□

If $A$ is local, the maximum principle takes the following form. The proof is rather straightforward.

**Proposition 6.6** If $A$ is local, then if $f \in \mathscr{D}(A)$ and $x$ is a point of relative maximum for $f$ then $Af(x) \le 0$.

If $E = \mathbb{R}^m$ the following proposition provides a way of computing the generator $A$, at least for a certain class of functions.

**Proposition 6.7** Let us assume that for every $R > 0$, $t \geq 0$ the limits

$$\lim_{h\to 0+} \frac{1}{h}\, p(t, t+h, x, B_R(x)^c) = 0 \tag{6.27}$$

$$\lim_{h\to 0+} \frac{1}{h} \int_{B_R(x)} (y_i - x_i)\, p(t, t+h, x, dy) = b_i(x, t) \tag{6.28}$$

$$\lim_{h\to 0+} \frac{1}{h} \int_{B_R(x)} (y_i - x_i)(y_j - x_j)\, p(t, t+h, x, dy) = a_{ij}(x, t) \tag{6.29}$$

exist. Then the matrix $a(x, t)$ is positive semidefinite for every $x, t$ and, if

$$L_t = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x, t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x, t)\frac{\partial}{\partial x_i}\ ,$$

we have, for every function $f \in C^2(\mathbb{R}^m) \cap C_b(\mathbb{R}^m)$,

$$\lim_{h\to 0+} \frac{1}{h}\, [T_{t,t+h}f(x) - f(x)] = L_t f(x) \qquad \text{for every } x \in D\ .$$

*Proof* By (6.27), for $f \in C^2(\mathbb{R}^m) \cap C_b(\mathbb{R}^m)$

$$\begin{aligned}
\lim_{h\to 0+} \frac{1}{h}\, [T_{t,t+h}f(x) - f(x)] &= \lim_{h\to 0+} \frac{1}{h} \int [f(y) - f(x)]\, p(t, t+h, x, dy) \\
&= \lim_{h\to 0+} \frac{1}{h} \int_{B_R(x)} [f(y) - f(x)]\, p(t, t+h, x, dy)\ .
\end{aligned}$$

Replacing $f$ with its Taylor development to the second order,

$$f(y) = f(x) + \sum_{i=1}^m \frac{\partial f}{\partial x_i}(x)\,(y_i - x_i) + \frac{1}{2}\sum_{i,j=1}^m \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\,(y_i - x_i)(y_j - x_j) + o(|x-y|^2)$$

we find

$$\lim_{h\to 0+} \frac{1}{h}\, [T_{t,t+h}f(x) - f(x)] = L_t f(x) + \lim_{h\to 0+} \frac{1}{h} \int_{B_R(x)} o(|x-y|^2)\, p(t, t+h, x, dy)\ .$$

Let us check that the rightmost limit is equal to 0. As the above computation holds for every $R$, let $R$ be small enough so that $|\frac{o(r)}{r}| \le \varepsilon$ for every $0 < r < R$. Then

$$\overline{\lim_{h\to 0+}} \frac{1}{h} \int_{B_R(x)} |o(|x-y|^2)|\, p(t,t+h,x,dy) \le \overline{\lim_{h\to 0+}}\, \varepsilon \frac{1}{h} \int_{B_R(x)} |x-y|^2\, p(t,t+h,x,dy)$$

$$= \overline{\lim_{h\to 0+}}\, \varepsilon \frac{1}{h} \int_{B_R(x)} \sum_{i=1}^m (x_i-y_i)^2\, p(t,t+h,x,dy) = \varepsilon \sum_{i=1}^m a_{ii}(x,t)$$

and the conclusion comes from the arbitrariness of $\varepsilon$. We still have to prove that the matrix $a(x,t)$ is positive semidefinite for every $x,t$. Let $\theta \in \mathbb{R}^m$ and $f \in C^2(\mathbb{R}^m) \cap C_b(\mathbb{R}^m)$ be a function such that $f(y) = -\langle \theta, y-x\rangle^2$ for $y$ in a neighborhood of $x$. Then, as the first derivatives of $f$ vanish at $x$, whereas $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = -2\theta_i\theta_j$,

$$-\langle a(x,t)\theta,\theta\rangle = -\sum_{i,j=1}^m a_{ij}(x,t)\theta_i\theta_j = L_t f(x) = \lim_{h\to 0+} \frac{1}{h}\left[T_{t,t+h}f(x) - f(x)\right].$$

But $x$ is a relative maximum for $f$, hence, by Proposition 6.6, $-\langle a(x,t)\theta,\theta\rangle = L_t f(x) \le 0$.

□

*Example 6.5* Let us compute the infinitesimal generator of an $m$-dimensional Brownian motion $B$. Let us first establish that it is local by checking condition (6.26) of Lemma 6.4. We have

$$p(h,x,B_R(x)^c) = \frac{1}{(2\pi h)^{m/2}} \int_{|x-y|\ge R} \mathrm{e}^{-\frac{1}{2h}|x-y|^2}\, dy$$

$$= \frac{1}{(2\pi h)^{m/2}} \int_{|y|\ge R} \mathrm{e}^{-\frac{1}{2h}|y|^2}\, dy = \frac{1}{(2\pi)^{m/2}} \int_{|z|\ge R/\sqrt{h}} \mathrm{e}^{-\frac{1}{2}|z|^2}\, dz$$

$$= \frac{1}{(2\pi)^{m/2}} \int_{|z|\ge R/\sqrt{h}} \mathrm{e}^{-\frac{1}{4}|z|^2}\mathrm{e}^{-\frac{1}{4}|z|^2}\, dz$$

$$\le \mathrm{e}^{-\frac{1}{4h}|R|^2} \frac{1}{(2\pi)^{m/2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\frac{1}{4}|z|^2}\, dz \le \mathrm{const}\, \mathrm{e}^{-\frac{1}{4h}|R|^2}$$

so that condition (6.26) is satisfied. Moreover,

$$\frac{1}{h}\int_{|y-x|\le R} (y_i - x_i)\, p(h,x,dy) = \frac{1}{h}\int_{|y-x|\le R} \frac{y_i - x_i}{(2\pi h)^{m/2}} \mathrm{e}^{-\frac{1}{2h}|x-y|^2}\, dy$$

(continued)

*Example 6.5* (continued)

$$= \frac{1}{h}\int_{|z|\le R} \frac{z_i}{(2\pi h)^{m/2}}\,\mathrm{e}^{-|z|^2/2h}\,dz = 0$$

as we are integrating an odd function on a symmetric set. Also, with the change of variable $z = \frac{y-x}{\sqrt{h}}$ so that $dz = \frac{dy}{h^{m/2}}$,

$$\begin{aligned}
&\lim_{h\to 0+}\frac{1}{h}\int_{|y-x|\le R}(y_i - x_i)(y_j - x_j)p(h,x,dy)\\
&= \lim_{h\to 0+}\frac{1}{h}\int_{|y-x|\le R}\frac{1}{(2\pi h)^{m/2}}\,(y_i - x_i)(y_j - x_j)\mathrm{e}^{-\frac{1}{2h}\,|x-y|^2}\,dy\\
&= \frac{1}{(2\pi)^{m/2}}\int_{|z|\le R/\sqrt{h}} z_i z_j \mathrm{e}^{-\frac{1}{2}|z|^2}\,dw \underset{h\to 0}{\to} \frac{1}{(2\pi)^{m/2}}\int_{\mathbb{R}^m} z_i z_j \mathrm{e}^{-\frac{1}{2}|z|^2}\,dw = \delta_{ij}
\end{aligned}$$

as we recognize in the last integral nothing else than the covariance matrix of an $N(0,I)$-distributed r.v. Going back to the notations of Proposition 6.7, we have $b_i = 0, a_{ij} = \delta_{ij}$. Therefore the Brownian motion has an infinitesimal generator given, for every function $f \in C^2 \cap C_b$, by

$$Lf = \tfrac{1}{2}\triangle f = \frac{1}{2}\sum_{i=1}^m \frac{\partial^2 f}{\partial x_i^2}\,\cdot$$

**Definition 6.4** Let

$$L_t = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x,t)\frac{\partial}{\partial x_i} \tag{6.30}$$

be a differential operator on $\mathbb{R}^m$ such that the matrix $a(x,t)$ is positive semidefinite. A *diffusion* associated to the generator $L_t$ is a Markov process such that

a) it is strongly Markovian,
b) for every function $f \in C_K^2(\mathbb{R}^m)$, if $(T_{s,t})_{t\ge s}$ is the semigroup associated to $X$,

$$T_{s,t}f(x) = f(x) + \int_s^t T_{s,u}L_u f(x)\,du\,. \tag{6.31}$$

We have, up to now, a candidate example of diffusion, which is Brownian motion. We know that it enjoys the Feller property (Example 6.4) and, being continuous, it is also strongly Markovian by Theorem 6.1. Its generator is the Laplace operator divided by 2 and we shall see soon that b) of Definition 6.4, which is more or less equivalent to the fact that $\frac{1}{2}\triangle$ is its generator, also holds.

It is natural now to wonder whether, given a more general differential operator $L_t$ as in (6.30), an associated diffusion process exists and is unique. In the next chapters we shall answer this question; the method will be the construction of a new process starting from Brownian motion, which is itself a diffusion.

## Exercises

**6.1** (p. 511) (When is a Gaussian process also Markov?) Let $X$ be a centered $m$-dimensional Gaussian process and let us denote by $K^{i,j}_{s,t} = \mathrm{E}(X^i_s X^j_t)$ its *covariance function*. Let us assume that the matrix $K_{s,s}$ is invertible for every $s$.

a) Prove that, for every $s, t$, $s \le t$, there exist a matrix $C_{t,s}$ and a Gaussian r.v. $Y_{t,s}$ independent of $X_s$ such that

$$X_t = C_{t,s}X_s + Y_{t,s}\ .$$

What is the conditional law of $X_t$ given $X_s = x$?

b) Prove that $X$ is a Markov process with respect to its natural filtration $(\mathscr{G}_t)_t$ if and only if for every $u \le s \le t$

$$K_{t,u} = K_{t,s}K^{-1}_{s,s}K_{s,u}\ . \tag{6.32}$$

b) Use the freezing Lemma 4.1; (6.32) is equivalent to requiring that $Y_{t,s}$ is orthogonal to $X_u$ for every $u \le s$.

**6.2** (p. 512) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and let

$$X_t = \int_0^t B_u\, du\ .$$

a) Is $(X_t)_t$ a martingale of the filtration $(\mathscr{F}_t)_t$?
b) Show that $(X_t)_t$ is a Gaussian process and compute its covariance function $K_{s,t} = \mathrm{Cov}(X_s, X_t)$. Show that $(X_t)_t$ *is not* a Markov process
c) Compute

$$\mathrm{Cov}(X_t, B_s)$$

(be careful to distinguish whether $s \le t$ or $s \ge t$).
d) Show that $(B_t, X_t)_t$ is a Markov process.

**6.3** (p. 514) Let $g, h : \mathbb{R}^+ \to \mathbb{R}^+$ be continuous functions with $g$ increasing and both vanishing at most at 0. Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and

$$X_t = h(t) B_{g(t)} \ .$$

a1) Prove that $X$ is a Markov process (with respect to which filtration?) and compute its transition function. Is it time homogeneous?

a2) Assume that $h(t) = \frac{\sqrt{t}}{\sqrt{g(t)}}$. What is the law of $X_t$ for a fixed $t$? Is $X$ a Brownian motion?

b) Assume that $\lim_{t\to+\infty} g(t) = +\infty$. What can be said about

$$\overline{\lim_{t\to+\infty}} \frac{X_t}{\sqrt{2g(t)h^2(t)\log\log g(t)}} \ ?$$

**6.4** (p. 515) Let $B$ be a Brownian motion, $\lambda > 0$, and let

$$X_t = \mathrm{e}^{-\lambda t} B_{\mathrm{e}^{2\lambda t}} \ .$$

a) Prove that $X$ is a time homogeneous Markov process and determine its transition function.

b1) Compute the law of $X_t$. What do you observe?

b2) Prove that for every $t_1 < t_2 < \cdots < t_m$ and $h > 0$ the two r.v.'s $(X_{t_1}, \dots, X_{t_m})$ and $(X_{t_1+h}, \dots, X_{t_m+h})$ have the same distribution (i.e. that $X$ is *stationary*).

c) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (Z_t)_t, (\mathrm{P}^x)_x)$ be a realization of the Markov process associated to the transition function computed in a). What is the law of $Z_t$ under $\mathrm{P}^x$? Prove that, whatever the starting point $x$,

$$Z_t \ \underset{t\to+\infty}{\overset{\mathscr{L}}{\to}} \ N(0,1) \ . \tag{6.33}$$

**6.5** (p. 516) (Brownian bridge again). We shall make use here of Exercises 4.15 and 6.1. Let $B$ be a real Brownian motion and let $X_t = B_t - tB_1$. We have already dealt with this process in Exercise 4.15, it is a Brownian bridge.

a) What is its covariance function $K_{s,t} = \mathrm{E}(X_s X_t)$? Compute the law of $X_t$ and the conditional law of $X_t$ given $X_s = x$, $0 \le s < t \le 1$.

b) Prove that $X$ is a non-homogeneous Markov process and compute its transition function.

- The generator of the Brownian bridge can be computed using Proposition 6.7. It will be easier to compute it later using a different approach, see Exercise 9.2.

**6.6** (p. 517) (Time reversal of a Brownian motion) Let $B$ be a Brownian motion and for $0 \le t \le 1$ let

$$X_t = B_{1-t} \ .$$

a) Prove that $X$ is a Gaussian process and compute its transition function.
b) Prove that $(X_t)_{t \le 1}$ is an inhomogeneous Markov process with respect to its natural filtration and compute its transition function.

**6.7** (p. 517) Let $X$ be an $\mathbb{R}^m$-valued Markov process associated to a transition function $p$ such that, for some $\beta > 0$, $\varepsilon > 0$ and $c > 0$,

$$\int_{\mathbb{R}^m} |x-y|^\beta \, p(s,t,x,dy) \le c\,|t-s|^{m+\varepsilon}$$

for every $x \in \mathbb{R}^m$ and $s \le t$. Then $X$ has a continuous version and its generator is local.

**6.8** (p. 518) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ be an $m$-dimensional time homogeneous diffusion and let us assume that its transition function $p$ satisfies the relation

$$p(t,x,A) = p(t,0,A-x)$$

as in Remark 6.3.

a) Prove that, for every bounded Borel function $f$,

$$\int f(y)\,p(t,x,dy) = \int_{\mathbb{R}^m} f(x+y)\,p(t,0,dy) \ . \tag{6.34}$$

b) Prove that the generator $L$ of $X$ has constant coefficients.

**6.9** (p. 519) (A transformation of Markov processes)

a) Let $p$ be a time homogeneous transition function on a measurable space $(E, \mathscr{E})$, $h : E \to \mathbb{R}$ a strictly positive measurable function and $\alpha$ a number such that for every $x \in \mathbb{R}$, $t > 0$,

$$\int_E h(y)\,p(t,x,dy) = \mathrm{e}^{\alpha t} h(x) \ . \tag{6.35}$$

Show that

$$p^h(t,x,A) = \frac{\mathrm{e}^{-\alpha t}}{h(x)} \int_A h(y)\,p(t,x,dy)$$

is a Markov transition function.

b) Let us assume that (6.35) holds and let us denote by $L$ and $L^h$ the generators of the semigroups $(T_t)_t$ and $(T^h_t)_t$ associated to $p$ and to $p^h$ respectively. Show that if $f \in \mathscr{D}(L)$, then $g = \frac{f}{h}$ belongs to $\mathscr{D}(L^h)$ and express $L^h g$ in terms of $Lf$.

c) Let us assume, moreover, that $E = \mathbb{R}^m$ and that $p$ is the Markov transition function of a diffusion of generator

$$L = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x)\frac{\partial}{\partial x_i}\,\cdot \tag{6.36}$$

Prove that, if $h$ is twice differentiable, $C^2_K \subset \mathscr{D}(L^h)$ and compute $L^h g$ for $g \in C^2_K$.

d) Let $p$ be the transition function of an $m$-dimensional Brownian motion (see Example 6.1) and let $h(x) = \mathrm{e}^{\langle v,x\rangle}$, where $v \in \mathbb{R}^m$ is some fixed vector. Show that (6.35) holds for some $\alpha$ to be determined and compute $L^h g$ for $g \in C^2_K$.

**6.10** (p. 520) Let $E$ be a topological space. We say that a time homogeneous $E$-valued Markov process $X$ associated to the transition function $p$ admits an *invariant* (or *stationary*) *measure* $\mu$ if $\mu$ is a $\sigma$-finite measure on $(E, \mathscr{B}(E))$ such that, for every compactly supported bounded Borel function $f$,

$$\int T_t f(x)\, d\mu(x) = \int_E f(x)\, d\mu(x) \tag{6.37}$$

for every $t$. Recall that $T_t f(x) = \int_E f(y)\, p(t, x, dy)$. If, moreover, $\mu$ is a probability, we say that $X$ admits an *invariant* (or *stationary*) *distribution*.

a) Prove that $\mu$ is a stationary distribution if and only if, if $X_0$ has law $\mu$ then $X_t$ also has law $\mu$ for every $t \ge 0$.

b) Prove that the Lebesgue measure of $\mathbb{R}^m$ is invariant for $m$-dimensional Brownian motion.

c) Prove that, if for every $x \in E$, the transition function of $X$ is such that

$$\lim_{t\to+\infty} p(t, x, A) = 0 \tag{6.38}$$

for every bounded Borel set $A \subset E$, then $X$ cannot have an invariant probability. Deduce that the $m$-dimensional Brownian motion cannot have an invariant probability.

d) Prove that if $X$ is a Feller process and there exists a probability $\mu$ on $(E, \mathscr{B}(E))$ such that, for every $x \in E$,

$$\lim_{t\to+\infty} p(t, x, \cdot) = \mu \tag{6.39}$$

in the sense of the weak convergence of probabilities, then $\mu$ is a stationary distribution.

**6.11** (p. 521) (When is a function of a Markov process also Markov?) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be a Markov process associated to the transition function $p$ and with values in $(E, \mathscr{E})$. Let $(G, \mathscr{G})$ be a measurable space and $\Phi : E \to G$ a surjective measurable map. Is it true that $Y = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (Y_t)_t, \mathrm{P})$ with $Y_t = \Phi(X_t)$ is also a Markov process? In this exercise we investigate this question. The answer is no, in general: the Markov property might be lost for $Y$. We have seen an example of this phenomenon in Exercise 6.2.

a) Prove that if the map $\Phi$ is bijective then $Y$ is a Markov process and determine its transition function.

b) Let us assume that for every $A \in \mathscr{G}$ the transition function $p$ of $X$ satisfies the relation

$$p(s, t, x, \Phi^{-1}(A)) = p(s, t, z, \Phi^{-1}(A)) \ . \tag{6.40}$$

for every $x, z$ such that $\Phi(x) = \Phi(z)$; let, for $\xi \in G$ and $A \in \mathscr{G}$, $q(s, t, \xi, A) = p(s, t, x, \Phi^{-1}(A))$, where $x$ is any element of $E$ such that $\Phi(x) = \xi$. )

b1) Prove that for every bounded measurable function $f : G \to \mathbb{R}$

$$\int_G f(y)\, q(s, t, \xi, dy) = \int_E f \circ \Phi(z)\, p(s, t, x, dz) \ , \tag{6.41}$$

where $x \in E$ is any element such that $\Phi(x) = \xi$.

b2) Show that, if $Y_t = \Phi(X_t)$, $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (Y_t)_t, \mathrm{P})$ is a Markov process associated to $q$, which therefore turns out to be a transition function (note that it is a Markov process with respect to the same filtration).

c) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion and $z \in \mathbb{R}^m$. Let $Y_t = |X_t - z|$; $Y_t$ is therefore the distance of $X_t$ from $z$. Show that $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (Y_t)_t, \mathrm{P})$ is a Markov process.

c) Note that if $Z \sim N(0, I)$, then the transition function of the Brownian motion can be written as

$$p(s, x, A) = \mathrm{P}(\sqrt{s} Z \in A - x) \ . \tag{6.42}$$

Then use the rotational invariance of the distributions $N(0, I)$.

- Condition (6.40) is called Dynkin's criterion. It is a simple sufficient condition for $Y$ to be a Markov process (not necessary however).
- The Markov process introduced in c) is called *Bessel process* of dimension $m$. It is a diffusion on $\mathbb{R}^+$ which will also be the object of Exercise 8.24.

# Chapter 7
# The Stochastic Integral

## 7.1 Introduction

Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a (continuous) standard Brownian motion fixed once and for all: the aim of this chapter is to give a meaning to expressions of the form

$$\int_0^T X_s(\omega)\, dB_s(\omega) \tag{7.1}$$

where the integrand $(X_s)_{0\le s\le T}$ is a process enjoying certain properties to be specified. As already remarked in Sect. 3.3, this cannot be done path by path as the function $t \mapsto B_t(\omega)$ does not have finite variation a.s. The r.v. (7.1) is a *stochastic integral* and it will be a basic tool for the construction and the investigation of new processes. For instance, once the stochastic integral is defined, it is possible to consider a *stochastic differential equation*: it will be something of the form

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dB_t\ , \tag{7.2}$$

where $b$ and $\sigma$ are suitable functions. To solve it will mean to find a process $(X_t)_t$ such that for every $t \ge 0$

$$X_t = X_0 + \int_0^t b(X_s)\, ds + \int_0^t \sigma(X_s)\, dB_s\ ,$$

which is well defined, once the stochastic integral is given a rigorous meaning. One can view the solution of (7.2) as a model to describe the behavior of objects following the ordinary differential equation $\dot x_t = b(x_t)$, but whose evolution is also influenced by random perturbations represented by the term $\sigma(X_t)\, dB_t$.

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_7

The idea of the construction of the stochastic integral is rather simple: imitating the definition the Riemann integral, consider first the integral of piecewise constant processes, i.e. of the form

$$X_t = \sum_{k=0}^{n-1} X_k 1_{[t_k,t_{k+1}[}(t) , \tag{7.3}$$

where $a = t_0 < t_1 < \dots < t_n = b$. For these it is natural to set

$$\int_a^b X_t\, dB_t = \sum_{k=0} X_k(B_{t_{k+1}} - B_{t_k}) .$$

Then with processes as in (7.3) we shall approximate more general situations. Although the idea is simple, technical difficulties will not be in short supply, as we shall see.

Once the integral is defined we shall be concerned with the investigation of the newly defined process $t \mapsto \int_0^t X_s\, dB_s$. The important feature is that, under suitable assumptions, it turns out to be a martingale.

## 7.2 Elementary processes

Let us define first the spaces of processes that will be the integrands of the stochastic integral.

We denote by $M^p_{loc}([a,b])$ the space of the equivalence classes of real-valued *progressively measurable* processes $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{a\le t\le b}, (X_t)_{a\le t\le b}, \mathrm{P})$ such that

$$\int_a^b |X_s|^p\, ds < +\infty \qquad \text{a.s.} \tag{7.4}$$

By $M^p([a,b])$ we conversely denote the subspace of $M^p_{loc}([a,b])$ of the processes such that

$$\mathrm{E}\Big(\int_a^b |X_s|^p\, ds\Big) < +\infty . \tag{7.5}$$

$M^p_{loc}([0,+\infty[)$ (resp. $M^p([0,+\infty[)$) will denote the space of the processes $(X_t)_t$ such that $(X_t)_{0\le t\le T}$ lies in $M^p_{loc}([0,T])$ (resp. $M^p([0,T])$) for every $T > 0$.

Speaking of equivalence classes, we mean that we identify two processes $X$ and $X'$ whenever they are indistinguishable, which can also be expressed by saying that

$$\int_a^b |X_s - X'_s|\, ds = 0 \qquad \text{a.s.}$$

*Remarks 7.1*

a) It is immediate that a continuous and adapted process $(X_t)_t$ belongs to $M^p_{loc}([a,b])$ for every $p \geq 0$. Indeed continuity, in addition to the fact of being adapted, implies progressive measurability (Proposition 2.1). Moreover, (7.4) is immediate as $s \mapsto |X_s(\omega)|^p$, being continuous, is automatically integrable on every bounded interval. By the same argument, multiplying a process in $M^p_{loc}([a,b])$ by a bounded progressively measurable process again gives rise to a process in $M^p_{loc}([a,b])$.

b) If $X \in M^p_{loc}$ (resp. $M^p$) and $\tau$ is a stopping time of the filtration $(\mathscr{F}_t)_t$, then the process $t \mapsto X_t 1_{\{t<\tau\}}$ also belongs to $M^p_{loc}$ (resp. $M^p$). Indeed the process $t \mapsto 1_{\{t<\tau\}}$ is itself progressively measurable (it is adapted and right-continuous) and moreover, as it vanishes for $t > \tau$,

$$\int_a^b |X_s|^p 1_{\{s<\tau\}}\, ds = \int_a^{b\wedge\tau} |X_s|^p\, ds \leq \int_a^b |X_s|^p\, ds$$

so that condition (7.4) (resp. (7.5)) is also satisfied.

c) It is not difficult to prove that $M^2$ is a Hilbert space with respect to the scalar product

$$\langle X, Y\rangle_{M^2} = \mathrm{E}\Big(\int_a^b X_s Y_s\, ds\Big) = \int_a^b \mathrm{E}(X_s Y_s)\, ds\,,$$

being a closed subspace of $L^2(\Omega \times \mathbb{R}^+, \mathrm{P} \otimes \lambda_{[a,b]})$, where $\lambda_{[a,b]}$ denotes the Lebesgue measure of $[a,b]$.

d) Quite often we shall deal with the problem of checking that a given progressively measurable process $(X_t)_t$ belongs to $M^2([a,b])$. This is not difficult. Indeed, by Fubini's theorem

$$\mathrm{E}\Big(\int_a^b |X_s|^2\, ds\Big) = \int_a^b \mathrm{E}(|X_s|^2)\, ds$$

and therefore it is enough to ascertain that $s \mapsto \mathrm{E}(|X_s|^2)$ is, for instance, continuous. The Brownian motion itself is in $M^2$, for instance.

Among the elements of $M^p_{loc}$ there are, in particular, those of the form

$$X_t(\omega) = \sum_{i=0}^{n-1} X_i(\omega) 1_{[t_i,t_{i+1}[}(t) , \tag{7.6}$$

where $a = t_0 < t_1 < \cdots < t_n = b$ and, for every $i$, $X_i$ is a real $\mathscr{F}_{t_i}$-measurable r.v. The condition that $X_i$ is $\mathscr{F}_{t_i}$-measurable is needed for the process to be adapted and ensures progressive measurability (a process as in (7.6) is clearly right-continuous). We shall call these processes *elementary*. As

$$\mathrm{E}\Big[\int_a^b X_t^2\,dt\Big] = \mathrm{E}\Big[\sum_{i=0}^{n-1} X_i^2(t_{i+1}-t_i)\Big] = \sum_{i=0}^{n-1} \mathrm{E}(X_i^2)(t_{i+1}-t_i)$$

we have that $X \in M^2([a,b])$ if and only if the r.v.'s $X_i$ are square integrable.

**Definition 7.1** Let $X \in M^2_{loc}([a,b])$ be an elementary process as in (7.6). The *stochastic integral* of $X$ (with respect to $B$), denoted $\int_a^b X_t\,dB_t$, is the r.v.

$$\sum_{i=0}^{n-1} X_i(B_{t_{i+1}} - B_{t_i}) . \tag{7.7}$$

It is easy to see that the stochastic integral is linear in $X$. Moreover,

**Lemma 7.1** If $X$ is an elementary process in $M^2([a,b])$, then

$$\mathrm{E}\Big(\int_a^b X_t\,dB_t \,\Big|\, \mathscr{F}_a\Big) = 0 ,$$

$$\mathrm{E}\Big[\Big(\int_a^b X_t\,dB_t\Big)^2 \,\Big|\, \mathscr{F}_a\Big] = \mathrm{E}\Big[\int_a^b X_t^2\,dt \,\Big|\, \mathscr{F}_a\Big] .$$

In particular, the stochastic integral of an elementary process of $M^2([a,b])$ is a centered square integrable r.v. and

$$\mathrm{E}\Big[\Big(\int_a^b X_t\,dB_t\Big)^2\Big] = \mathrm{E}\Big[\int_a^b X_t^2\,dt\Big] . \tag{7.8}$$

*Proof* Let $X_t = \sum_{i=0}^{n-1} X_i 1_{[t_i,t_{i+1}[}(t)$; as $X_i$ is square integrable and $\mathscr{F}_{t_i}$-measurable and $B_{t_{i+1}} - B_{t_i}$ is independent of $\mathscr{F}_{t_i}$, we have $\mathrm{E}[X_i(B_{t_{i+1}} - B_{t_i})\,|\,\mathscr{F}_{t_i}] = X_i\mathrm{E}[B_{t_{i+1}} - B_{t_i}] = 0$ and therefore

$$\mathrm{E}(X_i(B_{t_{i+1}} - B_{t_i})\,|\,\mathscr{F}_a) = \mathrm{E}\big[\mathrm{E}[X_i(B_{t_{i+1}} - B_{t_i})\,|\,\mathscr{F}_{t_i}]\,|\,\mathscr{F}_a\big] = 0\;,$$

from which the first relation follows. We also have

$$\mathrm{E}\Big[\Big(\int_a^b X_t\,dB_t\Big)^2\,\Big|\,\mathscr{F}_a\Big] = \mathrm{E}\Big[\sum_{i,j=0}^{n-1} X_iX_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\,\Big|\,\mathscr{F}_a\Big]\,. \tag{7.9}$$

Note first that, as $X_i$ is $\mathscr{F}_{t_i}$-measurable and therefore independent of $B_{t_{i+1}} - B_{t_i}$, the r.v.'s $X_i^2(B_{t_{i+1}} - B_{t_i})^2$ are integrable being the product of integrable independent r.v.'s (Proposition 1.3). Therefore the r.v. $X_iX_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})$ is also integrable, being the product of the square integrable r.v.'s $X_i(B_{t_{i+1}} - B_{t_i})$ and $X_j(B_{t_{j+1}} - B_{t_j})$. We have, for $j > i$,

$$\begin{aligned}
&\mathrm{E}\big[X_iX_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\,|\,\mathscr{F}_a\big]\\
&= \mathrm{E}\big[\mathrm{E}\big(X_iX_j(B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\,|\,\mathscr{F}_{t_j}\big)\,\big|\,\mathscr{F}_a\big]\\
&= \mathrm{E}[X_iX_j(B_{t_{i+1}} - B_{t_i})\underbrace{\mathrm{E}(B_{t_{j+1}} - B_{t_j}\,|\,\mathscr{F}_{t_j})}_{=0}\,|\,\mathscr{F}_a] = 0\;.
\end{aligned}$$

Therefore in (7.9) the contribution of the terms with $i \neq j$ vanishes. Moreover,

$$\mathrm{E}\big(X_i^2(B_{t_{i+1}} - B_{t_i})^2\,|\,\mathscr{F}_a\big) = \mathrm{E}\big[\mathrm{E}\big[X_i^2(B_{t_{i+1}} - B_{t_i})^2\,|\,\mathscr{F}_{t_i}]\,|\,\mathscr{F}_a\big] = \mathrm{E}[X_i^2(t_{i+1} - t_i)\,|\,\mathscr{F}_a]\;.$$

Therefore, going back to (7.9),

$$\mathrm{E}\Big[\Big(\int_a^b X_t\,dB_t\Big)^2\,\Big|\,\mathscr{F}_a\Big] = \mathrm{E}\Big[\sum_{i=0}^{n-1} X_i^2(t_{i+1} - t_i)\,\Big|\,\mathscr{F}_a\Big] = \mathrm{E}\Big[\int_a^b X_t^2\,dt\,\Big|\,\mathscr{F}_a\Big]\,.$$

□

In the proof of Lemma 7.1, note the crucial role played by the assumption that the r.v.'s $X_i$ are $\mathscr{F}_{t_i}$ measurable, which is equivalent to requiring that $X$ is progressively measurable.

## 7.3 The stochastic integral

In order to define the stochastic integral for processes in $M^p$ we need an approximation result stating that a process in $M^p$ can be suitably approximated with elementary processes.

Given a function $f \in L^p([a,b])$, $p \geq 1$, and an equi-spaced grid $\pi_n = \{a = t_0 < t_1 < \cdots < t_n = b\}$ such that $t_{i+1} - t_i = \frac{1}{n}(b-a)$, let

$$G_n f = \sum_{i=0}^{n-1} f_i 1_{[t_i, t_{i+1}[} ,$$

where $f_i = 0$ if $i = 0$ and

$$f_i = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(s)\, ds = \frac{n}{b-a} \int_{t_{i-1}}^{t_i} f(s)\, ds$$

is the average of $f$ on the interval *at the left of* $t_i$. As, by Jensen's inequality,

$$|f_i|^p = \Big| \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(s)\, ds \Big|^p \leq \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} |f(s)|^p\, ds = \frac{n}{b-a} \int_{t_{i-1}}^{t_i} |f(s)|^p\, ds ,$$

we have

$$\int_a^b |G_n f(s)|^p\, ds = \sum_{i=0}^{n-1} \frac{|f_i|^p}{n} (b-a) \leq \sum_{i=1}^{n} \int_{t_i}^{t_{i+1}} |f(s)|^p\, ds = \int_a^b |f(s)|^p\, ds \tag{7.10}$$

and therefore also $G_n f \in L^p([a,b])$. As $G_{2^n} f = G_{2^n}(G_{2^{n+1}} f)$, the same argument leading to (7.10) gives

$$\int_a^b |G_{2^n} f(s)|^p\, ds \leq \int_a^b |G_{2^{n+1}} f(s)|^p\, ds . \tag{7.11}$$

$G_n f$ is a step function and let us to prove that

$$\lim_{n\to\infty} \int_a^b |G_n f(s) - f(s)|^p\, ds = 0 . \tag{7.12}$$

Actually, this is immediate if $f$ is continuous, owing to uniform continuity: let, for a fixed $\varepsilon$, $n$ be such that $|f(x) - f(y)| \leq \varepsilon$ if $|x - y| \leq \frac{2}{n}(b-a)$ and $\int_a^{t_1} |f(s)|^p\, ds \leq \varepsilon^p$.

For $i \geq 1$, if $s \in [t_i, t_{i+1}[$ and $u \in [t_{i-1}, t_i[$, then $|u - s| \leq \frac{2}{n}(b-a)$ so that $|f(s) - f(u)| \leq \varepsilon$ and also $|G_n f(s) - f(s)| \leq \varepsilon$, as $G_n f(s)$ is the average of $f$ on $[t_{i-1}, t_i[$. Then we have

$$\int_a^b |G_n f(s) - f(s)|^p\, ds = \int_a^{t_1} |f(s)|^p\, ds + \int_{t_1}^b |G_n f(s) - f(s)|^p\, ds \leq \varepsilon^p (1 + b - a)$$

which proves (7.12) if $f$ is continuous. In order to prove (7.12) for a general function $f \in L^p([a,b])$ one has just to recall that continuous functions are dense in $L^p$, the details are left to the reader.

**Lemma 7.2** Let $X \in M^p_{loc}([a,b])$, then there exists a sequence of elementary processes $(X_n)_n \subset M^p_{loc}([a,b])$ such that

$$\lim_{n\to\infty} \int_a^b |X_t - X_n(t)|^p\,dt = 0 \qquad \text{a.s.} \tag{7.13}$$

If $X \in M^p([a,b])$ then there exists a sequence of elementary processes $(X_n)_n \subset M^p([a,b])$ such that

$$\lim_{n\to\infty} \mathrm{E}\Big(\int_a^b |X_t - X_n(t)|^p\,dt\Big) = 0\,. \tag{7.14}$$

Moreover, the elementary processes $X_n$ can be chosen in such a way that $n \mapsto \int_a^b |X_n|^p\,ds$ is increasing.

*Proof* If $X \in M^p_{loc}$ then $t \mapsto X_t(\omega)$ is a function in $L^p$ a.s. Let us define $X_n = G_n X$. Such a process $X_n$ is adapted as on the interval $[t_i, t_{i+1}[$ it takes the value $\frac{1}{t_i - t_{i+1}} \int_{t_{i-1}}^{t_i} X_s\,ds$, which is $\mathscr{F}_{t_i}$-measurable, as explained in Remark 2.2. Finally, (7.13) follows from (7.12).

Let $X \in M^p$ and let $\pi_n$ be the equi-spaced grid defined above. We know already that $X_n = G_n X$ is adapted and moreover, thanks to (7.10),

$$\int_a^b |X_n(s)|^p\,ds \le \int_a^b |X(s)|^p\,ds$$

so that $X_n \in M^p([a,b)]$ and, moreover, by (7.10) we have for every $n$

$$\int_a^b |X_t - X_n(t)|^p\,dt \le 2^{p-1}\Big(\int_a^b |X_t|^p\,dt + \int_a^b |X_n(t)|^p\,dt\Big) \le 2^p \int_a^b |X_t|^p\,dt\,,$$

hence we can take the expectation in (7.13) and obtain (7.14) using Lebesgue's theorem.

Finally the last statement is a consequence of (7.11). □

We can now define the stochastic integral for every process $X \in M^2([a,b])$: (7.8) states that the stochastic integral is an isometry between the elementary processes of $M^2([a,b])$ and $L^2(\mathrm{P})$. Lemma 7.2 says that these elementary processes are dense in $M^2([a,b])$, so that the isometry can be extended to the whole $M^2([a,b])$, thus defining the stochastic integral for every $X \in M^2([a,b])$.

In practice, in order to compute the stochastic integral of a process $X \in M^2([a, b])$, by now, we must do the following: we must first determine the sequence $G_n X$ of elementary processes approximating $X$ in the sense of (7.14), and then compute the limit

$$\lim_{n\to\infty} \int_a^b X_n(s)\, dB_s$$

in the $L^2$ sense. This procedure does not look appealing, but soon we shall see (in the next chapter) other ways of computing the stochastic integral. This is similar to what happens with the ordinary Riemann integral: first one defines the integral through an approximation with step functions and then finds much simpler and satisfactory ways of making the actual computation with the use of primitives.

Let us investigate the first properties of the stochastic integral. The following extends to general integrands the properties already known for the stochastic integral of elementary processes.

**Theorem 7.1** If $X \in M^2([a, b])$ then

$$\mathrm{E}\Big(\int_a^b X_t\, dB_t \,\Big|\, \mathscr{F}_a\Big) = 0\,, \tag{7.15}$$

$$\mathrm{E}\Big[\Big(\int_a^b X_t\, dB_t\Big)^2 \,\Big|\, \mathscr{F}_a\Big] = \mathrm{E}\Big[\int_a^b X_t^2\, dt \,\Big|\, \mathscr{F}_a\Big]\,. \tag{7.16}$$

In particular,

$$\mathrm{E}\Big[\Big(\int_a^b X_t\, dB_t\Big)^2\Big] = \mathrm{E}\Big[\int_a^b X_t^2\, dt\Big]\,. \tag{7.17}$$

*Proof* Let $(X_n)_n$ be a sequence of elementary processes approximating $X$ as stated in Lemma 7.2. Then we know by Lemma 7.1 that (7.15) is true for the stochastic integral of processes of $X_n$ and we can take the limit as $n \to \infty$ as $\int_a^b X_n(t)\, dB_t \to \int_a^b X_t\, dB_t$ in $L^2$. As for (7.16) we have

$$\mathrm{E}\Big[\Big(\int_a^b X_n(t)\, dB_t\Big)^2 \,\Big|\, \mathscr{F}_a\Big] = \mathrm{E}\Big[\int_a^b X_n^2(t)\, dt \,\Big|\, \mathscr{F}_a\Big]\,.$$

We can take the limit as $n \to \infty$ at the left-hand side as $\int_a^b X_n(t)\, dB_t \to \int_a^b X_t\, dB_t$ in $L^2$ hence $\big(\int_a^b X_n(t)\, dB_t\big)^2 \to \big(\int_a^b X_t\, dB_t\big)^2$ in $L^1$ and we can use Remark 4.3. As for the right-hand side we can assume that $n \mapsto \int_a^b |X_n(t)|^2\, dt$ is increasing and then use Beppo Levi's theorem for the conditional expectation, Proposition 4.2 a). $\square$

*Remark 7.1* If $X, Y \in M^2([a,b])$, then

$$E\Big(\int_a^b X_s\,dB_s \cdot \int_a^b Y_s\,dB_s\Big) = \int_a^b E(X_sY_s)\,ds\,. \tag{7.18}$$

Indeed, it is sufficient to apply (7.17) to $X+Y$ and $X-Y$ and then develop and subtract:

$$\Big(\int_a^b (X_s+Y_s)\,dB_s\Big)^2 = \Big(\int_a^b X_s\,dB_s\Big)^2 + \Big(\int_a^b Y_s\,dB_s\Big)^2 + 2\int_a^b X_s\,dB_s \cdot \int_a^b Y_s\,dB_s$$
$$\Big(\int_a^b (X_s-Y_s)\,dB_s\Big)^2 = \Big(\int_a^b X_s\,dB_s\Big)^2 + \Big(\int_a^b Y_s\,dB_s\Big)^2 - 2\int_a^b X_s\,dB_s \cdot \int_a^b Y_s\,dB_s$$

and therefore

$$\begin{aligned} &4E\Big(\int_a^b X_s\,dB_s \cdot \int_a^b Y_s\,dB_s\Big) \\ &= E\Big[\Big(\int_a^b (X_s+Y_s)\,dB_s\Big)^2\Big] - E\Big[\Big(\int_a^b (X_s-Y_s)\,dB_s\Big)^2\Big] \\ &= E\Big[\int_a^b (X_s+Y_s)^2\,ds\Big] - E\Big[\int_a^b (X_s-Y_s)^2\,ds\Big] = 4E\Big[\int_a^b X_sY_s\,ds\Big]\,. \end{aligned}$$

*Examples 7.1*

a) Note that the motion Brownian $B$ itself belongs to $M^2$. Is it true that

$$\int_0^1 B_s\,dB_s = \frac{1}{2}B_1^2\ ? \tag{7.19}$$

Of course no, as

$$E\Big(\int_0^1 B_s\,dB_s\Big) = 0 \quad \text{whereas} \quad E(B_1^2) = 1\,.$$

We shall see, however, that (7.19) becomes true if an extra term is added.

b) Also $(B_t^2)_t$ belongs to $M^2$. What is the value of

$$E\Big[\int_0^1 B_s\,dB_s \int_0^1 B_s^2\,dB_s\Big]\ ?$$

(continued)

*Examples 7.1* (continued)
Easy

$$\mathrm{E}\Big[\int_0^1 B_s\,dB_s \int_0^1 B_s^2\,dB_s\Big] = \mathrm{E}\Big[\int_0^1 B_s B_s^2\,ds\Big] = \int_0^1 \mathrm{E}[B_s^3]\,ds = 0\,.$$

c) What is the value of

$$\mathrm{E}\Big[\int_0^2 B_s\,dB_s \int_1^3 B_s\,dB_s\Big]\;?$$

The two stochastic integrals are over different intervals. If we set

$$X_s = \begin{cases} B_s & \text{if } 0 \le s \le 2 \\ 0 & \text{if } 2 < s \le 3 \end{cases} \qquad\qquad Y_s = \begin{cases} 0 & \text{if } 0 \le s < 1 \\ B_s & \text{if } 1 \le s \le 3 \end{cases}$$

then

$$\begin{aligned} \mathrm{E}\Big[\int_0^2 B_s\,dB_s \int_1^3 B_s\,dB_s\Big] &= \mathrm{E}\Big[\int_0^3 X_s\,dB_s \int_0^3 Y_s\,dB_s\Big] \\ &= \int_0^3 \mathrm{E}[X_s Y_s]\,ds = \int_1^2 s\,ds = \frac{3}{2}\,\cdot \end{aligned}$$

Of course, this argument can be used in general in order to compute the expectation of the product of stochastic integrals over different intervals. In particular, the expectation of the product of two stochastic integrals of processes of $M^2$ over disjoint intervals vanishes.

*Remark 7.2* As the stochastic integral is an isometry $M^2 \to L^2(\Omega, \mathscr{F}, \mathrm{P})$, we can ask whether every r.v. of $L^2(\Omega, \mathscr{F}, \mathrm{P})$ can be obtained as the stochastic integral of a process of $M^2$. One sees immediately that this cannot be as, by (7.15), every stochastic integral of a process in $M^2$ has zero mean. We shall see, however, in Sect. 12.3 that *if* $(\mathscr{F}_t)_t$ is the natural augmented filtration of the Brownian motion $B$, then every r.v. $Z \in L^2(\Omega, \mathscr{F}_T, \mathrm{P})$ can be represented in the form

$$Z = c + \int_0^T X_s\,dB_s$$

with $X \in M^2([0,T])$ and $c \in \mathbb{R}$. This representation has deep and important applications that we shall see later.

By the isometry property of the stochastic integral (7.17)

$$\mathrm{E}\Big[\Big(\int_a^b X_n(s)\,dB_s-\int_a^b X_s\,dB_s\Big)^2\Big]=\mathrm{E}\Big[\Big(\int_a^b (X_n(s)-X_s)\,dB_s\Big)^2\Big]$$
$$=\mathrm{E}\Big(\int_a^b |X_n(t)-X_t|^2\,dt\Big)$$

so that

**Corollary 7.1** If $X_n, X \in M^2([a,b])$ are processes such that

$$\lim_{n\to\infty}\mathrm{E}\Big(\int_a^b |X_n(t)-X_t|^2\,dt\Big)=0$$

then

$$\int_a^b X_n(s)\,dB_s \quad \overset{L^2}{\underset{n\to\infty}{\to}} \quad \int_a^b X_s\,dB_s\ .$$

Note that we cannot say that the value of the integral at $\omega$ depends only on the paths $t \mapsto X_t(\omega)$ and $t \mapsto B_t(\omega)$, as the integral is not defined pathwise. However, we have the following

**Theorem 7.2** Let $X, Y \in M^2([a,b])$ and $A \in \mathscr{F}$ such that $X_t = Y_t$ on $A$ for every $t \in [a,b]$. Then on $A$ we have a.s.

$$\int_a^b X_t\,dB_t = \int_a^b Y_t\,dB_t\ .$$

*Proof* Let $(X_n)_n$ and $(Y_n)_n$ be the sequences of elementary processes that respectively approximate $X$ and $Y$ in $M^2([a,b])$ constructed on p. 186, i.e. $X_n = G_nX$, $Y_n = G_nY$. A closer look at the definition of $G_n$ shows that $G_nX$ and $G_nY$ also coincide on $A$ for every $t \in [a,b]$ and for every $n$, and therefore also $\int_a^b X_n\,dB_t$ and

$\int_a^b Y_n\,dB_t$. By definition we have

$$\int_a^b X_n(t)\,dB_t \quad \underset{n\to\infty}{\overset{L^2}{\to}} \quad \int_a^b X(t)\,dB_t$$

$$\int_a^b Y_n(t)\,dB_t \quad \underset{n\to\infty}{\overset{L^2}{\to}} \quad \int_a^b Y(t)\,dB_t$$

and then just recall that $L^2$-convergence implies a.s. convergence for a subsequence. □

## 7.4 The martingale property

Let $X \in M^2([0,T])$, then the restriction of $X$ to $[0,t]$, $t \le T$ also belongs to $M^2([0,t])$ and we can consider its integral $\int_0^t X_s\,dB_s$; let, for $0 \le t \le T$, the real-valued process $I$ be defined as

$$I(t) = \int_0^t X_s\,dB_s\;.$$

It is clear that

a) if $t > s$, $I(t) - I(s) = \int_s^t X_u\,dB_u$,
b) $I(t)$ is $\mathscr{F}_t$-measurable for every $t$. Indeed if $(X_n)_n$ is the sequence of elementary processes that approximates $X$ in $M^2([0,t])$ and $I_n(t) = \int_0^t X_n(s)\,dB_s$, then it is immediate that $I_n(t)$ is $\mathscr{F}_t$-measurable, given the definition of the stochastic integral for the elementary processes. Since $I_n(t) \to I(t)$ in $L^2$ as $n \to \infty$, there exists a subsequence $(n_k)_k$ such that $I_{n_k}(t) \to I(t)$ a.s. as $k \to \infty$; therefore $I(t)$ is also $\mathscr{F}_t$-measurable (remember that we assume that the Brownian motion $B$ is standard so that, in particular, $\mathscr{F}_t$ contains the negligible events of $\mathscr{F}$ and changing an $\mathscr{F}_t$-measurable r.v. on a negligible event still produces an $\mathscr{F}_t$-measurable r.v.)

   Note that if

$$X_t = \sum_{i=0}^{n-1} X_i 1_{[t_i,t_{i+1}[}(t)$$

   is an elementary process then we can write

$$I(t) = \int_0^t X_s\,dB_s = \sum_{i=0}^{n-1} X_i(B_{t_{i+1}\wedge t} - B_{t_i\wedge t})\;. \tag{7.20}$$

In particular, if $X$ is an elementary process then $I$ is a continuous process adapted to the filtration $(\mathscr{F}_t)_t$.

A key property of the process $I$ is the following.

**Theorem 7.3** If $X \in M^2([0,T])$, then $I$ is an $(\mathscr{F}_t)_t$-square integrable martingale. Moreover, $I$ has a continuous version and its associated increasing process is

$$A_t = \int_0^t X_s^2\, ds\,. \tag{7.21}$$

*Proof* If $t > s$ we have, thanks to Theorem 7.1,

$$\mathrm{E}[I(t) - I(s)\,|\,\mathscr{F}_s] = \mathrm{E}\Big(\int_s^t X_u\, dB_u \,\Big|\, \mathscr{F}_s\Big) = 0 \quad \text{a.s.}$$

and therefore we have the martingale relation

$$\mathrm{E}[I(t)\,|\,\mathscr{F}_s] = I(s) + \mathrm{E}[I(t) - I(s)\,|\,\mathscr{F}_s] = I(s) \quad \text{a.s.}$$

Theorem 7.1 also states that $I(t)$ is square integrable. In order to check that $A$ is the associated increasing process to the martingale $I$, we need to verify that $Z_t = I(t)^2 - A(t)$ is a martingale. With the decomposition $I(t)^2 - A(t) = [I(s) + (I(t) - I(s))]^2 - A(s) - (A(t) - A(s))$ and remarking that $I(s)$ and $A(s)$ are $\mathscr{F}_s$-measurable,

$$\begin{aligned}
&\mathrm{E}[I(t)^2 - A(t)\,|\,\mathscr{F}_s]\\
&= I(s)^2 - A(s) + \mathrm{E}\big[2I(s)(I(t) - I(s)) + (I(t) - I(s))^2 - (A(t) - A(s))\,|\,\mathscr{F}_s\big]\\
&= I(s)^2 - A(s) + 2I(s)\underbrace{\mathrm{E}[I(t) - I(s)\,|\,\mathscr{F}_s]}_{=0} + \mathrm{E}\big[(I(t) - I(s))^2 - (A(t) - A(s))\,|\,\mathscr{F}_s\big]
\end{aligned}$$

which allows us to conclude that the process defined in (7.21) is the associate increasing process as, thanks to (7.16), we have a.s.

$$\begin{aligned}
\mathrm{E}[(I(t) - I(s))^2\,|\,\mathscr{F}_s] &= \mathrm{E}\Big[\Big(\int_s^t X_u\, dB_u\Big)^2 \,\Big|\, \mathscr{F}_s\Big] = \mathrm{E}\Big[\int_s^t X_u^2\, du \,\Big|\, \mathscr{F}_s\Big]\\
&= \mathrm{E}[A(t) - A(s)\,|\,\mathscr{F}_s]\,.
\end{aligned}$$

Let us prove the existence of a continuous version. We already know that this is true for elementary processes. Let $(X_n)_n$ be a sequence of elementary processes in

$M^2([0,T])$ such that

$$\lim_{n\to\infty} \mathrm{E}\Big(\int_0^T |X_n(t)-X(t)|^2\,dt\Big) = 0\,.$$

Let $I_n(t) = \int_0^t X_n(s)\,dB_s$. The idea of the existence of a continuous version is simple: just find a subsequence $(n_k)_k$ such that the sequence of processes $(I_{n_k})_k$ converges uniformly a.s. The limit, $J$ say, will therefore be a continuous process. As we know already that

$$\lim_{k\to\infty} I_{n_k}(t) = I(t) \quad \text{in } L^2$$

we shall have that $J(t) = I(t)$ a.s., so that $J$ will be the required continuous version. In order to prove the uniform convergence of the subsequence $(I_{n_k})_k$ we shall write

$$I_{n_k}(t) = I_{n_1}(t) + \sum_{i=1}^{n_k}(I_{n_i}(t) - I_{n_{i-1}}(t)) \tag{7.22}$$

and prove that $\sup_{0\le t\le T}|I_{n_i}(t) - I_{n_{i-1}}(t)|$ is the general term of a convergent series.

As $(I_n(t) - I_m(t))_t$ is a square integrable continuous martingale, by the maximal inequality (5.16) applied to the supermartingale $(-|I_n(t)-I_m(t)|^2)_t$ and for $\lambda = \varepsilon^2$,

$$\begin{aligned}
\mathrm{P}\Big(\sup_{0\le t\le T}|I_n(t)-I_m(t)| > \varepsilon\Big) &= \mathrm{P}\Big(\inf_{0\le t\le T} -|I_n(t)-I_m(t)|^2 < -\varepsilon^2\Big)\\
\le \frac{1}{\varepsilon^2}\,\mathrm{E}\big[|I_n(T)-I_m(T)|^2\big] &= \frac{1}{\varepsilon^2}\,\mathrm{E}\Big[\Big(\int_0^T \big(X_n(s)-X_m(s)\big)\,dB_s\Big)^2\Big]\\
&= \frac{1}{\varepsilon^2}\,\mathrm{E}\Big(\int_0^T |X_n(s)-X_m(s)|^2\,ds\Big)\,.
\end{aligned}$$

As $(X_n)_n$ is a Cauchy sequence in $M^2$, the right-hand side can be made arbitrarily small for $n$ and $m$ large. If we choose $\varepsilon = 2^{-k}$ then there exists an increasing sequence $(n_k)_k$ such that, for every $m > n_k$,

$$\mathrm{P}\Big(\sup_{0\le t\le T}|I_{n_k}(t)-I_m(t)| > 2^{-k}\Big) \le 2^{2k}\mathrm{E}\Big(\int_0^T |X_{n_k}(s)-X_m(s)|^2\,ds\Big) \le \frac{1}{k^2}$$

and therefore

$$\mathrm{P}\Big(\sup_{0\le t\le T}|I_{n_k}(t)-I_{n_{k+1}}(t)| > 2^{-k}\Big) \le \frac{1}{k^2}\,\cdot$$

But $\frac{1}{k^2}$ is the general term of a convergent series and, by the Borel–Cantelli lemma,

$$\mathrm{P}\Big(\sup_{0\le t\le T}|I_{n_k}(t)-I_{n_{k+1}}(t)|>2^{-k}\text{ infinitely many times}\Big)=0\,.$$

Therefore a.s. we have eventually

$$\sup_{0\le t\le T}|I_{n_k}(t)-I_{n_{k+1}}(t)|\le 2^{-k}\,.$$

Therefore the series in (7.22) converges uniformly, which concludes the proof. □

From now on, by $I(t)$ or $\int_0^t X_s\,dB_s$ we shall understand the continuous version.

Theorem 7.3 allows us to apply to the stochastic integral $(I(t))_t$ all the nice properties of square integrable continuous martingales that we have pointed out in Sect. 5.6. First of all that the paths of $I$ do not have finite variation (Theorem 5.15) unless they are a.s. constant, which can happen only if $X_s = 0$ a.s for every $s$.

The following inequalities also hold

$$\begin{aligned}&\mathrm{P}\Big(\sup_{0\le t\le T}\Big|\int_0^t X_s\,dB_s\Big|>\lambda\Big)\le\frac{1}{\lambda^2}\,\mathrm{E}\Big(\int_0^T X_s^2\,ds\Big),\qquad\text{for every }\lambda>0\\&\mathrm{E}\Big[\sup_{0\le t\le T}\Big(\int_0^t X_s\,dB_s\Big)^2\Big]\le 4\mathrm{E}\Big[\int_0^T X_s^2\,ds\Big]\,.\end{aligned}\tag{7.23}$$

In fact the first relation follows from the maximal inequality (5.16) applied to the supermartingale $M_t = -\big|\int_0^t X_s\,dB_s\big|^2$: (5.16) states that, for a continuous supermartingale $M$,

$$\lambda\mathrm{P}\big(\inf_{0\le t\le T}M_t\le-\lambda\big)\le\mathrm{E}[|M_T|]$$

so that

$$\mathrm{P}\Big(\sup_{0\le t\le T}\Big|\int_0^t X_s\,dB_s\Big|>\lambda\Big)=\mathrm{P}\Big(\sup_{0\le t\le T}\Big|\int_0^t X_s\,dB_s\Big|^2\ge\lambda^2\Big)$$

$$=\mathrm{P}\Big(\inf_{0\le t\le T}-\Big|\int_0^t X_s\,dB_s\Big|^2\le-\lambda^2\Big)\le\frac{1}{\lambda^2}\,\mathrm{E}\Big(\Big|\int_0^T X_s\,dB_s\Big|^2\Big)=\frac{1}{\lambda^2}\,\mathrm{E}\Big(\int_0^T X_s^2\,ds\Big)\,.$$

As for the second one, we have, from Doob's inequality (Theorem 5.12),

$$\mathrm{E}\Big[\sup_{0\le t\le T}\Big(\int_0^t X_s\,dB_s\Big)^2\Big] \le 4\sup_{0\le t\le T}\mathrm{E}\Big[\Big(\int_0^t X_s\,dB_s\Big)^2\Big]$$
$$= 4\sup_{0\le t\le T}\mathrm{E}\Big[\int_0^t X_s^2\,ds\Big] = 4\mathrm{E}\Big[\int_0^T X_s^2\,ds\Big]\,.$$

As a consequence, we have the following strengthening of Corollary 7.1.

**Theorem 7.4** Let $X_n, X \in M^2([0,T])$ and let us assume that

$$\mathrm{E}\Big[\int_0^T |X_n(s) - X_s|^2\,ds\Big] \underset{n\to\infty}{\to} 0$$

then

$$\lim_{n\to\infty}\mathrm{E}\Big(\sup_{0\le t\le T}\Big|\int_0^t X_n(s)\,dB_s - \int_0^t X_s\,dB_s\Big|^2\Big) = 0\,.$$

Obviously if $f \in L^2([0,T])$, then $f \in M^2([0,T])$. In this case the stochastic integral enjoys an important property.

**Proposition 7.1** If $f \in L^2([0,T])$ then the process

$$I_t = \int_0^t f(s)\,dB_s \tag{7.24}$$

is Gaussian.

*Proof* Let us prove that, for every choice of $0 \le s_1 < \cdots < s_m \le T$, the r.v. $\mathscr{I} = (I(s_1),\dots,I(s_m))$ is Gaussian. This fact is immediate if $f$ is piecewise constant: if $f(t) = \sum_{i=1}^n \lambda_i 1_{[t_{i-1},t_i[}$, then

$$I_s = \sum_{i=1}^n \lambda_i (B_{t_i\wedge s} - B_{t_{i-1}\wedge s})$$

and the vector $\mathscr{I}$ is therefore Gaussian, being a linear function of the r.v.'s $B_{t_i\wedge s_j}$ that are jointly Gaussian. We know that there exists a sequence $(f_n)_n \subset L^2([0,T])$ of piecewise constant functions converging to $f$ in $L^2([0,T])$ and therefore in $M^2$.

Let $I_n(t) = \int_0^t f_n(s)\,dB_s$, $I(t) = \int_0^t f(s)\,dB_s$. Then, by the isometry property of the stochastic integral (Theorem 7.1), for every $t$,

$$\mathrm{E}(|I_n(t) - I_t|^2) = \int_0^t (f_n(u) - f(u))^2\,ds \le \|f_n - f\|_2^2 \underset{n\to\infty}{\overset{L^2}{\to}} 0$$

so that $I_n(t) \to_{n\to\infty} I_t$ in $L^2$ and therefore, for every $t$, $I_t$ is Gaussian by the properties of the Gaussian r.v.'s under $L^2$ convergence (Proposition 1.9). Moreover, if $0 \le s_1 \le \dots \le s_m$, then $\mathscr{I}_n = (I_n(s_1), \dots, I_n(s_m))$ is a jointly Gaussian r.v. As it converges, for $n \to \infty$, to $\mathscr{I} = (I(s_1), \dots, I(s_m))$ in $L^2$, the random vector $\mathscr{I}$ is also jointly Gaussian, which concludes the proof. □

*Remark 7.3* Proposition 7.1 implies in particular that

$$\mathrm{E}\Big[\exp\Big(\int_0^t f(s)\,dB_s\Big)\Big] = \exp\Big(\frac{1}{2}\int_0^t f(s)^2\,ds\Big)\,. \tag{7.25}$$

Indeed we recognize on the left-hand side the mean of the exponential of a centered Gaussian r.v. (i.e. its Laplace transform computed at $\theta = 1$), which is equal to the exponential of its variance divided by 2 (see Exercise 1.6).

If $\tau$ is a stopping time of the filtration $(\mathscr{F}_t)_t$ then, thanks to Corollary 5.6, the process

$$I(t \wedge \tau) = \int_0^{t\wedge\tau} X_s\,dB_s$$

is a $(\mathscr{F}_t)_t$-martingale. In particular, $\mathrm{E}\big(\int_0^{t\wedge\tau} X_s\,dB_s\big) = 0$. The following statement is often useful.

**Theorem 7.5** Let $\tau$ be a stopping time of the filtration $(\mathscr{F}_t)_t$, with $\tau \le T$. Then if $X \in M^2([0,T])$ also $(X_t 1_{\{t<\tau\}})_t \in M^2([0,T])$ and

$$\int_0^\tau X_s\,dB_s = \int_0^T X_s 1_{\{s<\tau\}}\,dB_s \quad \text{a.s.} \tag{7.26}$$

*Proof* $(X_t 1_{\{t<\tau\}})_t \in M^2([0,T])$ because the process $t \to 1_{\{t<\tau\}}$ is bounded, adapted and right-continuous hence progressively measurable.

(7.26) is immediate if $\tau$ takes at most finitely many values $t_1, \dots, t_n$. Actually we have a.s.

$$\int_0^\tau X_s\,dB_s = \sum_{k=1}^n 1_{\{\tau=t_k\}} \int_0^{t_k} X_s\,dB_s \overset{\downarrow}{=} \sum_{k=1}^n 1_{\{\tau=t_k\}} \int_0^{t_k} X_s 1_{\{s<\tau\}}\,dB_s$$
$$= \int_0^T X_s 1_{\{s<\tau\}}\,dB_s\;,$$

where the equality marked with $\downarrow$ is a consequence of Theorem 7.2, as on $\{\tau = t_k\}$ the processes $(X_s)_s$ and $(X_s 1_{\{s<\tau\}})_s$ coincide on the time interval $[0, t_k[$.

For a general stopping time $\tau$, let $(\tau_n)_n$ be a sequence of stopping times taking at most finitely many values and decreasing to $\tau$ (see Lemma 3.3); possibly replacing $\tau_n$ with $\tau_n \wedge T$ we can assume $\tau_n \le T$ for every $n$. Then, the paths of the stochastic integral being continuous,

$$\lim_{n\to\infty} \int_0^{\tau_n} X_s\,dB_s = \int_0^\tau X_s\,dB_s \qquad \text{a.s.} \tag{7.27}$$

Moreover,

$$\int_0^T \big|X_s 1_{\{s<\tau_n\}} - X_s 1_{\{s<\tau\}}\big|^2\,ds = \int_0^T X_s^2 1_{\{\tau\le s<\tau_n\}}\,ds\;.$$

As $\lim_{n\to\infty} 1_{\{\tau\le s<\tau_n\}} = 1_{\{\tau=s\}}$, by Lebesgue's theorem

$$\lim_{n\to\infty} \int_0^T \big|X_s 1_{\{s<\tau_n\}} - X_s 1_{\{s<\tau\}}\big|^2\,ds = \int_0^T X_s^2 1_{\{\tau=s\}}\,ds = 0$$

and again by Lebesgue's theorem

$$\lim_{n\to\infty} \mathrm{E}\Big[\int_0^T \big|X_s 1_{\{s<\tau_n\}} - X_s 1_{\{s<\tau\}}\big|^2\,ds\Big] = 0\;.$$

Therefore, by Corollary 7.1,

$$\lim_{n\to\infty} \int_0^T X_s 1_{\{s<\tau_n\}}\,dB_s = \int_0^T X_s 1_{\{s<\tau\}}\,dB_s \qquad \text{in } L^2\;.$$

This relation together with (7.27) allows us to conclude the proof of the lemma. □

Let $X \in M^2([0,T])$ and let $\tau_1$ and $\tau_2$ be stopping times with $\tau_1 \le \tau_2 \le T$.

The following properties follow easily from the stopping theorem, Theorem 7.5, and the martingale property, Theorem 7.3, of the stochastic integral.

$$\mathrm{E}\Big(\int_0^{\tau_1} X_t\,dB_t\Big) = 0 \tag{7.28}$$

$$\mathrm{E}\Big[\Big(\int_0^{\tau_1} X_t\,dB_t\Big)^2\Big] = \mathrm{E}\Big(\int_0^{\tau_1} X_t^2\,dt\Big) \tag{7.29}$$

$$\mathrm{E}\Big(\int_0^{\tau_2} X_t\,dB_t \,\Big|\, \mathscr{F}_{\tau_1}\Big) = \int_0^{\tau_1} X_t\,dB_t, \qquad \text{a.s.} \tag{7.30}$$

$$\mathrm{E}\Big[\Big(\int_{\tau_1}^{\tau_2} X_t\,dB_t\Big)^2 \,\Big|\, \mathscr{F}_{\tau_1}\Big] = \mathrm{E}\Big(\int_{\tau_1}^{\tau_2} X_t^2\,dt \,\Big|\, \mathscr{F}_{\tau_1}\Big), \qquad \text{a.s.} \tag{7.31}$$

Let $[a, b]$ be an interval such that $X_t(\omega) = 0$ for almost every $a \le t \le b$. Is it true that $t \mapsto I_t$ is constant on $[a, b]$?

Let $\tau = \inf\big\{t; t > a, \int_a^t X_s^2\,ds > 0\big\}$ with the understanding that $\tau = b$ if $\{\ \} = \emptyset$. Then by Lemma 7.5 and (7.23)

$$\mathrm{E}\Big(\sup_{a\le t\le b} |I_{t\wedge\tau} - I_a|^2\Big) = \mathrm{E}\Big(\sup_{a\le t\le b} \Big|\int_a^{t\wedge\tau} X_u\,dB_u\Big|^2\Big)$$

$$= \mathrm{E}\Big(\sup_{a\le t\le b} \Big|\int_a^{t} X_u 1_{\{u<\tau\}}\,dB_u\Big|^2\Big) \le 4\mathrm{E}\Big(\int_a^{b} X_u^2 1_{\{u<\tau\}}\,du\Big) = 4\mathrm{E}\Big(\int_a^{\tau} X_u^2\,du\Big) = 0\,,$$

where the last equality follows from the fact that $X_u(\omega) = 0$ for almost every $u \in ]a, \tau(\omega)[$. Therefore there exists a negligible event $N_{a,b}$ such that, for $\omega \notin N_{a,b}$, $t \mapsto I_t(\omega)$ is constant on $]a, \tau[$.

**Proposition 7.2** Let $X \in M^2([0, T])$. Then there exists a negligible event $N$ such that, for every $\omega \notin N$ and for every $0 \le a < b \le T$, if $X_t(\omega) = 0$ for almost every $t \in ]a, b[$, then $t \mapsto I_t(\omega)$ is constant on $]a, b[$.

*Proof* We know already that for every $r, q \in \mathbb{Q} \cap [0, T]$ there exists a negligible event $N_{r,q}$ such that if $X_t(\omega) = 0$ for almost every $t \in ]r, q[$, then $t \mapsto I_t(\omega)$ is constant on $]r, q[$.

Let $N$ be the union of the events $N_{r,q}$, $r, q \in \mathbb{Q} \cap [0, T]$. $N$ is the negligible event we were looking for, as if $\omega \notin N$, then if $X_t(\omega) = 0$ on $]a, b[$, $t \mapsto I_t(\omega)$ is constant a.s. on every interval $]r, s[\subset]a, b[$ having rational endpoints and therefore also on $]a, b[$. □

## 7.5 The stochastic integral in $M^2_{loc}$

We can now define the stochastic integral of a process $X \in M^2_{loc}([0,T])$. The main idea is to approximate $X \in M^2_{loc}([0,T])$ by processes in $M^2([0,T])$.

Let, for every $n > 0$, $\tau_n = \inf\{t \le T; \int_0^t X_s^2\, ds > n\}$ with the understanding $\tau_n = T$ if $\int_0^T X_s^2\, ds \le n$. Then $\tau_n$ is a stopping time and the process $X_n(t) = X_t 1_{\{t<\tau_n\}}$ belongs to $M^2([0,T])$. Indeed, thanks to Theorem 7.5,

$$\int_0^T X_n(s)^2\, ds = \int_0^T X(s)^2 1_{\{s<\tau_n\}}\, ds = \int_0^{\tau_n \wedge T} X_s^2\, ds \le n$$

so that $X_n \in M^2$. We can therefore define, for every $n \ge 0$, the stochastic integral

$$I_n(t) = \int_0^t X_n(s)\, dB_s\, .$$

Let us observe that, as $\int_0^T X_s^2\, ds < +\infty$ by hypothesis, $\tau_n \nearrow T$ a.s. as $n \to +\infty$. Moreover, if $n > m$ then $\tau_n > \tau_m$, so that $\{\tau_n = T\} \supset \{\tau_m = T\}$ and the processes $X_n$ and $X_m$ coincide on $\{\tau_m = T\}$. By Proposition 7.2 $I_n$ and $I_m$ also coincide on $\{\tau_m = T\}$ a.s. for every $t$. As $\Omega = \bigcup_n \{\tau_n = T\}$, the sequence $(I_n)_n$ is eventually constant and therefore converges a.s. to some r.v. $I$.

**Definition 7.2** Let $X \in M^2_{loc}([0,T])$, then its stochastic integral is defined as

$$I_t = \int_0^t X_s\, dB_s = \lim_{n\to\infty} \int_0^t X_s 1_{\{s<\tau_n\}}\, dB_s \qquad \text{a.s.}$$

The stochastic integral of a process $X \in M^2_{loc}([0,T])$ is obviously continuous, as it coincides with $I_n$ on $\{\tau_n > T\}$ and $I_n$ is continuous as a stochastic integral of a process in $M^2$.

If $X \in M^2$ then also $X \in M^2_{loc}$. Let us verify that in this case Definition 7.2 coincides with the definition given for processes of $M^2$ in Sect. 7.3, p. 187. Note that if $X \in M^2([0,T])$, then

$$\mathrm{E}\Big(\int_0^T |X_s - X_n(s)|^2\, ds\Big) = \mathrm{E}\Big(\int_{\tau_n}^T |X_s|^2\, ds\Big) \tag{7.32}$$

since $X_s = X_n(s)$ if $s < \tau_n$, whereas $X_n(s) = 0$ on $s \ge \tau_n$. Therefore the right-hand side in (7.32) tends to 0 as $n \to \infty$, thanks to Lebesgue's theorem, so that

$$\lim_{n\to\infty} \mathrm{E}\Big(\int_0^T |X_s - X_n(s)|^2\, ds\Big) = 0\, .$$

Thanks to Corollary 7.1,

$$\lim_{n\to\infty} \mathrm{E}\Big(\sup_{0\le t\le T}\Big|\int_0^t X_s\,dB_s - \int_0^t X_n(s)\,dB_s\Big|^2\Big) = 0$$

and therefore if $X \in M^2([0,T])$ the two definitions coincide.

*Remark 7.4* The statement of Theorem 7.2 remains true for stochastic integrals of processes of $M^2_{loc}([0,T]$. Indeed if $X_t = Y_t$ on $\Omega_0$ for every $t \in [a,b]$, then this also true for the approximants $X_n$, $Y_n$. Therefore the stochastic integrals

$$\int_0^T X_n(s)\,dB_s \quad \text{and} \quad \int_0^T Y_n(s)\,dB_s$$

also coincide on $\Omega_0$ and therefore their a.s. limits also coincide.

We now point out some properties of the stochastic integral when the integrand is a process in $X \in M^2_{loc}([a,b])$. Let us first look for convergence results of processes $X_n, X \in M^2_{loc}([a,b])$. We shall see that if the processes $X_n$ suitably approximate $X$ in $M^2_{loc}([a,b])$, then the stochastic integrals converge in probability. The key tool in this direction is the following.

**Lemma 7.3** If $X \in M^2_{loc}([a,b])$, then for every $\varepsilon > 0$, $\rho > 0$

$$\mathrm{P}\Big(\Big|\int_a^b X_t\,dB_t\Big| > \varepsilon\Big) \le \mathrm{P}\Big(\int_a^b X_t^2\,dt > \rho\Big) + \frac{\rho}{\varepsilon^2}\,\cdot$$

*Proof* Let $\tau_\rho = \inf\{t; t \ge a, \int_a^t X_s^2\,ds \ge \rho\}$. Then we can write

$$\begin{aligned} &\mathrm{P}\Big(\Big|\int_a^b X_t\,dB_t\Big| > \varepsilon\Big) \\ &= \mathrm{P}\Big(\Big|\int_a^b X_t\,dB_t\Big| > \varepsilon, \tau_\rho > T\Big) + \mathrm{P}\Big(\Big|\int_a^b X_t\,dB_t\Big| > \varepsilon, \tau_\rho \le T\Big) \\ &\le \mathrm{P}\Big(\Big|\int_a^b X_t\,dB_t\Big| > \varepsilon, \tau_\rho > T\Big) + \mathrm{P}(\tau_\rho \le T)\,. \end{aligned} \tag{7.33}$$

As the two processes $t \mapsto X_t$ and $t \mapsto X_t 1_{\{\tau_\rho > t\}}$ coincide on $\Omega_0 = \{\tau_\rho > T\}$ and by Remark 7.4,

$$\mathrm{P}\Big(\Big|\int_a^b X_t\,dB_t\Big| > \varepsilon, \tau_\rho > T\Big) = \mathrm{P}\Big(\Big|\int_a^b X_t 1_{\{\tau_\rho > t\}}\,dB_t\Big| > \varepsilon, \tau_\rho > T\Big)$$
$$\le \mathrm{P}\Big(\Big|\int_a^b X_t 1_{\{\tau_\rho > t\}}\,dB_t\Big| > \varepsilon\Big)\,.$$

Therefore, as

$$\mathrm{E}\Big(\Big|\int_a^b X_t 1_{\{\tau_\rho > t\}}\,dB_t\Big|^2\Big) = \mathrm{E}\Big(\int_a^{b\wedge\tau_\rho} X_t^2\,dt\Big) \le \rho\,,$$

by Chebyshev's inequality

$$\mathrm{P}\Big(\Big|\int_a^b X_t\,dB_t\Big| > \varepsilon, \tau_\rho > T\Big) \le \frac{\rho}{\varepsilon^2}$$

and as

$$\mathrm{P}(\tau_\rho \le T) = \mathrm{P}\Big(\int_a^b X_t^2\,dt > \rho\Big)\,,$$

going back to (7.33) we can conclude the proof. □

**Proposition 7.3** Let $X, X_n \in M^2_{loc}([a,b])$, $n \ge 1$, and let us assume that

$$\int_a^b |X(t) - X_n(t)|^2\,dt \ \underset{n\to\infty}{\overset{\mathrm{P}}{\to}}\ 0\,.$$

Then

$$\int_a^b X_n(t)\,dB_t \ \underset{n\to\infty}{\overset{\mathrm{P}}{\to}}\ \int_a^b X(t)\,dB_t\,.$$

*Proof* By Lemma 7.3

$$\mathrm{P}\Big(\Big|\int_a^b X_n(t)\,dB_t - \int_a^b X_t\,dB_t\Big| > \varepsilon\Big) = \mathrm{P}\Big(\Big|\int_a^b (X_n(t) - X_t)\,dB_t\Big| > \varepsilon\Big)$$
$$\le \mathrm{P}\Big(\int_a^b |X_n(t) - X_t|^2\,dt > \rho\Big) + \frac{\rho}{\varepsilon^2}\,\cdot$$

Let $\eta > 0$ and let us first choose $\rho$ so that $\frac{\rho}{\varepsilon^2} \le \frac{\eta}{2}$ and then $n_0$ such that for $n > n_0$

$$\mathrm{P}\Big(\int_a^b |X_t - X_n(t)|^2\,dt > \rho\Big) < \frac{\eta}{2}\,\cdot$$

Therefore for $n > n_0$

$$\mathrm{P}\Big(\Big|\int_a^b X_n(t)\,dB_t - \int_a^b X_t\,dB_t\Big| > \varepsilon\Big) < \eta\,,$$

which allows us to conclude the proof. □

The following proposition states, similarly to the Lebesgue integral, that if the integrand is continuous then the integral is the limit of the Riemann sums. Note, however, that the limit is only in probability and that the integrand must be computed at the left end point of every small interval (see Exercise 7.9 to see what happens if this rule is not respected).

**Proposition 7.4** If $X \in M^2_{loc}([a,b])$ is a *continuous* process, then for every sequence $(\pi_n)_n$ of partitions $a = t_{n,0} < t_{n,1} < \cdots < t_{n,m_n} = b$ with $|\pi_n| = \max|t_{n_{k+1}} - t_{n_k}| \to 0$ we have

$$\sum_{k=0}^{m_n-1} X(t_{n,k})(B_{t_{n,k+1}} - B_{t_{n,k}}) \underset{n\to\infty}{\overset{\mathrm{P}}{\to}} \int_a^b X(t)\,dB_t\,.$$

*Proof* Let

$$X_n(t) = \sum_{k=0}^{m_n-1} X(t_{n,k}) 1_{[t_{n,k},t_{n,k+1}[}(t)\,.$$

$X_n$ is an elementary process and

$$\sum_{k=0}^{m_n-1} X(t_{n,k})(B_{t_{n,k+1}} - B_{t_{n,k}}) = \int_a^b X_n(t)\,dB_t\,.$$

As the paths are continuous,

$$\int_a^b |X_n(t) - X(t)|^2\,dt \underset{n\to\infty}{\to} 0 \ \text{ a.s.}$$

Hence by Proposition 7.3,

$$\sum_{k=0}^{m_n-1} X(t_{n,k})(B_{t_{n,k+1}} - B_{t_{n,k}}) = \int_a^b X_n(t)\, dB_t \ \underset{n\to\infty}{\overset{\mathrm{P}}{\to}} \ \int_a^b X(t)\, dB_t \ .$$

□

*Remark 7.5* It is worthwhile to point out that we defined the stochastic integral under the assumption that the integrand is progressively measurable, which implies, intuitively, that its value does not depend on the future increments of the Brownian motion.

This hypothesis is essential in the derivation of some of the most important properties of the stochastic integral, as already pointed out at the end of the proof of Lemma 7.1. Exercise 7.9 shows that strange things happen when one tries to mimic the construction of the stochastic integral without this assumption.

For the usual integral multiplying constants can be taken in and out of the integral sign. This is also true for the stochastic integral, but a certain condition is necessary, which requires attention.

**Theorem 7.6** Let $Z$ be a real $\mathscr{F}_a$-measurable r.v.; then for every $X \in M^2_{loc}([a,b])$

$$\int_a^b ZX_t\, dB_t = Z \int_a^b X_t\, dB_t \ . \tag{7.34}$$

*Proof* Let us consider the case of an elementary integrand

$$X_t = \sum_{i=1}^m X_i 1_{[t_i,t_{i+1}[}(t) \ .$$

Then $ZX$ is still an elementary process. Indeed $ZX_t = \sum_{i=1}^n ZX_i 1_{[t_i,t_{i+1}[}(t)$ and, as $Z$ is $\mathscr{F}_{t_i}$-measurable for every $i = 1, \dots, m$, the r.v.'s $ZX_i$ remain $\mathscr{F}_{t_i}$-measurable (here the hypothesis that $Z$ is $\mathscr{F}_a$-measurable and therefore $\mathscr{F}_{t_i}$-measurable for every $i$ is crucial). It is therefore immediate that the statement is true for elementary processes. Once the statement is proved for elementary processes it can be extended first to processes in $M^2$ and then in $M^2_{loc}$ by straightforward methods of approximation. The details are left to the reader. □

Note also that if the condition “$Z$ $\mathscr{F}_a$-measurable” is not satisfied, the relation (7.34) is not even meaningful: the process $t \mapsto ZX_t$ might not be adapted and the stochastic integral in this case has not been defined.

## 7.6 Local martingales

We have seen that, if $X \in M^2([0, +\infty[)$, $(I_t)_t$ is a martingale and this fact has been fundamental in order to derive a number of important properties of the stochastic integral as a process. This is not true in general if $X \in M^2_{loc}([0, +\infty[)$: $I_t$ in this case might not be integrable. Nevertheless, let us see in this case how $(I_t)_t$ can be approximated with martingales.

**Definition 7.3** A process $M = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (M_t)_t, \mathrm{P})$ is said to be a *local martingale* if there exists an increasing sequence $(\tau_n)_n$ of stopping times such that

i) $\tau_n \nearrow +\infty$ as $n \to \infty$ a.s.
ii) $(M_{t\wedge\tau_n})_t$ is a $(\mathscr{F}_t)_t$-martingale for every $n$.

We shall say that a sequence $(\tau_n)_n$ as above *reduces* the local martingale $M$.

It is clear that every martingale is a local martingale (just choose $\tau_n \equiv +\infty$).

The important motivation of this definition is that the stochastic integral of a process $X \in M^2_{loc}$ is a local martingale: going back to Definition 7.2, if $\tau_n = \inf\{t; \int_0^t X_s^2\,ds > n\}$ and

$$I_t = \int_0^t X_s\,dB_s$$

then

$$I_{t\wedge\tau_n} = \int_0^{t\wedge\tau_n} X_s\,dB_s = \int_0^t X_s 1_{\{s<\tau_n\}}\,dB_s$$

and the right-hand side is a (square integrable) martingale, being the stochastic integral of the process $s \mapsto X_s 1_{\{s<\tau_n\}}$ which belongs to $M^2$.

*Remark 7.6*

a) If $M$ is a local martingale, then $M_t$ might be non-integrable. However, $M_0$ is integrable. In fact $M_0 = M_{0\wedge\tau_n}$ and $(M_{t\wedge\tau_n})_t$ is a martingale.
b) Every local martingale has a right-continuous modification. Actually, this is true for all the stopped martingales $(M_{t\wedge\tau_n})_t$ by Theorem 5.14, so that $t \mapsto M_t$ has a right continuous modification for $t \le \tau_n$ for every $n$. Now just observe that $\tau_n \to +\infty$ as $n \to \infty$ a.s.
c) In Definition 7.3 we can always assume that $(M_{t\wedge\tau_n})_t$ is a uniformly integrable martingale for every $n$. If $(\tau_n)_n$ reduces $M$ the same is true for $\sigma_n = \tau_n \wedge n$. Indeed, condition i) is immediate. Also the fact that $(M_{t\wedge\sigma_n})_t$ is a martingale is immediate, being the martingale $(M_{\tau_n\wedge t})_t$ stopped at time $n$. Therefore $(\sigma_n)_n$ also reduces $M$. Moreover, $M_{\sigma_n}$ is integrable because $M_{\sigma_n} = M_{n\wedge\tau_n}$ and $(M_{t\wedge\tau_n})_t$ is a martingale. We also have

$$M_{t\wedge\sigma_n} = \mathrm{E}[M_{\sigma_n} | \mathscr{F}_t] \ . \tag{7.35}$$

In fact for $t \le n$ we have $\mathrm{E}[M_{\sigma_n}|\mathscr{F}_t] = \mathrm{E}[M_{n\wedge\tau_n}|\mathscr{F}_t] = M_{t\wedge\sigma_n}$, thanks to the martingale property of $(M_{t\wedge\tau_n})_t$, whereas, if $t > n$, $M_{\sigma_n} = M_{n\wedge\tau_n}$ is already $\mathscr{F}_t$-measurable. (7.35) implies that $(M_{t\wedge\sigma_n})_t$ is uniformly integrable.
d) If, moreover, $M$ is continuous, then we can assume that the stopped martingale $(M_{t\wedge\tau_n})_t$ is bounded. Let $\rho_n = \inf\{t; |M_t| \ge n\}$ and $\sigma_n = \tau_n \wedge \rho_n$, then the sequence $(\sigma_n)_n$ again satisfies the conditions of Definition 7.3 and, moreover, $|M_{\sigma_n\wedge t}| \le n$ for every $t$.

We shall always assume that the stopped martingales $(M_{t\wedge\tau_n})_t$ are uniformly integrable and, if $(M_t)_t$ is continuous, that they are also bounded, which is always possible thanks to the previous remark.

As remarked above, in general a local martingale need not be integrable. However, it is certainly integrable if it is positive, which is one of the consequences of the following result.

**Proposition 7.5** A positive local martingale is a supermartingale.

*Proof* Let $(\tau_n)_n$ be a sequence of stopping times that reduces $M$. Then $M_{t\wedge\tau_n} \to M_t$ as $n \to +\infty$ and by Fatou's lemma

$$\mathrm{E}(M_t) \le \lim_{n\to\infty} \mathrm{E}(M_{t\wedge\tau_n}) = \mathrm{E}(M_0) \ ,$$

therefore $M_t$ is integrable. We must now prove that if $A \in \mathscr{F}_s$ and $s < t$, then $\mathrm{E}(M_t 1_A) \le \mathrm{E}(M_s 1_A)$, i.e. that

$$\mathrm{E}[(M_t - M_s)1_A] \le 0\,. \tag{7.36}$$

We have

$$\mathrm{E}[1_{A\cap\{\tau_n>s\}}(M_{t\wedge\tau_n} - M_s)] = \mathrm{E}[1_{A\cap\{\tau_n>s\}}(M_{t\wedge\tau_n} - M_{s\wedge\tau_n})] = 0\,. \tag{7.37}$$

Indeed $A \cap \{\tau_n > s\} \in \mathscr{F}_s$ and $(M_{t\wedge\tau_n})_t$ is a martingale, which justifies the last equality. Also we used the fact that on $\{\tau_n > s\}$ the r.v.'s $M_s$ and $M_{s\wedge\tau_n}$ coincide. As

$$\lim_{n\to\infty} 1_{A\cap\{\tau_n>s\}}(M_{t\wedge\tau_n} - M_s) = 1_A(M_t - M_s)$$

and

$$1_{A\cap\{\tau_n>s\}}(M_{t\wedge\tau_n} - M_s) \ge -M_s\,,$$

we can apply Fatou's lemma (we proved above that $-M_s$ is integrable), which gives

$$0 = \varliminf_{n\to\infty} \mathrm{E}[1_{A\cap\{\tau_n>s\}}(M_{t\wedge\tau_n} - M_s)] \ge \mathrm{E}[1_A(M_t - M_s)]\,,$$

completing the proof. □

*Remark 7.7* A bounded local martingale $M$ is a martingale. This is almost obvious as, if the sequence $(\tau_n)_n$ reduces $M$, then

$$\mathrm{E}(M_{t\wedge\tau_n}|\mathscr{F}_s) = M_{s\wedge\tau_n}$$

and we can take the limit as $n \to \infty$ using Lebesgue's theorem for conditional expectations (Proposition 4.2).

Note, however, that other apparently strong assumptions are not sufficient to guarantee that a local martingale is a martingale. For instance, there exist uniformly integrable local martingales that are not martingales (see Example 8.10). A condition for a local martingale to be a true martingale is provided in Exercise 7.15.

**Proposition 7.6** Let $M$ be a continuous local martingale. Then there exists a unique continuous increasing process $(A_t)_t$, which we continue to call the *associated increasing process to* $M$, such that $X_t = M_t^2 - A_t$ is a continuous local martingale.

The proof is not difficult and consists in the use of the definition of a local martingale in order to approximate $M$ with square integrable martingales, for which Theorem 5.16 holds. Instead of giving the proof, let us see what happens in the situation that is of greatest interest to us.

**Proposition 7.7** If $X \in M^2_{loc}([0,T])$ and $I_t = \int_0^t X_s\,dB_s$, then $(I_t)_{0\le t\le T}$ is a local martingale whose increasing process is

$$A_t = \int_0^t X_s^2\,ds\,.$$

*Proof* We know already that $I$ is a local martingale. In order to complete the proof we must prove that $(I_t^2 - A_t)_t$ is a local martingale. Let $\tau_n = \inf\{t; \int_0^t X_s^2\,ds > n\}$, then we know already that $(\tau_n)_n$ reduces $I$ and that

$$I_{t\wedge\tau_n} = \int_0^{t\wedge\tau_n} X_s\,dB_s = \int_0^t 1_{\{s<\tau_n\}} X_s\,dB_s$$

is a square integrable $(\mathscr{F}_t)_t$-martingale. By Proposition 7.3 and Lemma 7.5,

$$I^2_{t\wedge\tau_n} - A_{t\wedge\tau_n} = \Big(\int_0^t 1_{\{s<\tau_n\}} X_s\,dB_s\Big)^2 - \int_0^t 1_{\{s<\tau_n\}} X_s^2\,ds$$

is a martingale, which means that $(\tau_n)_n$ reduces $(I_t^2 - A_t)_t$, which is therefore a local martingale. □

We shall still denote by $(\langle M\rangle_t)_t$ the associated increasing process of the local martingale $M$.

**Corollary 7.2** If $M$ and $N$ are continuous local martingales, then there exists a unique process $A$ with finite variation such that $Z_t = M_tN_t - A_t$ is a continuous local martingale.

The proof of the corollary boils down to the observation that $M_tN_t = \frac{1}{4}((M_t+N_t)^2 - (M_t - N_t)^2)$, so that $A_t = \frac{1}{4}(\langle M+N\rangle_t - \langle M-N\rangle_t)$ satisfies the requirement. $(A_t)_t$ is a process with finite variation, being the difference of two increasing processes.

We will denote by $\langle M,N\rangle$ the process with finite variation of Corollary 7.2.

## Exercises

**7.1** (p. 523) Let $B$ be a Brownian motion.

a) Compute

$$Z=\int_0^1 1_{\{B_t=0\}}\,dB_t\ .$$

b) Let

$$Z=\int_0^1 1_{\{B_t\ge 0\}}\,dB_t\ .$$

Compute $\mathrm{E}[Z]$ and $\mathrm{Var}(Z)$.

- Note that the processes $t\mapsto 1_{\{B_t=0\}}$ and $t\mapsto 1_{\{B_t\ge 0\}}$ are progressively measurable, thanks to Exercise 2.3.

**7.2** (p. 523) Compute

$$\mathrm{E}\Big(B_s\int_0^t B_u\,dB_u\Big)\ .$$

**7.3** (p. 523) Let $B$ be a standard Brownian motion and $Y_t=\int_0^t \mathrm{e}^s\,dB_s$. If

$$Z_t=\int_0^t Y_s\,dB_s\ ,$$

compute $\mathrm{E}(Z_t)$, $\mathrm{E}(Z_t^2)$ and $\mathrm{E}(Z_tZ_s)$.

**7.4** (p. 524)

a) Compute, for $s\le t$,

$$\mathrm{E}\Big[B_s^2\Big(\int_s^t B_u\,dB_u\Big)^2\Big]\ .$$

b) Prove that, if $X\in M^2([s,t])$, the r.v. $Z=\int_s^t X_u\,dB_u$ is uncorrelated with $B_v$, $v\le s$, but, in general, it is *not* independent of $\mathscr{F}_s$.

**7.5** (p. 525) Let

$$X_t=\int_0^t \mathrm{e}^{-B_s^2}\,dB_s\ .$$

Prove that $X$ is a square integrable martingale. Is it bounded in $L^2$?

**7.6** (p. 525) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a two-dimensional standard Brownian motion and let

$$Z_t = \int_0^t \frac{1}{1+4s}\, dB_1(s) \int_0^s \mathrm{e}^{-B_1(u)^2}\, dB_2(u)\ .$$

a) Is $Z$ a martingale? Determine the processes $(\langle Z\rangle_t)_t$ and $(\langle Z, B_1\rangle_t)_t$.
b) Prove that the limit $Z_\infty = \lim_{t\to+\infty} Z_t$ exists a.s. and in $L^2$ and compute $\mathrm{E}[Z_\infty]$ and $\mathrm{Var}(Z_\infty)$.

**7.7** (p. 526)

a) Let $f \in L^2([s,t])$; show that $W = \int_s^t f(u)\, dB_u$ is independent of $\mathscr{F}_s$.
b) Let $\Phi : \mathbb{R}^+ \to \mathbb{R}^+$ be a strictly increasing differentiable function such that $\Phi(0) = 0$ and let

$$B_t^{\Phi} = \int_0^{\Phi^{-1}(t)} \sqrt{\Phi'(u)}\, dB_u\ .$$

Then $B^{\Phi}$ is a Brownian motion. With respect to which filtration?

**7.8** (p. 527) Show that the r.v. $\int_0^t B_u^2\, du$ is orthogonal in $L^2(\Omega, \mathscr{F}, \mathrm{P})$ to every r.v. of the form $\int_0^s f(u)\, dB_u$, for every $s > 0$ and $f \in L^2([0,s])$.

**7.9** (p. 528) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t)$ be a Brownian motion and let $0 = t_0 < t_1 < \cdots < t_n = t$ be a partition of the interval $[0,t]$ and $|\pi| = \max_{i=0,\dots,n-1} |t_{i+1} - t_i|$ the amplitude of the partition. Mr. Whynot decides to approximate the stochastic integral

$$\int_0^t X_s\, dB_s$$

for a continuous adapted process $X$ with the Riemann sums

$$\sum_{i=0}^{n-1} X_{t_{i+1}}(B_{t_{i+1}} - B_{t_i}) \tag{7.38}$$

instead of the sums

$$\sum_{i=0}^{n-1} X_{t_i}(B_{t_{i+1}} - B_{t_i}) \tag{7.39}$$

as indicated (and recommended!) in Proposition 7.4. He argues that, if this were an ordinary integral instead of a stochastic integral, then the result would be the same.

a) Compute the limit of the sums (7.38) as the amplitude of the partition tends to 0 when $X = B$.
b) Do the same thing assuming that the process $X$ has paths with finite variation.

**7.10** (p. 529) Let $B$ be a Brownian motion and $f \in L^2(\mathbb{R}^+)$.

a) Show that the limit

$$\lim_{t\to+\infty} \int_0^t f(s)\, dB_s$$

exists in $L^2$. Does it also exist a.s.?

b) Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a function such that $\int_0^T g(s)^2\, ds < +\infty$ for every $T > 0$.

b1) Show that

$$Y_t = \mathrm{e}^{\int_0^t g(s)\, dB_s - \frac{1}{2}\int_0^t g(s)^2\, ds}$$

is a martingale. Compute $\mathrm{E}[Y_t^2]$.

b2) Prove that if $g \in L^2(\mathbb{R}^+)$ then $(Y_t)_t$ is uniformly integrable.

b3) Assume, conversely, that $\int_0^{+\infty} g(s)^2\, ds = +\infty$. Compute $E[Y_t^\alpha]$ for $\alpha < 1$. What is the value of $\lim_{t\to+\infty} Y_t$ now? Is $(Y_t)_t$ uniformly integrable in this case?

**7.11** (p. 530) a) For $t < 1$ let $B$ a Brownian motion and

$$Y_t = (1-t)\int_0^t \frac{dB_s}{1-s}\,\cdot$$

a1) Show that $(Y_t)_t$ is a Gaussian process and compute $\mathrm{E}(Y_t Y_s)$, $0 \le s < t < 1$. Does this remind you of a process we have already met?

a2) Show that the limit $\lim_{t\to1-} Y_t$ exists in $L^2$ and compute it.

b) Let $A(s) = \frac{s}{1+s}$ and let

$$W_s = \int_0^{A(s)} \frac{dB_u}{1-u}\,\cdot$$

Show that $(W_s)_s$ is a Brownian motion and deduce that $\lim_{t\to1-} Y_t = 0$ a.s.

**7.12** (p. 531) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and $\lambda > 0$. Let

$$Y_t = \int_0^t \mathrm{e}^{-\lambda(t-s)}\, dB_s, \qquad Z_t = \int_0^t \mathrm{e}^{-\lambda s}\, dB_s\,.$$

a) Prove that, for every $t > 0$, $Y_t$ and $Z_t$ have the same law and compute it.
b) Show that $(Z_t)_t$ is a martingale. And $(Y_t)_t$?
c) Show that $\lim_{t\to+\infty} Z_t$ exists a.s. and in $L^2$.
d1) Show that $\lim_{t\to+\infty} Y_t$ exists in law and determine the limit law.
d2) Show that

$$\lim_{t\to+\infty} \mathrm{E}[(Y_{t+h} - Y_t)^2] = \frac{1}{\lambda}\,(1 - \mathrm{e}^{-\lambda h})$$

and therefore $(Y_t)_t$ cannot converge in $L^2$.

**7.13** (p. 532) Let $B$ be a Brownian motion and let $\widetilde{B}_0 = 0$ and

$$\widetilde{B}_t = \int_0^t \Big(3 - \frac{12u}{t} + \frac{10u^2}{t^2}\Big)\, dB_u\,.$$

a) Show that $\widetilde{B}$ is a natural Brownian motion.
b) Let $Y = \int_0^1 u\, dB_u$. Show that $Y$ is independent of $\widetilde{B}_t$ for every $t \ge 0$.
c) Show that the $\sigma$-algebra generated by $\widetilde{B}_t$, $t \ge 0$ is strictly smaller than the $\sigma$-algebra generated by $B_t$, $t \ge 0$.

**7.14** (p. 533) Let $X \in M^2([0,T])$. We know that $\int_0^t X_s\, dB_s$ is square integrable. Is the converse true? That is if $X \in M^2_{loc}([0,T])$ and its stochastic integral is square integrable, does this imply that $X \in M^2([0,T])$ and that its stochastic integral is a square integrable martingale? The answer is no (a counterexample is given in Exercise 8.21). This exercise goes deeper into this question.

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and let $X \in M^2_{loc}([0,T])$ and $M_t = \int_0^t X_s\, dB_s$.

a) Let $\tau_n = \inf\{t; \int_0^t X_s^2\, ds > n\}$ with the understanding that $\tau_n = T$ if $\int_0^T X_s^2\, ds \le n$. Prove that

$$\mathrm{E}\Big[\int_0^{\tau_n} X_s^2\, ds\Big] \le \mathrm{E}\Big[\sup_{0\le t\le T} M_t^2\Big]\,.$$

b) Prove that $X \in M^2([0,T])$ *if and only if*

$$\mathrm{E}\Big[\sup_{0\le t\le T} M_t^2\Big] < +\infty\,. \tag{7.40}$$

**7.15** (p. 533) Let $M = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (M_t)_t, \mathrm{P})$ be a local martingale and assume that for every $t > 0$ the family $(M_{t\wedge\tau})_\tau$ is uniformly integrable, with $\tau$ ranging among all stopping times of $(\mathscr{F}_t)_t$. Prove that $(M_t)_t$ is a martingale.

# Chapter 8
# Stochastic Calculus

## 8.1 Ito's formula

Let $X$ be a process such that, for every $0 \le t_1 < t_2 \le T$,

$$X_{t_2} - X_{t_1} = \int_{t_1}^{t_2} F_t\,dt + \int_{t_1}^{t_2} G_t\,dB_t\ ,$$

where $F \in M^1_{loc}([0,T])$ and $G \in M^2_{loc}([0,T])$. We say then that $X$ admits the *stochastic differential*

$$dX_t = F_t\,dt + G_t\,dB_t\ .$$

A process admitting a stochastic differential is called an *Ito process*. An Ito process is therefore the sum of a process with finite variation and of a local martingale.

*Remark 8.1* The stochastic differential is unique: if there existed $A_1, A_2 \in M^1_{loc}([0,T])$, $G_1, G_2 \in M^2_{loc}([0,T])$ such that

$$A_1(t)\,dt + G_1(t)\,dB_t = A_2(t)\,dt + G_2(t)\,dB_t\ ,$$

then, for every $0 \le t \le T$, we would have a.s.

$$A_t \stackrel{def}{=} \int_0^t (A_1(s) - A_2(s))\,ds = \int_0^t (G_2(s) - G_1(s))\,dB_s \stackrel{def}{=} G_t$$

and this is not possible, unless the two integrands are identically zero, as the left-hand side is a process with finite variation, whereas the right-hand side, which is a local martingale, is not.

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_8

We shall write

$$\langle X\rangle_t = \int_0^t G_s^2\, ds\ .$$

$\langle X\rangle$ is nothing else than the increasing process (see Proposition 7.6) associated to the local martingale appearing in the definition of $X$, and it is well defined thanks to the previous remark. Similarly, if $Y$ is another Ito process with stochastic differential

$$dY_t = H_t\, dt + K_t\, dB_t,$$

we shall set

$$\langle X, Y\rangle_t = \int_0^t G_s K_s\, ds\ .$$

*Example 8.1* Let us compute the stochastic differential of $X_t = B_t^2$. The analogy with standard calculus might suggest that $dX_t = 2B_t\, dB_t$, but we have already remarked that this cannot be as it would give $B_t^2 = 2\int_0^t B_s\, dB_s$, which is impossible because the stochastic integral is a centered r.v. whereas $B_t^2$ is not.

Let $0 \le t_1 < t_2 \le T$ and $\pi_n = \{t_1 = t_{n,1} < t_{n,2} < \dots < t_{n,m_n} = t_2\}$ be a partition of $[t_1, t_2]$ such that $|\pi_n| \to 0$, then by Proposition 7.4, meaning the limits in probability,

$$\begin{aligned}\int_{t_1}^{t_2} B_t\, dB_t &= \lim_{n\to\infty} \sum_{k=1}^{m_n-1} B_{t_{n,k}}[B_{t_{n,k+1}} - B_{t_{n,k}}]\\ &= \frac{1}{2} \lim_{n\to\infty} \sum_{k=1}^{m_n-1} \Big([B^2_{t_{n,k+1}} - B^2_{t_{n,k}}] - [B_{t_{n,k+1}} - B_{t_{n,k}}]^2\Big)\\ &= \frac{1}{2}[B^2_{t_2} - B^2_{t_1}] - \frac{1}{2} \lim_{n\to\infty} \sum_{k=1}^{m_n-1} [B_{t_{n,k+1}} - B_{t_{n,k}}]^2\ .\end{aligned}$$

By Proposition 3.4 the rightmost limit is equal to $t_2 - t_1$ in $L^2$; therefore

$$\int_{t_1}^{t_2} B_t\, dB_t = \frac{1}{2}(B^2_{t_2} - B^2_{t_1}) - \frac{1}{2}(t_2 - t_1)\ ,$$

i.e.

$$dB_t^2 = dt + 2B_t\, dB_t\ . \tag{8.1}$$

*Example 8.2* $X_t = tB_t$. With the notations of the previous example

$$\int_{t_1}^{t_2} t\,dB_t = \lim_{n\to\infty} \sum_{k=1}^{m_n-1} t_{n,k}(B_{t_{n,k+1}} - B_{t_{n,k}})$$

$$\int_{t_1}^{t_2} B_t\,dt = \lim_{n\to\infty} \sum_{k=1}^{m_n-1} B_{t_{n,k}}(t_{n,k+1} - t_{n,k})\ .$$

Let us compute the limit, in probability,

$$\begin{aligned}
&\lim_{n\to\infty} \Big( \sum_{k=1}^{m_n-1} B_{t_{n,k}}(t_{n,k+1} - t_{n,k}) + \sum_{k=1}^{m_n-1} t_{n,k}(B_{t_{n,k+1}} - B_{t_{n,k}}) \Big) \\
&= \lim_{n\to\infty} \Big\{ \sum_{k=1}^{m_n-1} \Big( B_{t_{n,k}}(t_{n,k+1} - t_{n,k}) + t_{n,k+1}(B_{t_{n,k+1}} - B_{t_{n,k}}) \Big) \\
&\qquad + \sum_{k=1}^{m_n-1} (t_{n,k} - t_{n,k+1})(B_{t_{n,k+1}} - B_{t_{n,k}}) \Big\}\ .
\end{aligned} \tag{8.2}$$

But as

$$\begin{aligned}
&\Big| \sum_{k=1}^{m_n-1} (t_{n,k+1} - t_{n,k})(B_{t_{n,k+1}} - B_{t_{n,k}}) \Big| \\
&\le \max_{k=1,\dots,m_n-1} |B_{t_{n,k+1}} - B_{t_{n,k}}| \underbrace{\sum_{k=1}^{m_n-1} |t_{n,k+1} - t_{n,k}|}_{=t_2-t_1}\ ,
\end{aligned}$$

thanks to the continuity of the paths of $B$, the quantity above converges to 0 a.s. Therefore the limit in (8.2) is equal to

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{k=1}^{m_n-1} \Big( B_{t_{n,k}}(t_{n,k+1} - t_{n,k}) + t_{n,k+1}(B_{t_{n,k+1}} - B_{t_{n,k}}) \Big) \\
&\qquad = \lim_{n\to\infty} \sum_{k=1}^{m_n-1} [t_{n,k+1}B_{t_{n,k+1}} - t_{n,k}B_{t_{n,k}}]\ .
\end{aligned}$$

(continued)

*Example 8.2* (continued)
Hence

$$\int_{t_1}^{t_2} B_t\,dt + \int_{t_1}^{t_2} t\,dB_t = \lim_{n\to\infty} \sum_{k=1}^{m_n-1} [t_{n,k+1}B_{t_{n,k+1}} - t_{n,k}B_{t_{n,k}}] = t_2B_{t_2} - t_1B_{t_1} ,$$

i.e.

$$d(tB_t) = B_t\,dt + t\,dB_t . \tag{8.3}$$

**Proposition 8.1** If $X_i$, $i = 1, 2$, are processes with stochastic differentials

$$dX_i(t) = F_i(t)\,dt + G_i(t)\,dB_t$$

then

$$\begin{aligned} d(X_1(t)X_2(t)) &= X_1(t)\,dX_2(t) + X_2(t)\,dX_1(t) + G_1(t)G_2(t)\,dt \\ &= X_1(t)\,dX_2(t) + X_2(t)\,dX_1(t) + d\langle X_1, X_2\rangle_t . \end{aligned} \tag{8.4}$$

*Proof* If $F_1$, $G_1$, $F_2$, $G_2$ are constant on $[t_1, t_2]$, the statement is a consequence of (8.1) and (8.3). If, conversely, they are elementary processes, let $I_1, \dots, I_r$ be the subintervals of $[t_1, t_2]$ on which $F_1$, $F_2$, $G_1$, $G_2$ are constants; the statement follows now using (8.1) and (8.3) applied to each of the intervals $I_k$ and taking the sum. In general, let $F_{i,n}$, $G_{i,n}$, $i = 1, 2$, be elementary processes in $M^1_{loc}([0, T])$ and $M^2_{loc}([0, T])$, respectively, and such that

$$\int_0^T |F_{i,n}(t) - F_i(t)|\,dt \underset{n\to\infty}{\to} 0 \quad \text{a.s.}$$

$$\int_0^T |G_{i,n}(t) - G_i(t)|^2\,dt \underset{n\to\infty}{\to} 0 \quad \text{a.s.}$$

and let

$$X_{i,n}(t) = X_i(0) + \int_0^t F_{i,n}(s)\,ds + \int_0^t G_{i,n}(s)\,dB_s .$$

By Theorem 7.4

$$\sup_{0\le t\le T} |X_{i,n}(t) - X_i(t)| \underset{n\to\infty}{\overset{\mathrm{P}}{\to}} 0$$

so that, possibly considering a subsequence, $X_{i,n}(t) \to X_i(t)$ as $n \to \infty$ uniformly a.s. Therefore

$$\int_{t_1}^{t_2} X_{1,n}(t)F_{2,n}(t)\,dt \underset{n\to\infty}{\to} \int_{t_1}^{t_2} X_1(t)F_2(t)\,dt \qquad \text{a.s.}$$
$$\int_{t_1}^{t_2} X_{2,n}(t)F_{1,n}(t)\,dt \underset{n\to\infty}{\to} \int_{t_1}^{t_2} X_2(t)F_1(t)\,dt \qquad \text{a.s.}$$
$$\int_{t_1}^{t_2} G_{1,n}(t)G_{2,n}(t)\,dt \underset{n\to\infty}{\to} \int_{t_1}^{t_2} G_1(t)G_2(t)\,dt \qquad \text{a.s.}$$

and by Theorem 7.4

$$\int_{t_1}^{t_2} X_{1,n}(t)G_{2,n}(t)\,dB_t \underset{n\to\infty}{\overset{\mathrm{P}}{\to}} \int_{t_1}^{t_2} X_1(t)G_2(t)\,dB_t$$
$$\int_{t_1}^{t_2} X_{2,n}(t)G_{1,n}(t)\,dB_t \underset{n\to\infty}{\overset{\mathrm{P}}{\to}} \int_{t_1}^{t_2} X_2(t)G_1(t)\,dB_t$$

and taking the limit as $n \to \infty$ we obtain (8.4). □

Note that the formula for the differential of the product of two processes is a bit different from the corresponding one for the differential of the product of two functions. Actually there is an extra term, namely

$$d\langle X_1, X_2\rangle_t \,.$$

Note also that if at least one of $G_1$ and $G_2$ vanishes, then this additional term also vanishes and we have the usual relation

$$dX_1(t)X_2(t) = X_1(t)\,dX_2(t) + X_2(t)\,dX_1(t)\,.$$

Let $f : \mathbb{R} \times \mathbb{R}^+$ be a regular function of $(x, t)$ and $X$ an Ito process. What is the stochastic differential of the process $t \mapsto f(X_t, t)$? The answer to this question is Ito's formula, which will be a key tool of computation from now on. In the next theorem note again that the differential $df(X_t, t)$ behaves as an ordinary differential plus an extra term.

**Theorem 8.1 (Ito's formula)** Let $X$ be a process with stochastic differential

$$dX_t = F(t)\,dt + G(t)\,dB_t$$

(continued)

**Theorem 8.1** (continued)
and let $f : \mathbb{R}_x \times \mathbb{R}_t^+ \to \mathbb{R}$ be a continuous function in $(x, t)$, once continuously differentiable in $t$ and twice in $x$. Then

$$\begin{aligned} &df(X_t, t) \\ &= \frac{\partial f}{\partial t}(X_t, t)\, dt + \frac{\partial f}{\partial x}(X_t, t)\, dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(X_t, t)G(t)^2\, dt \\ &= \Big(\frac{\partial f}{\partial t}(X_t, t) + \frac{\partial f}{\partial x}(X_t, t)F(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(X_t, t)G(t)^2\Big)\, dt \\ &\qquad + \frac{\partial f}{\partial x}(X_t, t)G(t)\, dB_t\ . \end{aligned} \tag{8.5}$$

*Proof (The main ideas)*

1° *Step*: let us assume $f(x, t) = x^n$; (8.5) becomes

$$dX_t^n = nX_t^{n-1}\, dX_t + \tfrac{1}{2}\, n(n-1)X_t^{n-2}G_t^2\, dt\ .$$

This relation is obvious for $n = 1$ and follows easily by induction by Proposition 8.1.

2° *Step*: (8.5) is therefore true if $f(x, t)$ is a polynomial $P(x)$. Let us assume now that $f$ is of the form $f(x, t) = P(x)g_t$. We have

$$\begin{aligned} dP(X_t) &= [P'(X_t)F_t + \tfrac{1}{2}\, P''(X_t)G_t^2]\, dt + P'(X_t)G_t\, dB_t \\ dg_t &= g_t'\, dt \end{aligned}$$

and again (8.5) follows easily for such $f$, thanks to Proposition 8.1.

3° *Step*: (8.5) is therefore true if $f$ is of the form

$$f(x, t) = \sum_{i=1}^{\ell} P_i(x)g_i(t) \tag{8.6}$$

where the $P_i$'s are polynomials and the $g_i$'s are differentiable functions. If $f$ is continuous in $(x, t)$, once continuously differentiable in $t$ and twice in $x$, one can prove that there exists a sequence $(f_n)_n$ of functions of this form such that

$$\begin{gathered} f_n(x, t) \to f(x, t), \\ \frac{\partial}{\partial x} f_n(x, t) \to \frac{\partial}{\partial x} f(x, t), \qquad \frac{\partial}{\partial t} f_n(x, t) \to \frac{\partial}{\partial t} f(x, t), \\ \frac{\partial^2}{\partial x^2} f_n(x, t) \to \frac{\partial^2}{\partial x^2} f(x, t) \end{gathered}$$

as $n \to \infty$ uniformly on compact sets of $\mathbb{R} \times \mathbb{R}^+$ (see Friedman 1975, p. 95 for a proof). We have, for $t_1 < t_2 \le T$,

$$\begin{aligned} &f_n(X_{t_2}, t_2) - f_n(X_{t_1}, t_1) \\ &= \int_{t_1}^{t_2} \Big[\frac{\partial f_n}{\partial t}(X_t, t) + \frac{\partial f_n}{\partial x}(X_t, t)F_t + \frac{1}{2}\,\frac{\partial^2 f_n}{\partial x^2}(X_t, t)G_t^2\Big]\,dt \\ &\qquad + \int_{t_1}^{t_2} \frac{\partial f_n}{\partial x}(X_t, t)G_t\,dB_t\,. \end{aligned} \tag{8.7}$$

As $n \to \infty$ the first integral on the right-hand side converges toward

$$\int_{t_1}^{t_2} \Big[\frac{\partial f}{\partial t}(X_t, t) + \frac{\partial f}{\partial x}(X_t, t)F_t + \frac{1}{2}\,\frac{\partial^2 f}{\partial x^2}(X_t, t)G_t^2\Big]\,dt\,.$$

Actually, $t \mapsto X_t$ is continuous and therefore bounded on $[t_1, t_2]$ for every $\omega$; therefore $\frac{\partial}{\partial t} f_n(X_t, t)$ converges uniformly to $f_t(X_t, t)$ and the same holds for the other derivatives. Moreover,

$$\int_{t_1}^{t_1} \Big|\frac{\partial f_n}{\partial x}(X_t, t) - \frac{\partial f}{\partial x}(X_t, t)\Big|^2 G_t^2\,dt \ \underset{n\to\infty}{\overset{\mathrm{P}}{\to}}\ 0$$

and by Theorem 7.4

$$\int_{t_1}^{t_2} \frac{\partial f_n}{\partial x}(X_t, t)G_t\,dB_t \ \underset{n\to\infty}{\overset{\mathrm{P}}{\to}}\ \int_{t_1}^{t_2} \frac{\partial f}{\partial x}(X_t, t)G_t\,dB_t\,.$$

We can therefore take the limit as $n \to \infty$ in (8.7) and the statement is proved. □

*Example 8.3* As a particular case, if $f : \mathbb{R} \to \mathbb{R}$ is twice differentiable and

$$dX_t = F_t\,dt + G_t\,dB_t$$

then

$$df(X_t) = f'(X_t)\,dX_t + \frac{1}{2}f''(X_t)G_t^2\,dt = f'(X_t)\,dX_t + \frac{1}{2}f''(X_t)d\langle X\rangle_t\,.$$

In particular,

$$df(B_t) = f'(B_t)\,dB_t + \frac{1}{2}f''(B_t)\,dt\,.$$

*Example 8.4* If $f : \mathbb{R}^+ \to \mathbb{R}$ is differentiable then

$$\int_0^T f(s)\, dB_s = f(T)B_T - \int_0^T B_s f'(s)\, ds\, .$$

Actually Ito's formula (or more simply (8.4)) gives

$$df(t)B_t = f'(t)B_t\, dt + f(t)\, dB_t\, .$$

*Example 8.5* Let us compute the stochastic differential of

$$Z_t = (1-t)\int_0^t \frac{dB_s}{1-s}\, \cdot$$

In Exercise 7.11 it is shown that $Y$ is a Brownian bridge. There are two possibilities in order to apply Ito's formula. The first one is to write $Z_t = X_t Y_t$ where

$$X_t = 1-t, \qquad Y_t = \int_0^t \frac{dB_s}{1-s}\, ,$$

i.e.

$$dX_t = -dt, \qquad dY_t = \frac{1}{1-t}\, dB_t\, .$$

Proposition 8.1 gives therefore

$$dZ_t = Y_t\, dX_t + X_t\, dY_t = -Y_t\, dt + dB_t\, . \tag{8.8}$$

Actually here $\langle X, Y\rangle_t \equiv 0$. Observing that $Z_t = (1-t)Y_t$, the previous relation becomes

$$dZ_t = -\frac{1}{1-t} Z_t\, dt + dB_t\, ,$$

which is our first example of a Stochastic Differential Equation.

A second possibility in order to compute the stochastic differential is to write $Z_t = g(Y_t, t)$, where $g(x,t) = (1-t)x$. Now

$$\frac{\partial g}{\partial t}(x,t) = -x, \quad \frac{\partial g}{\partial x}(x,t) = 1-t, \quad \frac{\partial^2 g}{\partial x^2}(x,t) = 0$$

(continued)

*Example 8.5* (continued)
and the computation gives again

$$dZ_t = \frac{\partial g}{\partial t}(Y_t, t)\, dt + \frac{\partial g}{\partial x}(Y_t, t)\, dY_t = -Y_t\, dt + dB_t = -\frac{1}{1-t} Z_t\, dt + dB_t\,.$$

*Example 8.6* In Exercise 5.13 we proved that

$$Y_t = tB_t - \int_0^t B_u\, du$$

is a martingale. Let us see how Ito's formula makes all these computations simpler. We have

$$dY_t = t\, dB_t + B_t\, dt - B_t\, dt = t\, dB_t$$

i.e., as $Y_0 = 0$,

$$Y_t = \int_0^t s\, dB_s$$

so that $Y$, being the stochastic integral of a (deterministic) process of $M^2$, is a square integrable martingale. This representation also makes it particularly easy to compute the variance of $Y$, which was done in Exercise 5.13 with some effort: we have

$$\mathrm{E}(Y_t^2) = \int_0^t s^2\, ds = \frac{t^3}{3}\,\cdot$$

Example 8.6 points out how Ito's formula allows to check that a process is a martingale: first compute its stochastic differential. If the finite variation term vanishes (i.e. there is no term in $dt$) then the process is a stochastic integral and necessarily a *local* martingale. If, moreover, it is the stochastic integral of a process in $M^2$, then it is a martingale (we shall see in the examples other ways of proving that a stochastic integral is actually a martingale). This method of checking that a process is a martingale also provides immediately the associated increasing process.

*Remark 8.2* It is possible to give a heuristic explanation of the extra term appearing in Ito's formula. Thinking of differentials in the old way we find

$$\begin{aligned} df(X_t) &= f'(X_t)\,dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 + \dots \\ &= f'(X_t)(A_t\,dt + G_t\,dB_t) + \frac{1}{2}f''(X_t)(A_t\,dt + G_t\,dB_t)^2 + \dots \\ &= f'(X_t)(A_t\,dt + G_t\,dB_t) + \frac{1}{2}f''(X_t)(A_t^2\,dt^2 + 2A_tG_t\,dt\,dB_t + G_t^2\,(dB_t)^2) + \dots \end{aligned}$$

At this point in ordinary calculus only the first term is considered, all the terms in $dt^2$ or of higher order being negligible in comparison to $dt$. But now it turns out that $dB_t$ "behaves as" $\sqrt{dt}$ so that the term $\frac{1}{2}f''(X_t)G_t^2\,(dB_t)^2$ is no longer negligible with respect to $dt$...

## 8.2 Application: exponential martingales

If $X \in M^2_{loc}([0,T])$ and $\lambda \in \mathbb{C}$ let

$$\begin{aligned} Z_t &= \lambda \int_0^t X_s\,dB_s - \frac{\lambda^2}{2}\int_0^t X_s^2\,ds \\ Y_t &= \mathrm{e}^{Z_t}\,. \end{aligned} \tag{8.9}$$

Let us compute the stochastic differential of $Y$. We must apply Ito's formula to $t \mapsto f(Z_t)$ with $f(x) = \mathrm{e}^x$: as $f'(x) = f''(x) = \mathrm{e}^x$, we find

$$\begin{aligned} dY_t &= \mathrm{e}^{Z_t}\,dZ_t + \frac{1}{2}\,\mathrm{e}^{Z_t}d\langle Z\rangle_t \\ &= \mathrm{e}^{Z_t}\Big(\lambda X_t\,dB_t - \frac{\lambda^2}{2}X_t^2\,dt\Big) + \frac{1}{2}\,\mathrm{e}^{Z_t}\,\lambda^2 X_t^2\,dt = \lambda X_t Y_t\,dB_t\,, \end{aligned} \tag{8.10}$$

i.e., keeping in mind that $Y_0 = 1$,

$$Y_t = 1 + \lambda \int_0^t Y_s X_s\,dB_s\,. \tag{8.11}$$

Therefore $Y$ is a local martingale. If $\lambda \in \mathbb{R}$ it is a positive local martingale and, by Proposition 7.5, a supermartingale; in particular, if $\lambda \in \mathbb{R}$, then

$$\mathrm{E}(Y_t) \le \mathrm{E}(Y_0) = 1$$

for every $t \le T$.

In the future, we shall often deal with the problem of proving that, if $\lambda \in \mathbb{R}$, $Y$ is a martingale. Indeed, there are two ways to do this:

1) by showing that $\mathrm{E}(Y_t) = 1$ for every $t$ (a supermartingale is a martingale if and only if it has constant expectation, Exercise 5.1);
2) by proving that $YX \in M^2([0,T])$. In this case $Y$ is even square integrable.

We know (Lemma 3.2) the behavior of the "tail" of the distribution of $B_t$:

$$\mathrm{P}(|B_t| \geq x) \leq \text{const}\ \mathrm{e}^{-\frac{1}{2t}x^2}\ ,$$

and the reflection principle allows us to state a similar behavior for the tail of the r.v. $\sup_{0\leq t\leq T}|B_t|$. The next statement allows us to say that a similar behavior is shared by stochastic integrals, under suitable conditions on the integrand.

**Proposition 8.2 (Exponential bound)** Let $X \in M^2([0,T])$ be such that $\int_0^T X_s^2\,ds \leq k$ a.s. for some constant $k > 0$. Then, for every $x > 0$,

$$\mathrm{P}\Big(\sup_{0\leq t\leq T}\Big|\int_0^t X_s\,dB_s\Big| \geq x\Big) \leq 2\mathrm{e}^{-\frac{x^2}{2k}}\ . \tag{8.12}$$

*Proof* Let $M_t = \int_0^t X_s\,dB_s$ and $A_t = \langle M\rangle_t = \int_0^t X_s^2\,ds$. Then as $A_t \leq k$, for every $\theta > 0$,

$$\{M_t \geq x\} = \{\mathrm{e}^{\theta M_t} \geq \mathrm{e}^{\theta x}\} \subset \{\mathrm{e}^{\theta M_t - \frac{1}{2}\theta^2 A_t} \geq \mathrm{e}^{\theta x - \frac{1}{2}\theta^2 k}\}\ .$$

The maximal inequality (5.15) applied to the supermartingale $t \mapsto \mathrm{e}^{\theta M_t - \frac{1}{2}\theta^2 A_t}$ gives

$$\mathrm{P}\Big(\sup_{0\leq t\leq T} M_t \geq x\Big) \leq \mathrm{P}\Big(\sup_{0\leq t\leq T} \mathrm{e}^{\theta M_t - \frac{1}{2}\theta^2 A_t} \geq \mathrm{e}^{\theta x - \frac{1}{2}\theta^2 k}\Big) \leq \mathrm{e}^{-\theta x + \frac{1}{2}\theta^2 k}\ .$$

This inequality holds for every $\theta > 0$. The minimum of $\theta \mapsto -\theta x + \frac{1}{2}\theta^2 k$ for $\theta > 0$ is attained at $\theta = \frac{x}{k}$ and its value is $-\frac{x^2}{2k}$. Substituting this value into the right-hand side we get

$$\mathrm{P}\Big(\sup_{0\leq t\leq T} M_t \geq x\Big) \leq \mathrm{e}^{-\frac{x^2}{2k}}\ . \tag{8.13}$$

The same argument applied to $-M$ gives

$$\mathrm{P}\Big(\inf_{0\leq t\leq T} M_t \leq -x\Big) \leq \mathrm{e}^{-\frac{x^2}{2k}}\ ,$$

which allows us to conclude the proof. $\square$

If the process $X$ appearing in (8.9) is of the form $\sum_{i=1}^m \lambda_i X_i(s)$ with $X_1, \dots, X_m \in M^2_{loc}([0,T])$, $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{C}^m$, then we obtain that if

$$Y_t = \exp\Big(\int_0^t \sum_{i=1}^m \lambda_i X_i(s)\, dB_s - \frac{1}{2}\int_0^t \sum_{i,j=1}^m \lambda_i\lambda_j X_i(s)X_j(s)\, ds\Big)\,,$$

we have

$$Y_t = 1 + \sum_{i=1}^m \lambda_i \int_0^t Y_s X_i(s)\, dB_s\,, \tag{8.14}$$

i.e. $Y$ is a local martingale. This is the key tool in order to prove the following important result.

**Theorem 8.2** Let $X \in M^2_{loc}([0,+\infty[)$ be such that

$$\int_0^{+\infty} X_s^2\, ds = +\infty \quad \text{a.s.} \tag{8.15}$$

Then, if $\tau(t) = \inf\{u; \int_0^u X_s^2\, ds > t\}$, the process

$$W_t = \int_0^{\tau(t)} X_s\, dB_s$$

is a (continuous) Brownian motion.

*Proof* We first compute, using (8.14), the finite-dimensional distributions of $W$, verifying that they coincide with those of a Brownian motion. Next we shall prove that $W$ is continuous. Let $0 \le t_1 < \dots < t_m$; as $\tau(t)$ is a finite stopping time for every $t$,

$$W_{t_j} = \int_0^{\tau(t_j)} X_s\, dB_s = \lim_{t\to+\infty}\int_0^{\tau(t_j)\wedge t} X_s\, dB_s = \lim_{t\to+\infty}\int_0^t X_s 1_{\{s<\tau(t_j)\}}\, dB_s\,. \tag{8.16}$$

In order to determine the joint distribution of $(W_{t_1}, \dots, W_{t_m})$ we investigate the characteristic function

$$\mathrm{E}\Big[\exp\Big(\sum_{j=1}^m i\theta_j W_{t_j}\Big)\Big] = \mathrm{E}\Big[\exp\Big(\sum_{j=1}^m i\theta_j \int_0^{\tau(t_j)} X_s\, dB_s\Big)\Big]\,.$$

Let us write (8.14) setting $X_j(s) = X_s 1_{\{s<\tau(t_j)\}}$ and $\lambda_j = i\theta_j$, $\theta_j \in \mathbb{R}$:

$$\begin{aligned}
Y_t &= \exp\Big[\sum_{j=1}^m i\theta_j \int_0^t X_j(s)\,dB_s + \frac{1}{2}\sum_{h,k=1}^m \theta_h\theta_k \int_0^t X_h(s)X_k(s)\,ds\Big] \\
&= \exp\Big[\sum_{j=1}^m i\theta_j \int_0^{t\wedge\tau(t_j)} X_s\,dB_s + \frac{1}{2}\sum_{h,k=1}^m \theta_h\theta_k \int_0^{t\wedge\tau(t_h)\wedge\tau(t_k)} X_s^2\,ds\Big]\,.
\end{aligned} \tag{8.17}$$

We have, by (8.14),

$$Y_t = 1 + \sum_{j=1}^m i\theta_j \int_0^t Y_s X_j(s)\,dB_s\,. \tag{8.18}$$

Moreover, $Y_t$ is bounded for every $t$ as

$$\int_0^t X_h(s)X_k(s)\,ds = \int_0^{t\wedge\tau(t_h)\wedge\tau(t_k)} X_s^2\,ds \le t_h \wedge t_k \tag{8.19}$$

and therefore, by (8.17),

$$|Y_t| \le \exp\Big[\frac{1}{2}\sum_{h,k=1}^m \theta_h\theta_k(t_h \wedge t_k)\Big]\,. \tag{8.20}$$

$Y$ is therefore a bounded local martingale, hence a martingale, from which we deduce that $\mathrm{E}(Y_t) = \mathrm{E}(Y_0) = 1$ for every $t$, i.e.

$$\mathrm{E}\Big[\exp\Big(\sum_{j=1}^m i\theta_j \int_0^{t\wedge\tau(t_j)} X_s\,dB_s + \frac{1}{2}\sum_{h,k=1}^m \theta_h\theta_k \int_0^{t\wedge\tau(t_h)\wedge\tau(t_k)} X_s^2\,ds\Big)\Big] = 1\,. \tag{8.21}$$

But

$$\lim_{t\to+\infty} \int_0^{t\wedge\tau(t_h)\wedge\tau(t_k)} X_s^2\,ds = \int_0^{\tau(t_h)\wedge\tau(t_k)} X_s^2\,ds = t_h \wedge t_k$$

and therefore, by (8.16) and Lebesgue's theorem (the r.v.'s $Y_t$ remain bounded thanks to (8.20)), taking the limit as $t \to +\infty$,

$$1 = \lim_{t\to+\infty} \mathrm{E}[Y_t] = \mathrm{E}\Big[\exp\Big(\sum_{j=1}^m i\theta_j W_{t_j} + \frac{1}{2}\sum_{h,k=1}^m \theta_h\theta_k(t_h \wedge t_k)\Big)\Big]\,,$$

i.e.

$$\mathrm{E}\Big[\exp\Big(\sum_{j=1}^m i\theta_j W_{t_j}\Big)\Big] = \exp\Big[-\frac{1}{2}\sum_{h,k=1}^m \theta_h\theta_k(t_h \wedge t_k)\Big]\,,$$

which is the characteristic function of a centered multivariate Gaussian r.v. with covariance matrix $\Gamma_{hk} = t_h \wedge t_k$. Therefore $W$ has the same finite-dimensional distributions as a Brownian motion and is actually a Brownian motion.

We now have to prove that the paths are continuous. Let $A_t = \int_0^t X_s^2\,ds$. If this process, which is obviously continuous, was strictly increasing then we would have $\tau(t) = A^{-1}(t)$. As the inverse of a continuous function is still continuous, $W$ would be continuous, being the composition of the two applications

$$s \mapsto \int_0^s X_u\,dB_u \qquad \text{and} \qquad t \mapsto \tau(t)\ .$$

If, conversely, $A$ is not strictly increasing, this simple argument does not work, as $t \mapsto \tau(t)$ might be discontinuous. For instance, if $A$ was constant on an interval $[a,b], a < b$ and $t = A_a = A_b$, then we would have

$$\tau(t) = \inf\{u; A_u > t\} \ge b\ ,$$

whereas $\tau(u) < a$ for every $u < t$; therefore necessarily $\lim_{u\to t-} \tau(u) \le a < b \le \tau(t)$. The arguments that follow are required in order to fix this technical point. However, the idea is simple: if $A$ is constant on $[a,b[$, then $X_t = 0$ a.e. on $[a,b[$, and therefore $t \mapsto \int_0^t X_u\,dB_u$ is itself constant on $[a,b[$.

If $a : \mathbb{R}^+ \to \mathbb{R}^+$ is a non-increasing right-continuous function, its *pseudo-inverse* is defined as

$$c(t) = \inf\{s; a(s) > t\}\ . \tag{8.22}$$

$\tau$ is therefore the pseudo-inverse of $A$.

**Proposition 8.3** Let $a$ be an increasing right-continuous function. Then

a) its pseudo-inverse $c$ is right-continuous.
b) If $\lim_{t\to t_0-} c(t) = L < c(t_0)$, then $a$ is constant $\equiv t_0$ on $[L, c(t_0)[$.

*Proof*

a) Let $(t_n)_n$ be a sequence decreasing to $t$ and let us prove that $\lim_{n\to\infty} c(t_n) = c(t)$; $c$ being increasing we have $\lim_{n\to\infty} c(t_n) \ge c(t)$; let us assume that $c(t_n) \searrow L > c(t)$ and let us prove that this is not possible. Let $u$ be a number such that $c(t) < u < L$. As $u < c(t_n)$ we have $a(u) \le t_n$ for every $n$. As $c(t) < u$ we have $t < a(u)$. These two inequalities are clearly incompatible.
b) If $L \le u < c(t_0)$, then $a(u) \ge t$ for every $t < t_0$ and therefore $a(u) \ge t_0$. On the other hand $u < c(t_0)$ implies $a(u) \le t_0$, and therefore $a(u) = t_0$. □

*End of the Proof of Theorem 8.2* As $\tau$ is the pseudo-inverse of $A$ and is therefore right-continuous, $W$ is right-continuous. If $\tau$ was left-continuous at $t_0$ then $W$ would be continuous at $t_0$. If, conversely, $\lim_{t\to t_0-}\tau(t) = L < \tau(t_0)$, then $A$ would be constant on $[L, \tau(t_0)[$ and therefore $X_s = 0$ a.e. on $[L, \tau(t_0)[$. Therefore by Proposition 7.2 $W_L = W_{\tau(t_0)}$ a.s. and $W$ is also left-continuous a.s. □

**Corollary 8.1** Under the hypotheses of Theorem 8.2, if $A_t = \int_0^t X_s^2\,ds$ we have

$$W_{A_t} = \int_0^t X_s\,dB_s\ .$$

*Proof* By definition, $A$ being increasing, $\tau(A_t) = \inf\{s; A_s > A_t\} \ge t$. If $\tau(A_t) = t$ the statement follows immediately from Theorem 8.2. If, conversely, $\tau(A_t) = L > t$ then $A$ is constant on $[L, t[$ and therefore $X_s \equiv 0$ on $[L, t[$ a.s.; then by Proposition 7.2 and Theorem 8.2

$$W_{A_t} = \int_0^{\tau(A_t)} X_s\,dB_s = \int_0^t X_s\,dB_s\ .$$

□

Corollary 8.1 states that a stochastic integral is a "time changed" Brownian motion, i.e. it is a process that "follows the same paths" of some Brownian motion $W$ but at a speed that changes as a function of $t$ and $\omega$. Moreover, the time change that determines this "speed" is given by a process $A$ that is nothing else than the associated increasing process of the stochastic integral.

Corollary 8.1 is obviously useful in order to obtain results concerning the regularity or the asymptotic of the paths of $t \mapsto \int_0^t X_s\,dB_s$, which can be deduced directly from those of the Brownian motion (Hölder continuity, Lévy's modulus of continuity, Iterated Logarithm Law, ...).

## 8.3 Application: $L^p$ estimates

We have seen in Chap. 7 some $L^2$ estimates for the stochastic integral (mostly thanks to the isometry property $M^2 \leftrightarrow L^2$). If, conversely, the integrand belongs to $M^p$, $p > 2$, is it true that the stochastic integral is in $L^p$? Ito's formula allows us to answer positively and to derive some useful estimates. We shall need them in the next chapter in order to derive regularity properties of the solutions of Stochastic Differential Equations with respect to parameters and initial conditions.

**Proposition 8.4** If $p \ge 2$ and $X \in M^2_{loc}([0,T])$, then

$$\mathrm{E}\Big(\sup_{0\le t\le T}\Big|\int_0^t X_s\,dB_s\Big|^p\Big) \le c_p \mathrm{E}\Big[\Big(\int_0^T |X_s|^2\,ds\Big)^{\frac{p}{2}}\Big] \le c_p T^{\frac{p-2}{2}} \mathrm{E}\Big(\int_0^T |X_s|^p\,ds\Big)$$

for some constant $c_p > 0$.

*Proof* One can of course assume $X \in M^p([0,T])$, otherwise the statement is obvious (the right-hand side is $= +\infty$). Let $I_t = \int_0^t X_s\,dB_s$ and define $I_t^* = \sup_{0\le s\le t} |I_s|$. $(I_t)_t$ is a square integrable martingale and by Doob's inequality (Theorem 5.12)

$$\begin{aligned}\mathrm{E}\Big(\sup_{0\le t\le T}\Big|\int_0^t X_s\,dB_s\Big|^p\Big) &= \mathrm{E}[I_t^{*p}] \le \Big(\frac{p}{p-1}\Big)^p \sup_{0\le t\le T} \mathrm{E}\big[|I_t|^p\big] \\ &= \Big(\frac{p}{p-1}\Big)^p \mathrm{E}[|I_T|^p]\,.\end{aligned} \tag{8.23}$$

Let us apply Ito's formula to the function $f(x) = |x|^p$ (which is twice differentiable, as $p \ge 2$) and to the process $I$ whose stochastic differential is $dI_t = X_t\,dB_t$.

We have $f'(x) = p\,\mathrm{sgn}(x)|x|^{p-1}$, $f''(x) = p(p-1)|x|^{p-2}$, where sgn denotes the "sign" function ($= 1$ for $x \ge 0$ and $-1$ for $x < 0$). Then by Ito's formula

$$\begin{aligned} d|I_t|^p &= f'(I_t)\,dI_t + \frac{1}{2} f''(I_t)\,d\langle I\rangle_t \\ &= |I_s|^{p-1}\,\mathrm{sgn}(I_s) X_s\,dB_s + \frac{1}{2} p(p-1)|I_s|^{p-2} X_s^2\,ds\,, \end{aligned}$$

i.e., as $I_0 = 0$,

$$|I_t|^p = p\int_0^t |I_s|^{p-1}\,\mathrm{sgn}(I_s) X_s\,dB_s + \frac{1}{2}p(p-1)\int_0^t |I_s|^{p-2} X_s^2\,ds\,. \tag{8.24}$$

Let us now first assume $|I_T^*| \le K$: this guarantees that $|I_s|^{p-1}\,\mathrm{sgn}(I_s) X_s \in M^2([0,T])$. Let us take the expectation in (8.24) recalling that the stochastic integral has zero mean. By Doob's inequality, (8.23) and Hölder's inequality with the exponents $\frac{p}{2}$ and $\frac{p}{p-2}$, we have

$$\begin{aligned} \mathrm{E}[I_T^{*p}] &\le \Big(\frac{p}{p-1}\Big)^p \mathrm{E}[|I_T|^p] = \underbrace{\frac{1}{2}\Big(\frac{p}{p-1}\Big)^p p(p-1)}_{:=c_0} \mathrm{E}\Big(\int_0^T |I_s|^{p-2} X_s^2\,ds\Big) \\ &\le c_0\,\mathrm{E}\Big(I_T^{*p-2}\int_0^T X_s^2\,ds\Big) \le c_0\,\mathrm{E}[I_T^{*p}]^{1-\frac{2}{p}}\,\mathrm{E}\Big[\Big(\int_0^T X_s^2\,ds\Big)^{\frac{p}{2}}\Big]^{\frac{2}{p}}\,. \end{aligned}$$

As we assume $|I_T^*| \le K$, $\mathrm{E}[I_T^{*p}] < +\infty$ and in the previous inequality we can divide by $\mathrm{E}[I_T^{*p}]^{1-\frac{2}{p}}$, which gives

$$\mathrm{E}[I_T^{*p}]^{\frac{2}{p}} \le c_0\,\mathrm{E}\Big[\Big(\int_0^t X_s^2\,ds\Big)^{\frac{p}{2}}\Big]^{\frac{2}{p}},$$

i.e.

$$\mathrm{E}[I_T^{*p}] \le c_0^{p/2}\,\mathrm{E}\Big[\Big(\int_0^t X_s^2\,ds\Big)^{\frac{p}{2}}\Big]. \tag{8.25}$$

If, conversely, $I$ is not bounded, then let $\tau_n = \inf\{t \le T; |I_t| \ge n\}$ ($\tau(n) = T$ if $\{\ \} = \emptyset$). $(\tau_n)_n$ is a sequence of stopping times increasing to $T$, as the paths of $I$ are continuous and then also bounded. We have therefore $I_{\tau_n \wedge t} \to I_t$ as $n \to \infty$ and

$$I_{t\wedge\tau_n} = \int_0^{t\wedge\tau_n} X_s\,dB_s = \int_0^t X_s 1_{\{s<\tau_n\}}\,dB_s$$

so that (8.25) gives

$$\mathrm{E}(I_{T\wedge\tau_n}^{*}{}^p) \le c_0^{p/2}\mathrm{E}\Big[\Big(\int_0^T |X_s|^2 1_{\{s<\tau_n\}}\,ds\Big)^{\frac{p}{2}}\Big] \le c_0^{p/2}\mathrm{E}\Big[\Big(\int_0^T |X_s|^2\,ds\Big)^{\frac{p}{2}}\Big],$$

and we can just apply Fatou's lemma. Finally, again by Hölder's inequality,

$$\mathrm{E}\Big[\Big(\int_0^T |X_s|^2\,ds\Big)^{\frac{p}{2}}\Big] \le T^{\frac{p-2}{p}}\mathrm{E}\Big[\int_0^T |X_s|^p\,ds\Big].$$

□

## 8.4 The multidimensional stochastic integral

Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a standard $d$-dimensional continuous Brownian motion and let $B_t = (B_1(t), \dots, B_d(t))$.

A matrix-valued process $\sigma(s) = (\sigma_{ij}(s))_{i,j}$ $i = 1, \dots, m$, $j = 1, \dots, d$, is said to be in $M_{loc}^p([a, b])$, $0 \le a < b$, if $(\sigma_{ij}(s))_s$ is progressively measurable for every $i, j$ and

i) for every $i, j$ $\sigma_{ij} \in M_{loc}^p$, i.e. $\displaystyle\int_a^b |\sigma_{ij}(s)|^p\,ds < +\infty$ a.s.

We say that $\sigma \in M^p([a, b])$ if

i') for every $i, j$ $\sigma_{ij} \in M^p$, i.e. $\mathrm{E}\Big[\displaystyle\int_a^b |\sigma_{ij}(s)|^p\,ds\Big] < +\infty$.

Similarly to the one-dimensional case, for fixed $d, m$, the space $M^2([a,b])$ turns to be a Hilbert space with norm

$$\|\sigma\|^2 = \mathrm{E}\Big[\int_a^b |\sigma(s)|^2\, ds\Big]\,,$$

where

$$|\sigma(s)|^2) = \sum_{i=1}^m \sum_{j=1}^d \sigma_{ij}(s)^2 = \mathrm{tr}(\sigma(s)\sigma(s)^*)\,.$$

If $\sigma \in M^2_{loc}([a,b])$ then, for every $i, j$, the stochastic integral $\int_a^b \sigma_{ij}(t)\, dB_j(t)$ is already defined. Let

$$\int_a^b \sigma(t)\, dB_t = \Big(\sum_{j=1}^d \int_a^b \sigma_{ij}(t)\, dB_j(t)\Big)_i\,. \tag{8.26}$$

The stochastic integral in (8.26) is an $\mathbb{R}^m$-valued r.v. Note that the matrix $\sigma(s)$ is, in general, rectangular and that the process defined by the stochastic integral has a dimension ($m$) which may be different from the dimension of the Brownian motion ($d$). It is clear, by the properties of the stochastic integral in dimension 1, that, if $\sigma \in M^2_{loc}([0,T])$, the process $I_t = \int_0^t \sigma(s)\, dB_s$ has a continuous version and every component is a local martingale (it is easy to see that a sum of local martingales with respect to the same filtration is still a local martingale).

In the next statements we determine the associated increasing processes to these local martingales.

**Lemma 8.1** Let $X_1, X_2 \in M^2([0,T])$ be real processes and

$$I_1(t) = \int_0^t X_1(u)\, dB_1(u), \quad I_2(t) = \int_0^t X_2(u)\, dB_2(u)\,.$$

Then, for $0 \le s \le t \le T$,

$$\mathrm{E}\Big(\int_s^t X_1(u)\, dB_1(u) \int_s^t X_2(u)\, dB_2(u) \,\Big|\, \mathscr{F}_s\Big) = 0\,. \tag{8.27}$$

In particular, the process

$$t \mapsto I_1(t) I_2(t)$$

is a martingale and $\langle I_1, I_2\rangle_t \equiv 0$.

*Proof* Let us first prove the statement for elementary processes. Let

$$X_1(t) = \sum_{j=1}^n X_i^{(1)} 1_{[t_i,t_{i+1}[}, \qquad X_2(t) = \sum_{j=1}^n X_i^{(2)} 1_{[t_i,t_{i+1}[} \,.$$

Then

$$\begin{aligned} &\mathrm{E}\Big[\int_s^t X_1(u)\,dB_1(u)\int_s^t X_2(u)\,dB_2(u)\,\Big|\,\mathscr{F}_s\Big] \\ &\mathrm{E}\Big[\sum_{i,j=1}^n X_i^{(1)}\big(B_1(t_{i+1})-B_1(t_i)\big)X_j^{(2)}\big(B_2(t_{j+1})-B_2(t_j)\big)\,\Big|\,\mathscr{F}_s\Big] \\ &= \sum_{i,j=1}^n \mathrm{E}\big[X_i^{(1)}\big(B_1(t_{i+1})-B_1(t_i)\big)X_j^{(2)}\big(B_2(t_{j+1})-B_2(t_j)\big)\,\big|\,\mathscr{F}_s\big] \,. \end{aligned} \tag{8.28}$$

Note first that all the terms appearing in the conditional expectation on the right-hand side are integrable, as the r.v.'s $X_i^{(k)}\big(B_k(t_{i+1})-B_k(t_i)\big)$ are square integrable, being the product of square integrable and independent r.v.'s (recall that $X_i^{(k)}$ is $\mathscr{F}_{t_i}$-measurable). Now if $t_i < t_j$ then, as the r.v.'s $X_i^{(1)}, B_1(t_{i+1})-B_1(t_i), X_j^{(2)}$ are already $\mathscr{F}_{t_j}$-measurable,

$$\begin{aligned} &\mathrm{E}\big[X_i^{(1)}\big(B_1(t_{i+1})-B_1(t_i)\big)X_j^{(2)}\big(B_2(t_{j+1})-B_2(t_j)\big)\,\big|\,\mathscr{F}_s\big] \\ &= \mathrm{E}\big[\mathrm{E}\big[X_i^{(1)}\big(B_1(t_{i+1})-B_1(t_i)\big)X_j^{(2)}\big(B_2(t_{j+1})-B_2(t_j)\big)\,\big|\,\mathscr{F}_{t_j}\big]\,\big|\,\mathscr{F}_s\big] \\ &= \mathrm{E}\big[X_i^{(1)}\big(B_1(t_{i+1})-B_1(t_i)\big)X_j^{(2)}\underbrace{\mathrm{E}\big[\big(B_2(t_{j+1})-B_2(t_j)\big)\,\big|\,\mathscr{F}_{t_j}\big]}_{=0}\,\big|\,\mathscr{F}_s\big] = 0 \,. \end{aligned}$$

Similarly one proves that the terms with $i = j$ in (8.28) also vanish thanks to the relation

$$\begin{aligned} &\mathrm{E}[(B_1(t_{i+1})-B_1(t_i))(B_2(t_{i+1})-B_2(t_i))\,|\,\mathscr{F}_{t_i}] \\ &= \mathrm{E}[(B_1(t_{i+1})-B_1(t_i))(B_2(t_{i+1})-B_2(t_i))] = 0 \,, \end{aligned}$$

where we use first the independence of $(B_1(t_{i+1})-B_1(t_i))(B_2(t_{i+1})-B_2(t_i))$ with respect to $\mathscr{F}_{t_i}$ and then the independence of $B_1(t_{i+1})-B_1(t_i)$ and $B_2(t_{i+1})-B_2(t_i)$.

(8.27) then follows as the stochastic integrals are the limit in $L^2$ of the stochastic integrals of approximating elementary processes. □

Let now $\sigma \in M^2([0,T])$ and

$$X_t = \int_0^t \sigma(u)\,dB_u \,. \tag{8.29}$$

What is the associated increasing process of the square integrable martingale $(X_i(t))_t$ (the $i$-th component of $X$)? If we define

$$I_k(t) = \int_0^t \sigma_{ik}(u)\, dB_k(u)$$

then $X_i(t) = I_1(t) + \cdots + I_d(t)$ and

$$X_i(t)^2 = \Big(\sum_{k=1}^d I_k(t)\Big)^2 = \sum_{h,k=1}^d I_h(t) I_k(t)\;.$$

By Lemma 8.1 $I_h I_k$, $h \neq k$, is a martingale. Moreover, (Proposition 7.3)

$$I_k(t)^2 - \int_0^t \sigma_{ik}^2(s)\, ds$$

is a martingale. Therefore

$$X_i(t)^2 - \sum_{k=1}^d \int_0^t \sigma_{ik}^2(s)\, ds \tag{8.30}$$

is a martingale, i.e.

$$\langle X_i \rangle_t = \sum_{k=1}^d \int_0^t \sigma_{ik}^2(s)\, ds\;. \tag{8.31}$$

A repetition of the approximation arguments of Sect. 7.6 gives that if, conversely, $\sigma \in M^2_{loc}([0,T])$, then the process (8.30) is a local martingale so that again its associated increasing process is given by (8.31). From (8.31) we have

$$\langle X_i + X_j \rangle_t = \sum_{k=1}^d \int_0^t (\sigma_{ik}(s) + \sigma_{jk}(s))^2\, ds$$

$$\langle X_i - X_j \rangle_t = \sum_{k=1}^d \int_0^t (\sigma_{ik}(s) - \sigma_{jk}(s))^2\, ds$$

and, using the formula $\langle X_i, X_j \rangle_t = \frac{1}{4}\,(\langle X_i + X_j \rangle_t - \langle X_i - X_j \rangle_t)$, we obtain

$$\langle X_i, X_j \rangle_t = \sum_{k=1}^d \int_0^t \sigma_{ik}(s)\sigma_{jk}(s)\, ds = \int_0^t a_{ij}(s)\, ds\;, \tag{8.32}$$

where $a(s) = \sigma(s)\sigma(s)^*$. As a consequence we have the following, which shows that also in the multidimensional case the stochastic integral is an isometry between $M^2$ and $L^2$.

**Proposition 8.5** If $\sigma \in M^2([0,T])$ and $X$ is as in (8.29) then

$$\mathrm{E}\big[(X_i(t_2) - X_i(t_1))(X_j(t_2) - X_j(t_1))\big] = \mathrm{E}\Big(\int_{t_1}^{t_2} a_{ij}(s)\,ds\Big)$$

(recall that $a(s) = \sigma(s)\sigma(s)^*$) and

$$\mathrm{E}\Big(\Big|\int_{t_1}^{t_2} \sigma(t)\,dB_t\Big|^2\Big) = \mathrm{E}\Big(\int_{t_1}^{t_2} |\sigma(t)|^2\,dt\Big) = \mathrm{E}\Big(\int_{t_1}^{t_2} \mathrm{tr}\,a(t)\,dt\Big)\,. \qquad (8.33)$$

*Proof* As both $X_i$ and $X_j$ are martingales, we have for $t_1 \le t_2$

$$\mathrm{E}[X_i(t_2)X_j(t_1)] = \mathrm{E}\big[\mathrm{E}[X_i(t_2)X_j(t_1)\,|\,\mathscr{F}_{t_1}]\big] = \mathrm{E}[X_i(t_1)X_j(t_1)]\,.$$

Therefore

$$\begin{aligned} &\mathrm{E}\big[(X_i(t_2) - X_i(t_1))(X_j(t_2) - X_j(t_1))\big] \\ &= \mathrm{E}\big[X_i(t_2)X_j(t_2) + X_i(t_1)X_j(t_1) - X_i(t_1)X_j(t_2) - X_i(t_2)X_j(t_1)\big] \\ &= \mathrm{E}\big[X_i(t_2)X_j(t_2) - X_i(t_1)X_j(t_1)\big] \end{aligned}$$

but as the process $(X_i(t)X_j(t))_t$ is equal to a martingale vanishing at 0 plus the process $\langle X_i, X_j\rangle_t$ we have

$$\cdots = \mathrm{E}\big[\langle X_i, X_j\rangle_{t_2} - \langle X_i, X_j\rangle_{t_1}\big] = \mathrm{E}\Big(\int_{t_1}^{t_2} a_{ij}(s)\,ds\Big)\,.$$

Now (8.33) follows because

$$\mathrm{E}(|X(t_2) - X(t_1)|^2) = \sum_{i=1}^m \mathrm{E}[(X_i(t_2) - X_i(t_1))^2] = \mathrm{E}\Big(\int_{t_1}^{t_2} \sum_{i=1}^m a_{ii}(t)\,dt\Big)\,.$$

□

As is the case for the stochastic integral in dimension 1, the martingale property allows us to derive some important estimates. The following is a form of Doob's maximal inequality.

**Proposition 8.6** If $\sigma \in M^2([0,T])$ then

$$\mathrm{E}\Big[\sup_{0\le t\le T}\Big|\int_0^t \sigma(s)\,dB_s\Big|^2\Big] \le 4\mathrm{E}\Big[\int_0^T |\sigma(s)|^2\,ds\Big]\,.$$

*Proof*

$$\mathrm{E}\Big[\sup_{0\le t\le T}\Big|\int_0^t \sigma(s)\,dB_s\Big|^2\Big] = \mathrm{E}\Big[\sup_{0\le t\le T}\sum_{i=1}^m\Big(\int_0^t\sum_{j=1}^d \sigma_{ij}(s)\,dB_j(s)\Big)^2\Big]$$
$$\le \sum_{i=1}^m \mathrm{E}\Big[\sup_{0\le t\le T}\Big(\int_0^t\sum_{j=1}^d \sigma_{ij}(s)\,dB_j(s)\Big)^2\Big]\,.$$

By Doob's inequality (Theorem 5.12),

$$\mathrm{E}\Big[\sup_{0\le t\le T}\Big(\int_0^t\sum_{j=1}^d \sigma_{ij}(s)\,dB_j(s)\Big)^2\Big] \le 4\mathrm{E}\Big[\Big(\int_0^T\sum_{j=1}^d \sigma_{ij}(s)\,dB_j(s)\Big)^2\Big]$$
$$= 4\int_0^T \mathrm{E}\Big[\sum_{h=1}^d \sigma_{ih}(s)\sigma_{ih}(s)\Big]\,ds = 4\int_0^T \mathrm{E}\big[(\sigma\sigma^*)_{ii}(s)\big]\,ds$$

and now just take the sum in $i$, recalling that $|\sigma(s)|^2 = \mathrm{tr}\,\sigma(s)\sigma(s)^*$. □

We say that an $\mathbb{R}^m$-valued process $X$ has stochastic differential

$$dX_t = F_t\,dt + G_t\,dB_t\,,$$

where $F = (F_1,\dots,F_m) \in M^1_{loc}([0,T])$ and $G = (G_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,d}} \in M^2_{loc}([0,T])$, if for every $0 \le t_1 < t_2 \le T$ we have

$$X_{t_2} - X_{t_1} = \int_{t_1}^{t_2} F_t\,dt + \int_{t_1}^{t_2} G_t\,dB_t\,.$$

Again do not be confused: the Brownian motion here is $d$-dimensional whereas the process $X$ turns out to be $m$-dimensional. Note that, also in the multidimensional case, the stochastic differential is unique: just argue coordinate by coordinate, the details are left to the reader.

In the multidimensional case Ito's formula takes the following form, which extends Theorem 8.1. We shall not give the proof (which is however similar).

**Theorem 8.3** Let $u : \mathbb{R}^m_x \times \mathbb{R}^+_t \to \mathbb{R}$ be a continuous function with continuous derivatives $u_t, u_{x_i}, u_{x_i x_j}, i, j = 1, \dots, m$. Let $X$ be a process having stochastic differential

$$dX_t = F_t\,dt + G_t\,dB_t\;.$$

Then the process $(u(X_t, t))_t$ has stochastic differential

$$\begin{aligned} du(X_t, t) &= u_t(X_t, t)\,dt + \sum_{i=1}^m u_{x_i}(X_t, t)\,dX_i(t) \\ &\quad + \frac{1}{2}\sum_{i,j=1}^m u_{x_i x_j}(X_t, t) A_{ij}(t)\,dt\;, \end{aligned} \tag{8.34}$$

where $A = GG^*$.

Thanks to (8.32) we have $A_{ij}(t)\,dt = d\langle X_i, X_j\rangle_t$, so that (8.34) can also be written as

$$du(X_t, t) = u_t(X_t, t)\,dt + \sum_{i=1}^m u_{x_i}(X_t, t)\,dX_i(t) + \frac{1}{2}\sum_{i,j=1}^m u_{x_i x_j}(X_t, t) d\langle X_i, X_j\rangle_t\;.$$

*Example 8.7* Let $u : \mathbb{R}^d \to \mathbb{R}$ be a twice differentiable function. Then, denoting by $u'$ the gradient of $u$,

$$du(B_t) = u'(B_t)\,dB_t + \frac{1}{2} \triangle\, u(B_t)\,dt\;. \tag{8.35}$$

Actually in this case $d\langle B_i, B_j\rangle_t = 0$ for $i \neq j$ and $d\langle B_i, B_i\rangle_t = dt$ so that

$$\frac{1}{2}\sum_{i,j=1}^d u_{x_i x_j}(B_t) d\langle B_i, B_j\rangle_t = \frac{1}{2} \triangle\, u(B_t)\,dt\;.$$

In particular, if $i \neq j$, applying the previous remark to the function $u(x) = x_i x_j$, whose Laplacian is equal to 0,

$$d(B_i(t)B_j(t)) = B_i(t)\,dB_j(t) + B_j(t)\,dB_i(t)\;. \tag{8.36}$$

We find again that $B_iB_j$ is a martingale for $i \neq j$ (see Exercise 5.24). In addition we obtain immediately the associated increasing process $A$ of this

(continued)

*Example 8.7* (continued)
martingale: as

$$B_i(t)^2 B_j(t)^2 = \Big(\int_0^t B_i(t)\,dB_j(t) + \int_0^t B_j(t)\,dB_i(t)\Big)^2$$

$$= \Big(\int_0^t B_j(t)\,dB_i(t)\Big)^2 + \Big(\int_0^t B_i(t)\,dB_j(t)\Big)^2 + 2\int_0^t B_j(t)\,dB_i(t)\int_0^t B_i(t)\,dB_j(t)$$

and we know that the last term on the right-hand side is already a martingale (Lemma 8.1), we have that if

$$A_t = \int_0^t (B_i(s)^2 + B_j(s)^2)\,ds$$

then $t \mapsto B_i(t)^2 B_j(t)^2 - A_t$ is a martingale. Hence $A$ is the requested associated increasing process.

*Example 8.8 (Bessel square processes)* What is the stochastic differential of $X_t = |B_t|^2$?

We must compute the stochastic differential of $u(B_t)$ where $u : \mathbb{R}^d \to \mathbb{R}$ is given by $u(x) = |x|^2 = x_1^2 + \dots + x_d^2$. Obviously

$$\frac{\partial u}{\partial x_i}(x) = 2x_i, \qquad \frac{\partial^2 u}{\partial x_i \partial x_j}(x) = 2\delta_{ij}$$

i.e. $u'(x) = 2x$ whereas $\triangle u = 2d$. Therefore by (8.35)

$$dX_t = d\,dt + 2B_t\,dB_t \tag{8.37}$$

and we find again another well-known fact: $|B_t|^2 - dt$ is a martingale. $X$ is a *Bessel square process*.

If $\theta = (\theta_1, \dots, \theta_m) \in \mathbb{R}^m$, $\sigma(s) = (\sigma_{ij}(s))_{\substack{i=1,\dots,m\\ j=1,\dots,d}} \in M^2_{loc}([0,T])$, let $a = \sigma\sigma^*$ and

$$X_t = \int_0^t \sigma(s)\,dB_s\ , \tag{8.38}$$

which is an $m$-dimensional process. Let

$$Y_t^\theta = \exp\Big(\Big\langle\theta, \int_0^t \sigma(s)\,dB_s\Big\rangle - \frac{1}{2}\int_0^t \langle a(s)\,\theta, \theta\rangle\,ds\Big)\ , \tag{8.39}$$

then, by Ito's formula applied to the function $u(x) = \mathrm{e}^{\langle \theta, x\rangle}$ and to the process

$$Z_t^\theta = \int_0^t \sigma(s)\, dB_s - \frac{1}{2}\int_0^t a(s)\theta\, ds$$

so that $Y_t^\theta = u(Z_t^\theta)$, we obtain

$$Y_t^\theta = 1 + \sum_{h=1}^m \theta_h \int_0^t Y_s^\theta\, dX_h(s) = 1 + \sum_{h=1}^m \sum_{j=1}^d \theta_h \int_0^t Y_s^\theta \sigma_{hj}(s)\, dB_j(s)\,. \tag{8.40}$$

$Y^\theta$ is a positive local martingale and therefore a supermartingale. In a way similar to Proposition 8.2 we get

**Proposition 8.7** Let $\sigma \in M^2([0,T])$ and $\theta \in \mathbb{R}^m$ with $|\theta| = 1$. Then if there exists a constant $k$ such that $\int_0^T \langle a(s)\,\theta, \theta\rangle\, ds \le k$ we have

$$\mathrm{P}\Big(\sup_{0\le t\le T} |\langle \theta, X_t\rangle| \ge x\Big) \le 2\mathrm{e}^{-\frac{x^2}{2k}}\,. \tag{8.41}$$

Moreover, if $\int_0^T \langle a(s)\,\theta, \theta\rangle\, ds \le k^*$ for every vector $\theta \in \mathbb{R}^m$ of modulus 1,

$$\mathrm{P}\Big(\sup_{0\le t\le T} |X_t| \ge x\Big) \le 2m\mathrm{e}^{-\frac{x^2}{2k^* m}}\,. \tag{8.42}$$

*Proof* The first part of the statement comes from a step by step repetition of the proof of Proposition 8.2: we have for $\lambda > 0$

$$\begin{aligned}
\mathrm{P}\Big(\sup_{0\le t\le T} |\langle \theta, X_t\rangle| \ge x\Big) &= \mathrm{P}\Big(\sup_{0\le t\le T} |\langle \lambda\theta, X_t\rangle| \ge \lambda x\Big)\\
&\le \mathrm{P}\Big(\sup_{0\le t\le T} |\langle \lambda\theta, X_t\rangle| - \frac{1}{2}\int_0^T \lambda^2 \langle a(s)\,\theta, \theta\rangle\, ds \ge \lambda x - \frac{1}{2}\lambda^2 k\Big)\\
&= \mathrm{P}\Big(\sup_{0\le t\le T} \exp\Big(|\langle \lambda\theta, X_t\rangle| - \frac{1}{2}\int_0^T \lambda^2 \langle a(s)\,\theta, \theta\rangle\, ds\Big) \ge \mathrm{e}^{\lambda x - \frac{1}{2}\lambda^2 k}\Big)\,.
\end{aligned}$$

Now $t \mapsto \exp(|\langle \lambda\theta, X_t\rangle| - \frac{1}{2}\int_0^T \lambda^2 \langle a(s)\,\theta, \theta\rangle\, ds)$ is a continuous supermartingale whose expectation is smaller than 1. Hence, by the maximal inequality (5.15),

$$\mathrm{P}\Big(\sup_{0\le t\le T} |\langle \theta, X_t\rangle| \ge x\Big) \le \mathrm{e}^{-\lambda x + \frac{1}{2}\lambda^2 k}$$

and taking the minimum of $\lambda \mapsto -\lambda x + \frac{1}{2}\lambda^2 k$ for $\lambda \in \mathbb{R}^+$ we obtain (8.41). Moreover, the inequality (8.41) applied to the coordinate vectors gives, for every $i = 1, \dots, m$,

$$\mathrm{P}\Big(\sup_{0\le t\le T} |X_i(t)| \ge \frac{x}{\sqrt{m}}\Big) \le 2\mathrm{e}^{-\frac{x^2}{2k^* m}}$$

and now just observe that if $\sup_{0\le t\le T} |X_t| \ge x$, then for one at least of the coordinates $i, i = 1, \dots, m$, necessarily $\sup_{0\le t\le T} |X_i(t)| \ge xm^{-1/2}$. Therefore

$$\mathrm{P}\Big(\sup_{0\le t\le T} |X_t| \ge x\Big) \le \sum_{i=1}^m \mathrm{P}\Big(\sup_{0\le t\le T} |X_i(t)| \ge \frac{x}{\sqrt{m}}\Big) \le 2m\mathrm{e}^{-\frac{x^2}{2k^* m}} .$$

□

**Theorem 8.4** Let $X = (X_1, \dots, X_d) \in M^2_{loc}([0, +\infty[)$ such that

$$\int_0^{+\infty} |X_s|^2\, ds = +\infty \qquad \text{a.s.}$$

Let $A_t = \int_0^t |X_s|^2\, ds$ and $\tau(t) = \inf\{s; A_s > t\}$; then the process

$$W_t = \int_0^{\tau(t)} X_s\, dB_s$$

is a (one-dimensional) Brownian motion such that $W(A_t) = \int_0^t X_s\, dB_s$.

*Proof* Identical to the proof of Theorem 8.2. □

In particular, if $|X_s| = 1$ a.s. for every $s$, then $A_t = t$ and therefore

**Corollary 8.2** Let $X = (X_1, \dots, X_d) \in M^2_{loc}([0, +\infty[)$ be a process such that $|X_s| = 1$ for almost every $s$ a.s. Then the process

$$W_t = \int_0^t X_s\, dB_s$$

is a Brownian motion.

Sometimes it would be useful to apply Corollary 8.2 to $X_t = \frac{1}{|B_t|}(B_1(t), \dots, B_d(t))$, which is obviously such that $|X_t| = 1$. Unfortunately this is not immediately possible, as $X$ is not defined if $|B_t| = 0$. In the following example we see how to go round this difficulty.

*Example 8.9 (Bessel square processes again)* We have seen in (8.37) that the process $X_t = |B_t|^2$ has the stochastic differential

$$dX_t = d\,dt + 2B_t\,dB_t\;. \tag{8.43}$$

Now let $z_0 \in \mathbb{R}^d$ be any vector such that $|z_0| = 1$ and let (recall that $\sqrt{X_t} = |B_t|$)

$$W_t = \int_0^t \Big(\frac{B_s}{\sqrt{X_s}}\,1_{\{X_s\neq 0\}} + z_0 1_{\{X_s=0\}}\Big)\,dB_s\;. \tag{8.44}$$

Therefore $W$ is a one-dimensional Brownian motion. Note that

$$\sqrt{X_t}\,dW(t) = B_t\,1_{\{X_t\neq 0\}}\,dB_t + \sqrt{X_t}\,1_{\{X_t=0\}}z_0\,dB_t = B_t\,dB_t$$

(of course $\sqrt{X_t}\,1_{\{X_t=0\}} = 0$) so that

$$B_t\,dB_t = \sum_{i=1}^d B_i(t)\,dB_i(t) = \sqrt{X_t}\,dW(t)\;.$$

(8.43) can now be written as

$$dX_t = d\,dt + 2\sqrt{X_t}\,dW_t\;. \tag{8.45}$$

Therefore the process $X_t = |B_t|^2$ is a solution of (8.45), another example of a Stochastic Differential Equation, which we shall investigate in the next chapter.

The same arguments developed in the proofs of Theorem 8.4 and Corollary 8.2 give the following

**Proposition 8.8** Let $(O_s)_s$ be a progressively measurable $O(d)$-valued (i.e. orthogonal matrix-valued) process. Then $(O_s)_s \in M^2([0,+\infty[)$ and

$$X_t = \int_0^t O_s\,dB_s$$

is itself an $(\mathscr{F}_t)_t$-$d$-dimensional Brownian motion.

*Proof* $(O_s)_s$ obviously belongs to $M^2([0,+\infty[)$ as all the coefficients of an orthogonal matrix are smaller, in modulus, than 1. By the criterion of Theorem 5.17, we just need to prove that, for every $\lambda \in \mathbb{R}^d$, $Y_t^\lambda = \mathrm{e}^{i\langle\lambda, X_t\rangle + \frac{1}{2}|\lambda|^2 t}$ is an $(\mathscr{F}_t)_t$-martingale. By Ito's formula, or directly by (8.39) and (8.40), we have

$$Y_t^\lambda = 1 + i\Big\langle\lambda, \int_0^t Y_s^\lambda O_s\, dB_s\Big\rangle\,.$$

Since, for $s \le t$, $|Y_s^\lambda| \le \mathrm{e}^{\frac{1}{2}|\lambda|^2 s}$, the process $(Y_s^\lambda O_s)_s$ is in $M^2([0,+\infty[)$ and therefore $Y^\lambda$ is an $(\mathscr{F}_t)_t$-martingale, which allows us to conclude the proof. □

We conclude this section with an extension to the multidimensional case of the $L^p$ estimates of Sect. 8.3.

**Proposition 8.9** If $\sigma \in M^2_{loc}([0,T])$, $p \ge 2$, then

$$\mathrm{E}\Big(\sup_{0\le t\le T}\Big|\int_0^t \sigma(s)\, dB_s\Big|^p\Big) \le c(p,m,d)T^{\frac{p-2}{2}}\,\mathrm{E}\Big(\int_0^t |\sigma(s)|^p\, ds\Big)\,.$$

*Proof* One can repeat the proof of Proposition 8.4, applying Theorem 8.3 to the function, $\mathbb{R}^m \to \mathbb{R}$, $u(x) = |x|^p$, or using the inequalities (also useful later)

$$\sum_{i=1}^m |x_i|^p \le |x|^p \le m^{\frac{p-2}{2}} \sum_{i=1}^m |x_i|^p \qquad x \in \mathbb{R}^m \tag{8.46}$$

$$|y_1 + \dots + y_d|^p \le d^{p-1} \sum_{j=1}^m |y_j|^p \qquad y \in \mathbb{R}^d\,. \tag{8.47}$$

By Proposition 8.4 and (8.46),

$$\begin{aligned}\mathrm{E}\Big(\sup_{0\le t\le T}\Big|\int_0^t \sigma(s)\, dB_s\Big|^p\Big) &\le m^{\frac{p-2}{2}}\mathrm{E}\Big(\sup_{0\le t\le T}\sum_{i=1}^m\Big|\sum_{j=1}^d\int_0^t \sigma_{ij}(s)\, dB_j(s)\Big|^p\Big)\\ &\le m^{\frac{p-2}{2}}\sum_{i=1}^m \mathrm{E}\Big(\sup_{0\le t\le T}\Big|\sum_{j=1}^d\int_0^t \sigma_{ij}(s)\, dB_j(s)\Big|^p\Big)\end{aligned}$$

and then by (8.47) and Proposition 8.4

$$\le m^{\frac{p-2}{2}} d^{p-1} \sum_{i=1}^m \sum_{j=1}^d \mathrm{E}\Big( \sup_{0\le t\le T} \Big| \int_0^t \sigma_{ij}(s)\, dB_j(s)\Big|^p\Big)$$

$$\le p^{3p/2}(2(p-1))^{-p/2} m^{\frac{p-2}{2}} d^{p-1} T^{\frac{p-2}{2}} \mathrm{E}\Big( \int_0^T \sum_{i=1}^m \sum_{j=1}^d |\sigma_{ij}(s)|^p\, ds\Big)$$

$$\le c(p,m,d) T^{\frac{p-2}{2}} \mathrm{E}\Big( \int_0^T |\sigma(s)|^p\, ds\Big)\,.$$

□

## 8.5 *A case study: recurrence of multidimensional Brownian motion

Let $B$ be an $m$-dimensional Brownian motion, with $m \ge 2$. We know (see Remark 3.4) that in dimension 1 the Brownian motion visits (infinitely many times) every real number with probability 1. What can be said in dimension $m \ge 2$? In dimension $m$ does Brownian motion also go everywhere?

There are two ways of precisely stating this kind of property for a generic process $X$.

p1) $X$ visits every open set $D \subset \mathbb{R}^m$ with probability 1.
p2) $X$ visits every point $x \in \mathbb{R}^m$ with probability 1.

A process satisfying p1) is said to be *recurrent*. A process that is not recurrent is said to be *transient*. Of course p2) entails p1).

Let us investigate whether the Brownian motion is transient or recurrent in dimension $\ge 2$. We shall treat the case $m \ge 3$, leaving the case $m = 2$ as an exercise (Exercise 8.23). Let $x \in \mathbb{R}^m$, $x \ne 0$: we want to investigate the probability for the Brownian motion to visit a neighborhood of $x$. It should not be surprising, as in other questions concerning the visit of a process to some set, that the problem boils down to finding the right martingale. Let

$$X_t = \frac{1}{|B_t - x|^{m-2}}\,\cdot \tag{8.48}$$

We can write $X_t = f(B_t)$, where $f(z) = |z-x|^{-(m-2)}$. $f$ has a singularity at $x$ but is bounded and infinitely many times differentiable outside every ball containing $x$. In order to apply Ito's formula, which in this case is given in (8.35), let us compute $f'(z)$ and $\triangle f(z)$, $z \ne x$. If $g(z) = |z-x|$, we have, for $z \ne x$,

$$\frac{\partial g}{\partial z_i}(z) = \frac{\partial}{\partial z_i}\sqrt{\sum_{j=1}^m (z_j - x_j)^2} = \frac{z_i - x_i}{|z-x|}$$

and therefore

$$\frac{\partial f}{\partial z_i}(z) = \frac{-(m-2)}{|z-x|^{m-1}}\,\frac{z_i - x_i}{|z-x|} = \frac{-(m-2)(z_i-x_i)}{|z-x|^m}$$

and

$$\frac{\partial^2 f}{\partial z_i^2}(z) = \frac{m(m-2)(z_i-x_i)^2}{|z-x|^{m+2}} - \frac{m-2}{|z-x|^m}$$

so that

$$\triangle f(z) = \frac{m(m-2)}{|z-x|^{m+2}}\sum_{i=1}^m (z_i-x_i)^2 - \frac{m(m-2)}{|z-x|^m} = 0\,.$$

$f$ is not a $C^2$ function on $\mathbb{R}^m$ (it is not even defined everywhere) so that we cannot directly apply Ito's formula in order to obtain the stochastic differential of $X_t = f(B_t)$. However, let $\widetilde{f}$ be a function that coincides with $f$ outside a small ball, $V_n(x)$, centered at $x$ and with radius $\frac{1}{n}$ with $\frac{1}{n} < |x|$ and extended inside the ball in order to be $C^2$. Then by Ito's formula (recall that $\widetilde{f}(0) = f(0) = |x|^{-(m-2)}$)

$$\widetilde{f}(B_t) = |x|^{-(m-2)} + \int_0^t \widetilde{f}'(B_s)\,dB_s + \int_0^t \frac{1}{2}\,\triangle\widetilde{f}(B_s)\,ds\,.$$

Let us denote by $\tau_n$ the entrance time of $B$ into the small ball. Then, as $f$ coincides with $\widetilde{f}$ outside the ball,

$$\begin{aligned} X_{t\wedge\tau_n} &= f(B_{t\wedge\tau_n}) = \widetilde{f}(B_{t\wedge\tau_n}) \\ &= |x|^{-(m-2)} + \int_0^{t\wedge\tau_n} f'(B_s)\,dB_s + \underbrace{\int_0^{t\wedge\tau_n} \frac{1}{2}\triangle f(B_s)\,ds}_{=0}\,. \end{aligned} \tag{8.49}$$

Therefore $(X_{t\wedge\tau_n})_t$ is a local martingale. As $f$ is bounded outside $V_n(x)$ ($f \le n^{m-2}$ outside the ball $V_n(x)$), $(X_{t\wedge\tau_n})_t$ is also bounded and is therefore a martingale (a bounded local martingale is a martingale, Remark 7.7). Let us deduce that $P(\tau_n < +\infty) < 1$ if $n$ is large enough. As $X_0 = |x|^{-(m-2)}$,

$$E(X_{t\wedge\tau_n}) = E(X_0) = |x|^{-(m-2)}$$

for every $t \ge 0$. As $X_{t\wedge\tau_n} = n^{m-2}$ on $\{\tau_n \le t\}$, we have $E(X_{t\wedge\tau_n}) \ge n^{m-2}P(\tau_n \le t)$ and therefore

$$|x|^{-(m-2)} \ge n^{m-2}P(\tau_n \le t) \tag{8.50}$$

so that $\mathrm{P}(\tau_n \le t) \le (n|x|)^{-(m-2)}$ and as $t \to +\infty$ we have $\mathrm{P}(\tau_n < +\infty) \le (n|x|)^{-(m-2)}$ and for $n$ large the last quantity is $< 1$. Therefore *with strictly positive probability the Brownian motion does not visit a small ball around* $x$. In dimension $\ge 3$ the Brownian motion is therefore transient.

From this estimate we deduce that, with probability 1, $B$ does not visit $x$: in order to visit $x$, $B$ must first enter into the ball of radius $\frac{1}{n}$ centered at $x$ so that for every $n$

$$\mathrm{P}(B_t = x \text{ for some } t > 0) \le \mathrm{P}(\tau_n < +\infty) \le \frac{1}{(n|x|)^{(m-2)}}$$

and, as $n$ is arbitrary,

$$\mathrm{P}(B_t = x \text{ for some } t > 0) = 0\,. \tag{8.51}$$

If $m = 2$, conversely, the Brownian motion is recurrent (see Exercise 8.23), but (8.51) is still true: even in dimension 2 the probability for the Brownian motion to visit a given point $x \ne 0$ is equal to 0.

The argument proving recurrence in dimension 2 is similar to the one developed here for dimension $\ge 3$ (find the right martingale...), but using the function $f_2(z) = -\log(|z - x|)$ instead. What makes the difference with the function $f$ of (8.48) is that $f_2$ is not bounded outside a neighborhood of $x$.

*Example 8.10 (A uniformly integrable local martingale which is not a martingale)* Since we now know that, with probability 1, $B$ does not visit $x$, the process

$$X_t = \frac{1}{|B_t - x|^{m-2}}$$

is well defined and, thanks to (8.49), is a local martingale. We now prove that it is uniformly integrable but not a martingale, which is a useful counterexample. In order to prove uniform integrability, let us show that $\sup_{t>0} \mathrm{E}(X_t^p) < +\infty$ for every $p < m(m-2)^{-1}$ and therefore that it is a process bounded in $L^p$ for some $p > 1$. Let $r = p(m-2)$; then

$$\mathrm{E}(X_t^p) = \mathrm{E}(|B_t - x|^{-r}) = \frac{1}{(2\pi t)^{m/2}} \int |y - x|^{-r} \mathrm{e}^{-|y|^2/2t}\, dy\,.$$

We split the integral into two parts: the first one on a ball $S_x$ centered at $x$ and with radius $\frac{1}{2}|x|$, the second one on its complement. As $|y - x|^{-r} \le |x|^{-r} 2^r$

(continued)

*Example 8.10* (continued)
outside of $S_x$,

$$\frac{1}{(2\pi t)^{m/2}}\int_{S_x^c}|y-x|^{-r}\mathrm{e}^{-|y|^2/2t}\,dy\le\frac{2^r}{|x|^r(2\pi t)^{m/2}}\int\mathrm{e}^{-|y|^2/2t}\,dy=\frac{2^r}{|x|^r}\,\cdot$$

The remaining term conversely is evaluated by observing that inside $S_x$ we have $|y|\ge\frac{1}{2}|x|$ and therefore $\mathrm{e}^{-|y|^2/2t}\le\mathrm{e}^{-|x|^2/4t}$. The integral over $S_x$ is therefore smaller than

$$\frac{1}{(2\pi t)^{m/2}}\,\mathrm{e}^{-|x|^2/4t}\int_{|y-x|\le\frac{1}{2}|x|}|y-x|^{-r}\,dy=\frac{1}{(2\pi t)^{m/2}}\,\mathrm{e}^{-|x|^2/4t}\int_{|z|\le\frac{1}{2}|x|}|z|^{-r}\,dy$$
$$=\mathrm{const}\,\frac{1}{(2\pi t)^{m/2}}\,\mathrm{e}^{-|x|^2/4t}\int_0^{\frac{1}{2}|x|}\rho^{m-1-r}\,d\rho.$$

If $p<m(m-2)^{-1}$, we have $r<m$ and the last integral is convergent. The quantity $\mathrm{E}(X_t^p)$ is bounded in $t$ and $X$ is uniformly integrable.

If $X$, which is uniformly integrable, were a martingale, then the limit $X_\infty=\lim_{t\to+\infty}X_t$ would exist a.s. and in $L^1$. But by the Iterated Logarithm Law there exists a.s. a sequence of times $(t_n)_n$ with $\lim_{n\to\infty}t_n=+\infty$ such that $\lim_{n\to\infty}|B_{t_n}|=+\infty$, therefore $X_\infty=0$ a.s., in contradiction with the fact that the convergence must also take place in $L^1$.

## Exercises

**8.1** (p. 533) Let $(\Omega,\mathscr{F},(\mathscr{F}_t)_t,(B_t)_t,\mathrm{P})$ be a real Brownian motion and let

$$M_t=(B_t+t)\,\mathrm{e}^{-(B_t+\frac{1}{2}t)}\,.$$

a) Compute the stochastic differential of $(M_t)_t$.
b) Show that $(M_t)_t$ is a martingale.

**8.2** (p. 534) Let $B=(\Omega,\mathscr{F},(\mathscr{F}_t)_t,(B_t)_t,\mathrm{P})$ be a Brownian motion and let $a,b>0$. Compute the stochastic differential of $X_t=(b-B_t)(a+B_t)+t$ and deduce that it is a martingale.

**8.3** (p. 535)

a) Let $(B_t)_t$ be a Brownian motion. Determine which of the following processes is a martingale

$$X_t=\mathrm{e}^{t/2}\sin B_t$$
$$Y_t=\mathrm{e}^{t/2}\cos B_t\,.$$

Compute $\mathrm{E}(X_t)$ and $\mathrm{E}(Y_t)$ for $t = 1$.

b) Prove that $X$ and $Y$ are Ito's processes and compute $\langle X, Y\rangle_t$.

**8.4** (p. 536)

a) Let $\sigma > 0$. Prove that

$$\int_0^t \mathrm{e}^{\sigma B_s - \frac{1}{2}\sigma^2 s}\, dB_s = \frac{1}{\sigma}\big(\mathrm{e}^{\sigma B_t - \frac{1}{2}\sigma^2 t} - 1\big)\,. \tag{8.52}$$

b)

$$\int_0^t \mathrm{e}^{\int_0^s B_u\, dB_u - \frac{1}{2}\int_0^s B_u^2\, du} B_s\, dB_s = ?$$

**8.5** (p. 536) Let $B$ be a Brownian motion with respect to the filtration $(\mathscr{F}_t)_t$.

a) Show that $X_t = B_t^3 - 3tB_t$ is an $(\mathscr{F}_t)_t$-martingale.

b) Prove that for every $n = 1, 2, \dots$ there exist numbers $c_{n,m}$, $m \le [n/2]$ such that the polynomial

$$P_n(x,t) = \sum_{m=0}^{[n/2]} c_{n,m} x^{n-2m} t^m \tag{8.53}$$

is such that $X_t = P_n(B_t, t)$ is an $(\mathscr{F}_t)_t$-martingale. Compute $P_n$ for $n = 1, 2, 3, 4$.

c) Let $a, b > 0$ and let $\tau$ be the exit time of $B$ from $]-a, b[$. In Example 5.3 and Exercise 5.10 we have computed the law of $B_\tau$ and $\mathrm{E}(\tau)$. Are $\tau$ and $B_\tau$ independent?

**8.6** (p. 537) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and $\lambda \in \mathbb{R}$.

a) Prove that

$$M_t = \mathrm{e}^{\lambda t} B_t - \lambda \int_0^t \mathrm{e}^{\lambda u} B_u\, du$$

is a square integrable martingale and compute its associated increasing process.

b) Prove that if $\lambda < 0$ then $M$ is uniformly integrable and determine the distribution of its limit as $t \to +\infty$.

c1) Prove that

$$Z_t = \mathrm{e}^{M_t - \frac{1}{4\lambda}(\mathrm{e}^{2\lambda t} - 1)}$$

is a martingale.

c2) Let $Z_\infty$ be the limit of the positive martingale $Z$ as $t \to +\infty$ and assume $\lambda < 0$. What is the law of $Z_\infty$?

c3) Assume $\lambda > 0$. Compute $\mathrm{E}[Z_t^p]$ for $p < 1$. What is the law of $Z_\infty$?

**8.7** (p. 538) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and let

$$Y_t = tB_t - \int_0^t B_u \, du \; .$$

In Exercise 5.13 and in Example 8.6 we proved that $Y$ is an $(\mathscr{F}_t)_t$-martingale.

Prove that

$$Z_t = \mathrm{e}^{Y_t - \frac{1}{6}t^3}$$

is an $(\mathscr{F}_t)_t$-martingale.

**8.8** (p. 539) Let $B$ be a Brownian motion and

$$X_t = \int_0^t \frac{dB_s}{\sqrt{2 + \sin s}} \cdot$$

What is the value of

$$\overline{\lim_{t \to +\infty}} \frac{X_t}{(2t \log \log t)^{1/2}} \; ?$$

For $-\pi < t < \pi$ a primitive of $(2 + \sin t)^{-1}$ is $t \mapsto 2 \cdot 3^{-1/2} \arctan(2 \cdot 3^{-1/2} \tan(\frac{t}{2}) + 3^{-1/2})$ hence

$$\int_{-\pi}^{\pi} \frac{ds}{2 + \sin s} = \frac{2\pi}{\sqrt{3}} \cdot$$

**8.9** (p. 540) Let, for $\lambda > 0$,

$$X_t = \int_0^t \mathrm{e}^{-\lambda s} B_s \, ds \; .$$

Does the limit

$$\lim_{t \to +\infty} X_t \tag{8.54}$$

exist? In what sense? Determine the limit and compute its distribution.

Integrate by parts...

**8.10** (p. 540) Let $B$ be a real Brownian motion and, for $\varepsilon > 0$ and $t \le T$,

$$X_t^\varepsilon = \sqrt{2} \int_0^t \sin \frac{s}{\varepsilon} \, dB_s \,.$$

a) Prove that, for every $0 < t \le T$, $X_t^\varepsilon$ converges in law as $\varepsilon \to 0$ to a limiting distribution to be determined.
b) Prove that, as $\varepsilon \to 0$, the law of the continuous process $X^\varepsilon$ for $t \in [0, T]$ converges to the Wiener measure.

**8.11** (p. 541) Let $B$ be a Brownian motion and

$$X_t = \int_0^t \frac{dB_s}{\sqrt{1+s}} \,\cdot$$

Compute $\mathrm{P}\big(\sup_{0\le t\le 3} X_t \ge 1\big)$.

**8.12** (p. 541) Let, for $\alpha > 0$,

$$X_t = \sqrt{\alpha + 1} \int_0^t u^{\alpha/2} \, dB_u \,.$$

a) Compute $\mathrm{P}(\sup_{0\le s\le 2} X_s \ge 1)$.
b) Let $\tau = \inf\{t > 0; X_t \ge 1\}$. Compute the density of $\tau$. For which values of $\alpha$ does $\tau$ have finite expectation?

**8.13** (p. 542) As the Brownian paths are continuous, and therefore locally bounded, we can consider their derivative in the sense of distributions, i.e. for every $\phi \in C_K^\infty(\mathbb{R}^+)$ and for every $\omega \in \Omega$ we can consider the distribution $B'(\omega)$ defined as

$$\langle \phi, B'(\omega) \rangle = -\int_0^{+\infty} \phi_t' B_t(\omega) \, dt \,.$$

Check that

$$\langle \phi, B' \rangle = \int_0^{+\infty} \phi_t \, dB_t \quad \text{a.s.}$$

**8.14** (p. 542) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and let, for $0 \le t < 1$,

$$Z_t = \frac{1}{\sqrt{1-t}} \exp\Big(-\frac{B_t^2}{2(1-t)}\Big) \,.$$

a) Prove that $Z$ is a continuous $(\mathscr{F}_t)_t$-martingale.
b) Compute, for $p > 0$, $\mathrm{E}[Z_t^p]$. Is $Z$ square integrable? Is $Z$ bounded in $L^2$?
c) What is the value of $\lim_{t\to 1-} Z_t$?
d) Find a process $(X_t)_t \in M^2_{loc}$ such that

$$Z_t = \mathrm{e}^{\int_0^t X_s\, dB_s - \frac{1}{2}\int_0^t X_s^2\, ds} .$$

**8.15** (p. 545) Let $B = (B_1, B_2)$ be a two-dimensional Brownian motion. Which one of the following is a Brownian motion?

a) $W_1(t) = \displaystyle\int_0^t \sin s\, dB_1(s) + \int_0^t \cos s\, dB_1(s)$.
b) $W_2(t) = \displaystyle\int_0^t \sin s\, dB_1(s) + \int_0^t \cos s\, dB_2(s)$.
c) $W_3(t) = \displaystyle\int_0^t \sin B_2(s)\, dB_1(s) + \int_0^t \cos B_2(s)\, dB_2(s)$.

**8.16** (p. 545) Let $(B_1, B_2, B_3)$ be a three-dimensional Brownian motion and

$$\begin{aligned} X_t &= \int_0^t \sin(B_3(s))\, dB_1(s) + \int_0^t \cos(B_3(s))\, dB_2(s) \\ Y_t &= \int_0^t \cos(B_3(s))\, dB_1(s) + \int_0^t \sin(B_3(s))\, dB_2(s) . \end{aligned} \tag{8.55}$$

a) Prove that $(X_t)_t$ and $(Y_t)_t$ are Brownian motions.
b1) Prove that, for every $s, t$, $X_s$ and $Y_t$ are uncorrelated.
b2) Is $(X_t, Y_t)_t$ a two-dimensional Brownian motion?
c) Assume instead that

$$Y_t = -\int_0^t \cos(B_3(s))\, dB_1(s) + \int_0^t \sin(B_3(s))\, dB_2(s) .$$

Is $(X_t, Y_t)_t$ now a two-dimensional Brownian motion?

**8.17** (p. 546) Let $B = (\Omega, \mathscr{F}, (\mathscr{G}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion with respect to its natural filtration $(\mathscr{G}_t)_t$ and let $\widetilde{\mathscr{G}}_t = \mathscr{G}_t \vee \sigma(B_1)$.

a) Let $0 < s < t < 1$ be fixed. Determine a square integrable function $\Phi$ and a number $\alpha$ (possibly depending on $s, t$) such that the r.v.

$$B_t - \int_0^s \Phi(u)\, dB_u - \alpha B_1$$

is orthogonal to $B_1$ and to $B_v$ for every $v \le s$.
b) Compute $\mathrm{E}[B_t | \widetilde{\mathscr{G}}_s]$. Is $B$ also a Brownian motion with respect to the filtration $(\widetilde{\mathscr{G}}_t)_t$? Is $B$ a martingale with respect to $(\widetilde{\mathscr{G}}_t)_t$?

c) Let, for $t \le 1$,

$$\widetilde{B}_t = B_t - \int_0^t \frac{B_1 - B_u}{1-u}\, du\ .$$

c1) Prove that, for $0 < s < t < 1$, $\widetilde{B}_t - \widetilde{B}_s$ is independent of $\widetilde{\mathscr{G}}_s$.
c2) Prove that, for $0 < t < 1$, $(\widetilde{B}_t)_t$ is a $(\widetilde{\mathscr{G}}_t)_t$-Brownian motion.
c3) Prove that $(B_t)_t$ is a Ito process with respect to $(\widetilde{\mathscr{G}}_t)_t$.

- This exercise shows that when switching to a larger filtration, in general, properties such as being a Brownian motion or a martingale are lost, whereas to be a Ito process is more stable.

**8.18** (p. 548) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion and $X$ a centered square integrable $\mathscr{F}_T$-measurable r.v.

a) Let us assume that there exists a process $(Y_t)_t$ in $M^2([0,T])$ such that

$$X = \int_0^T Y_s\, dB_s\ . \tag{8.56}$$

Show that such a process is unique.
b1) Let $X = B_T^3$. What is the value of $X_s = \mathrm{E}[X|\mathscr{F}_s]$ for $0 \le s \le T$? What is the stochastic differential of $(X_s)_{0\le s\le T}$?
b2) Determine a process $(Y_t)_t$ such that (8.56) holds in this case.
c) Determine a process $(Y_t)_t$ such that (8.56) holds for $X = \mathrm{e}^{\sigma B_T} - \mathrm{e}^{\frac{\sigma^2}{2}T}$, for $\sigma \ne 0$.

**8.19** (p. 549) Let $B$ be a Brownian motion and

$$Z = \int_0^T B_s\, ds\ .$$

a1) Compute $M_t = \mathrm{E}[Z|\mathscr{F}_t]$.
a2) Compute the stochastic differential of $(M_t)_t$ and determine a process $X \in M^2$ such that

$$Z = \int_0^T X_s\, dB_s\ .$$

b) Determine a process $X \in M^2$ such that

$$\int_0^T B_s^2\, ds = \frac{T^2}{2} + \int_0^T X_s\, dB_s\ .$$

**8.20** (p. 550) In Exercise 5.24 we proved that, if $B$ is an $m$-dimensional Brownian motion, then $M_t = B_i(t)B_j(t)$, $1 \le i, j \le m$, $i \ne j$, is a square integrable martingale and in Example 8.7 we have shown that

$$\langle M \rangle_t = \int_0^t (B_i^2(s) + B_j^2(s))\, ds \,.$$

Let $\theta > 0$. Determine an increasing process $(A_t)_t$ such that

$$Z_t = \mathrm{e}^{\theta B_i(t) B_j(t) - A_t}$$

is a local martingale.

**8.21** (p. 551)

a) Let $Z$ be an $N(0, 1)$-distributed r.v. Show that

$$\mathrm{E}[Z^2 \mathrm{e}^{\alpha Z^2}] = \begin{cases} (1-2\alpha)^{-3/2} & \text{if } \alpha < \frac{1}{2} \\ +\infty & \text{if } \alpha \ge \frac{1}{2} \,. \end{cases}$$

b) Let $(B_t)_t$ be a Brownian motion and let

$$H_t = \frac{B_t}{(1-t)^{3/2}} \exp\Big( - \frac{B_t^2}{2(1-t)} \Big) \,.$$

Prove that $\lim_{t \to 1-} H_t = 0$ a.s. Prove that $H \in M^2_{loc}([0, 1])$ but $H \notin M^2([0, 1])$.

c) Let, for $t < 1$,

$$X_t = \frac{1}{\sqrt{1-t}} \exp\Big( - \frac{B_t^2}{2(1-t)} \Big) \,.$$

Prove that, for $t < 1$,

$$X_t = 1 - \int_0^t H_s \, dB_s \,. \tag{8.57}$$

Prove that $\lim_{t \to 1-} X_t = 1$ a.s.

d) Prove that

$$\int_0^1 H_s \, dB_s = 1$$

so that $H$ is an example of an integrand not belonging to $M^2$ whose stochastic integral is square integrable.

**8.22** (p. 552) (Poincaré's lemma) Let $B_1, B_2, \dots$ be a sequence of one-dimensional independent Brownian motions defined on the same probability space and let $B_t^{(n)} = (B_1(t), \dots, B_n(t))$; $B^{(n)}$ is therefore an $n$-dimensional Brownian motion. Let $\tau_n$ be the exit time of $B^{(n)}$ from the ball of radius $\sqrt{n}$.

a) Show that $M_n(t) = |B_t^{(n)}|^2 - nt$ is a martingale. Deduce that $\mathrm{E}(\tau_n) = 1$ for every $n$.
b) If $R_n(t) = \frac{1}{n}|B_t^{(n)}|^2$, show that there exists a Brownian motion $W$ such that

$$dR_n(t) = dt + \frac{2}{\sqrt{n}}\sqrt{R_n(t)}\,dW_t\,. \tag{8.58}$$

c) Deduce that

$$\mathrm{E}\Big[\sup_{0\le t\le \tau_n} |R_n(t) - t|^2\Big] \le \frac{16}{n}$$

and therefore $\tau_n \to 1$ in probability as $n \to \infty$.
d) (Gallardo 1981). Deduce the following classical lemma of H. Poincaré. Let $\sigma_n$ denote the normalized $(n-1)$-dimensional Lebesgue measure of the spherical surface of radius $\sqrt{n}$ of $\mathbb{R}^n$. Let $d$ be a fixed positive integer and $\pi_{n,d}$ the projection $\mathbb{R}^n \to \mathbb{R}^d$.

Then as $n \to \infty$ the image of $\sigma_n$ through $\pi_{n,d}$ converges to the $N(0, I)$ law of $\mathbb{R}^d$.

c) $Z_t = R_n(t) - t$, suitably stopped, is a martingale.
d) $\sigma_n$ is the law of $B_{\tau_n}^{(n)}$; $\pi_{n,d}(\sigma_n)$ is therefore the law of $(B_1(\tau_n), \dots, B_d(\tau_n))$.

**8.23** (p. 553) In this exercise we investigate the properties of recurrence/transience of Brownian motion in dimension 2, hence completing the analysis of Sect. 8.5. It turns out that in dimension 2 Brownian motion visits every open set with probability 1 (which is different from the case of dimension $m \ge 3$), and is therefore recurrent, but in contrast to dimension 1, every point $x \ne 0$ is visited with probability 0.

Let $B$ be a two-dimensional Brownian motion and $x \in \mathbb{R}^2, x \ne 0$. Let $\tau_n$ be the entrance time of $B$ in the ball of radius $\frac{1}{n}$ centered at $x$, with $\frac{1}{n} < |x|$, so that this ball does not contain the origin. Let $X_t = -\log|B_t - x|$, when defined.

a) Prove that $(X_{t\wedge\tau_n})_t$ is a martingale.
b) Let, for $\frac{1}{n} < M$, $\tau_{n,M}$ be the exit time of $B$ from the annulus $\{z; \frac{1}{n} < |z-x| \le M\}$ with $M > |x|$ so that this annulus contains the origin (Fig. 8.1). Prove that $\tau_{n,M} < +\infty$ a.s. and compute $\mathrm{P}(|B_{\tau_{n,M}} - x| = \frac{1}{n})$.
c) Deduce that $\mathrm{P}(\tau_n < +\infty) = 1$ for every $n$ and therefore that $B$ is recurrent.
d) Let $x \in \mathbb{R}^2$, $x \ne 0$, and let us prove that $B$ does not visit $x$ with probability 1. Let $k > 0$ be such that $\frac{1}{k^k} < |x| < k$. Let

$$\zeta_k = \tau_{k^k} = \inf\{t; |B_t - x| = \tfrac{1}{k^k}\}$$

$$\sigma_k = \inf\{t; |B_t - x| = k\}\,.$$

d1) Deduce from the computation of b) the value of $P(\zeta_k < \sigma_k)$. Compute the limit of this quantity as $k \to \infty$.
d2) Let $\tau = \inf\{t; B_t = x\}$. Note that, for every $k$, $P(\tau < \sigma_k) \le P(\zeta_k < \sigma_k)$ and deduce that $P(\tau < +\infty) = 0$.

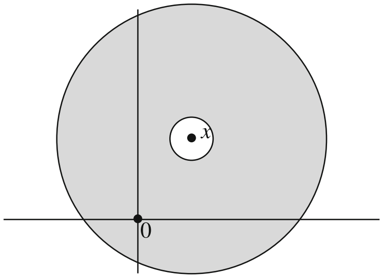


**Fig. 8.1** $\tau_{n,M}$ is the exit time of the Brownian motion from the shaded annulus. The large circle has radius $M$, the small one $\frac{1}{n}$

**8.24** (p. 554) (Bessel processes) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, P)$ be an $m$-dimensional standard Brownian motion with $m > 1$ and let $x \in \mathbb{R}^m$ with $|x| > 0$. Then, if $X_t = |B_t + x|$, $X$ is a Markov process by Exercise 6.11 and $X_t > 0$ a.s. for every $t \ge 0$ a.s. as seen in Sect. 8.5 and Exercise 8.23.

a) Show that $X_t = |B_t + x|$ is an Ito process and determine its stochastic differential.
b) Show that if $f \in C^2_K(]0, +\infty[)$ and $\xi = |x|$, then

$$\frac{1}{t}\big(E[f(X_t)] - f(\xi)\big) \underset{t\to+\infty}{\to} Lf(\xi) ,$$

where $L$ is a differential operator to be determined.

# Chapter 9
# Stochastic Differential Equations

In this chapter we introduce the notion of a Stochastic Differential Equation. In Sects. 9.4, 9.5, 9.6 we investigate existence and uniqueness. In Sect. 9.8 we obtain some $L^p$ estimates that will allow us to specify the regularity of the paths and the dependence from the initial conditions. In the last sections we shall see that the solution of a stochastic differential equation is a Markov process and even a diffusion associated to a differential operator that we shall specify.

## 9.1 Definitions

Let $b(x,t) = (b_i(x,t))_{1\le i\le m}$ and $\sigma(x,t) = (\sigma_{ij}(x,t))_{\substack{1\le i\le m\\ 1\le j\le d}}$ be measurable functions defined on $\mathbb{R}^m \times [0,T]$ and $\mathbb{R}^m$- and $M(m,d)$-valued respectively (recall that $M(m,d) = m\times d$ real matrices).

**Definition 9.1** The process $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,T]}, (\xi_t)_{t\in[u,T]}, (B_t)_t, \mathrm{P})$ is said to be a solution of the Stochastic Differential Equation (SDE)

$$\begin{aligned} d\xi_t &= b(\xi_t,t)\,dt + \sigma(\xi_t,t)\,dB_t \\ \xi_u &= x \qquad\qquad\qquad\qquad x\in\mathbb{R}^m \end{aligned} \tag{9.1}$$

if

a) $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ is a $d$-dimensional standard Brownian motion, and
b) for every $t\in[u,T]$ we have

$$\xi_t = x + \int_u^t b(\xi_s,s)\,ds + \int_u^t \sigma(\xi_s,s)\,dB_s\,.$$

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_9

Of course in Definition 9.1 we require implicitly that $s \mapsto b(\xi_s, s)$ and $s \mapsto \sigma(\xi_s, s)$ are processes in $M^1_{loc}([u,T])$ and $M^2_{loc}([u,T])$, respectively.

Note that the Brownian motion with respect to which the stochastic integrals are taken and the probability space on which it is defined *are not given a priori*.

Note also that the solution of a SDE as above is necessarily a continuous process. In case someone wanted to model some real life phenomenon with such a SDE it is important to realize this point: (9.1) is not fit to model quantities that make jumps. Discontinuous behaviors must be modeled using different SDEs (with the Brownian motion replaced by some more suitable stochastic process).

Some terminology: $\sigma$ is the *diffusion coefficient*, $b$ is the *drift*.

**Definition 9.2** We say that (9.1) has *strong solutions* if for every standard Brownian motion $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ there exists a process $\xi$ that satisfies Eq. (9.1).

We shall speak of *weak solutions*, meaning those in the sense of Definition 9.1.

If $\xi$ is a solution, strong or weak, we can consider the law of the process (see Sect. 3.2): recall that the map $\psi_\xi : \Omega \to \mathscr{C}([0,T], \mathbb{R}^m)$ defined as

$$\omega \mapsto (t \mapsto \xi_t(\omega))$$

is measurable (Proposition 3.3) and the *law of the process* $\xi$ is the probability on $\mathscr{C}([0,T], \mathbb{R}^m)$ that is the image of P through the map $\psi_\xi$.

**Definition 9.3** We say that for the SDE (9.1) there is *uniqueness in law* if, given two solutions $\xi^i = (\Omega^i, \mathscr{F}^i, (\mathscr{F}^i_t)_t, (\xi^i_t)_t, (B^i_t)_t, \mathrm{P}^i)$, $i = 1, 2$, (possibly defined on different probability spaces and/or with respect to different Brownian motions) $\xi^1$ and $\xi^2$ have the same law.

**Definition 9.4** We say that for the SDE (9.1) there is *pathwise uniqueness* if, given two solutions $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (\xi^i_t)_t, (B_t)_t, \mathrm{P}), i = 1, 2$, defined on the same probability space and with respect to the same Brownian motion, $\xi_1$ and $\xi_2$ are indistinguishable, i.e. $\mathrm{P}(\xi^1_t = \xi^2_t \text{ for every } t \in [u,T]) = 1$.

Note that, whereas the existence of strong solutions immediately implies the existence of weak solutions, it is less obvious that pathwise uniqueness implies uniqueness in law, since for the latter we must compare solutions that are defined on different probability spaces and with respect to different Brownian motions. It is, however, possible to prove that pathwise uniqueness implies uniqueness in law.

*Remark 9.1* If $\xi$ is a solution of the SDE (9.1) and $f : \mathbb{R}^m \times \mathbb{R}^+ \to \mathbb{R}$ is a $C^{2,1}$ function, then the stochastic differential of $t \mapsto f(\xi_t, t)$ is

$$df(\xi_t, t) = \frac{\partial f}{\partial t}(\xi_t, t)\, dt + \sum_i \frac{\partial f}{\partial x_i}(\xi_t, t)\, d\xi_i(t) + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i x_j}(\xi_t, t) a_{ij}(\xi_t, t)\, dt\;,$$

where $a_{ij} = \sum_\ell \sigma_{i\ell}\sigma_{j\ell}$, i.e. $a = \sigma\sigma^*$. Hence we can write

$$df(\xi_t, t) = \Big(\frac{\partial f}{\partial t}(\xi_t, t) + Lf(\xi_t, t)\Big)\, dt + \sum_{i,j} \frac{\partial f}{\partial x_i}(\xi_t, t)\sigma_{ij}(\xi_t, t)\, dB_j(t)\;,$$

where $L$ denotes the second-order differential operator

$$L = \frac{1}{2} \sum_{i,j=1}^m a_{ij}(x, t) \frac{\partial^2}{\partial x_i x_j} + \sum_{i=1}^m b_i(x, t) \frac{\partial}{\partial x_i}\, \cdot$$

In particular, if we look for a function $f$ such that $t \mapsto f(\xi_t, t)$ is a martingale, a first requirement is that

$$\frac{\partial f}{\partial t} + Lf = 0\;,$$

which ensures that $t \mapsto f(\xi_t, t)$ is a local martingale.

The operator $L$ will play an important role from now on. Note that if $d\xi_t = dB_t$, then $L = \frac{1}{2} \triangle$, $\triangle$ denoting the Laplace operator.

In this chapter we shall generally denote stochastic processes by the symbols $\xi$, $(\xi_t)_t$. The notations $X$, $(X_t)_t$, most common in the previous chapters, are now reserved to denote the canonical process. From now on $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ will denote a $d$-dimensional standard Brownian motion, fixed once and for all.

## 9.2 Examples

In the next sections we shall investigate the existence and the uniqueness of solutions of a SDE, but let us first see some particular equations for which it is possible to find an explicit solution.

*Example 9.1* Let us consider the equation, in dimension 1,

$$\begin{aligned} d\xi_t &= -\lambda\xi_t\,dt + \sigma\,dB_t \\ \xi_0 &= x \end{aligned} \tag{9.2}$$

where $\lambda, \sigma \in \mathbb{R}$; i.e. we assume that the drift is linear in $\xi$ and the diffusion coefficient constant. To solve (9.2) we can use a method which is similar to the variation of constants for ordinary differential equations: here the stochastic part of the equation, $\sigma\,dB_t$, plays the role of the constant term. The "homogeneous" equation would be

$$\begin{aligned} d\xi_t &= -\lambda\xi_t\,dt \\ \xi_0 &= x \end{aligned}$$

whose solution is $\xi_t = \mathrm{e}^{-\lambda t}x$. Let us look for a "particular solution" of (9.2) of the form $\xi_t = \mathrm{e}^{-\lambda t}Z_t$. As, by Ito's formula,

$$d(\mathrm{e}^{-\lambda t}Z_t) = -\lambda\mathrm{e}^{-\lambda t}Z_t\,dt + \mathrm{e}^{-\lambda t}\,dZ_t$$

$\xi_t = \mathrm{e}^{-\lambda t}Z_t$ is a solution if $\mathrm{e}^{-\lambda t}dZ_t = \sigma\,dB_t$, i.e.

$$Z_t = \int_0^t \mathrm{e}^{\lambda s}\sigma\,dB_s$$

and, in conclusion,

$$\xi_t = \mathrm{e}^{-\lambda t}x + \mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{\lambda s}\sigma\,dB_s\,. \tag{9.3}$$

This is the *Ornstein–Uhlenbeck process* (in the literature this name is sometimes reserved for the case $\lambda > 0$). It is a process of particular interest, being the natural model in many applications, and we shall see in the exercises some of its important properties (see Exercises 9.1, 9.2, 9.24 and 9.25). Now let us observe that the computation above can be repeated if we consider an $m$-dimensional analog of (9.2), i.e. the equation

$$\begin{aligned} d\xi_t &= -\Lambda\xi_t\,dt + \sigma\,dB_t \\ \xi_0 &= x \end{aligned} \tag{9.4}$$

where $\Lambda$ and $\sigma$ are $m \times m$ and $m \times d$ matrices, respectively. By a step by step repetition of the computation above, we have that a solution is

(continued)

*Example 9.1* (continued)

$$\xi_t = e^{-At}x + e^{-At}\int_0^t e^{As}\sigma\, dB_s\,, \tag{9.5}$$

$e^{At}$ denoting the exponential of the matrix $At$. This is a Gaussian process by Proposition 7.1, as the stochastic integral above has a deterministic integrand.

*Example 9.2* Let us now consider instead the equation in dimension 1

$$\begin{aligned} d\xi_t &= b\xi_t\, dt + \sigma\xi_t\, dB_t \\ \xi_0 &= x\,. \end{aligned} \tag{9.6}$$

Now both drift and diffusion coefficient are linear functions of $\xi$. Dividing both sides by $\xi_t$ we have

$$\frac{d\xi_t}{\xi_t} = b\, dt + \sigma\, dB_t\,. \tag{9.7}$$

The term $\frac{d\xi_t}{\xi_t}$ is suggestive of the stochastic differential of $\log \xi_t$. It is not quite this way, as we know that stochastic differentials behave differently from the usual ones. Anyway, if we compute the stochastic differential of $\log \xi_t$, assuming that $\xi$ is a solution of (9.6), Ito's formula gives

$$d(\log \xi_t) = \frac{d\xi_t}{\xi_t} - \frac{1}{2\xi_t^2}\,\sigma^2\xi_t^2\, dt = \Big(b - \frac{\sigma^2}{2}\Big)\, dt + \sigma\, dB_t\,.$$

Therefore $\log \xi_t = \log x + \big(b - \frac{\sigma^2}{2}\big)t + \sigma B_t$ and

$$\xi_t = x e^{(b-\frac{\sigma^2}{2})t+\sigma B_t}\,. \tag{9.8}$$

In fact this derivation of the solution is not correct: we cannot apply Ito's formula to the logarithm function, which is not twice differentiable on $\mathbb{R}$ (it is not even defined on the whole of $\mathbb{R}$). But once the solution (9.8) is derived, it is easy, always by Ito's formula, to check that it is actually a solution of Eq. (9.6). Note that, if $x > 0$, then the solution remains positive for every $t \geq 0$.

This process is *geometric Brownian motion* and it is one of the processes to be taken into account as a model for the evolution of quantities that must

(continued)

*Example 9.2* (continued)
always stay positive (it is not the only one enjoying this property). Because of this, it is also used to describe the evolution of prices in financial markets.

Note that, for every $t$, $\xi_t$ appears to be the exponential of a Gaussian r.v. It has therefore a *lognormal* distribution (see Exercise 1.11).

This section provides examples of SDE's for which it is possible to obtain an explicit solution. This is not a common situation. However developing the arguments of these two examples it is possible to find a rather explicit solution of a SDE when the drift and diffusion coefficient are both linear-affine functions. Complete details are given, in the one-dimensional case, in Exercise 9.11. Other examples of SDE's for which an explicit solution can be obtained are developed in Exercises 9.6 and 9.13.

Note that in the two previous examples we have found a solution but we still know nothing about uniqueness.

## 9.3 An a priori estimate

In this section we prove some properties of the solution of an SDE before looking into the question of existence and uniqueness.

**Assumption (A)** We say that the coefficients $b$ and $\sigma$ satisfy Assumption (A) if they are measurable in $(x,t)$ and if there exist constants $L > 0, M > 0$ such that for every $x, y \in \mathbb{R}^m$, $t \in [0,T]$,

$$|b(x,t)| \le M(1+|x|) \qquad |\sigma(x,t)| \le M(1+|x|) \tag{9.9}$$

$$|b(x,t)-b(y,t)| \le L|x-y| \qquad |\sigma(x,t)-\sigma(y,t)| \le L|x-y| \,. \tag{9.10}$$

Note that (9.10) requires the coefficients $b$ and $\sigma$ to be Lipschitz continuous in $x$ uniformly in $t$, whereas (9.9) requires that they have a sublinear growth at infinity. Note also that if $b$ and $\sigma$ do not depend on $t$ then (9.10) implies (9.9). However, it is useful to keep these two conditions separate in order to clarify their role in what follows. In the remainder we shall often make use of the inequality, for $x_1, \dots, x_m \in \mathbb{R}^d$,

$$|x_1 + \cdots + x_m|^p \le m^{p-1}(|x_1|^p + \cdots + |x_m|^p) \,. \tag{9.11}$$

A key tool in our investigation of SDE's is the following classical inequality.

**Lemma 9.1 (Gronwall's inequality)** Let $w, v$ be non-negative real functions on the interval $[a, b]$ with $w$ integrable and $v$ bounded measurable; let $c \geq 0$ and let us assume that

$$v(t) \leq c + \int_a^t w(s)v(s)\,ds\,.$$

Then

$$v(t) \leq c\,\mathrm{e}^{\int_a^t w(s)\,ds}\,.$$

Note in particular in the previous statement that if $c = 0$ then necessarily $v \equiv 0$.

Gronwall's inequality is used with ordinary differential equations in order to prove bounds and uniqueness of the solutions. For SDE's its use is similar.

In particular, Gronwall's inequality allows to prove the following Theorem 9.1: it states that if there exists a solution (weak or strong) of the SDE (9.1) with an initial value $\eta$ that is in $L^p$ with $p \geq 2$ and independent of the Brownian motion, then under the condition of sublinear growth (9.10) the solution is necessarily a process in $M^p$. Note that sublinear growth of the coefficients is required, but no assumption concerning Lipschitz continuity is made. And still we do not know whether a solution actually exists.

**Theorem 9.1** Let $\xi$ be a solution of

$$\xi_t = \eta + \int_u^t b(\xi_r, r)\,dr + \int_u^t \sigma(\xi_r, r)\,dB_r$$

where the coefficients $b, \sigma$ are measurable in $x, t$ and satisfy (9.9) (sublinear growth) and $\eta$ is an $\mathscr{F}_u$-measurable r.v. of $L^p$. Then for every $p \geq 2$

$$\mathrm{E}\Big(\sup_{u\leq s\leq T} |\xi_s|^p\Big) \leq c(p, T, M)(1 + \mathrm{E}(|\eta|^p)) \tag{9.12}$$

$$\mathrm{E}\Big(\sup_{u\leq s\leq t} |\xi_s - \eta|^p\Big) \leq c(p, T, M)(t-u)^{p/2}(1 + \mathrm{E}(|\eta|^p))\,. \tag{9.13}$$

*Proof* The idea is to find some inequalities in order to show that the function $v(t) = \mathrm{E}[\sup_{u\leq s\leq t} |\xi_s|^p]$ satisfies an inequality of the kind

$$v(t) \leq c_1(1 + \mathrm{E}[|\eta|^p]) + c_2 \int_u^t v(s)\,ds$$

and then to apply Gronwall's inequality. There is, however, a difficulty, as in order to apply Gronwall's inequality we must know beforehand that such a function $v$ is bounded. Otherwise it might be $v \equiv +\infty$. In order to circumvent this difficulty we shall be obliged stop the process when it takes large values as described below.

Let, for $R > 0$, $\xi_R(t) = \xi_{t\wedge\tau_R}$ where $\tau_R = \inf\{t; u \le t \le T, |\xi_t| \ge R\}$ denotes the exit time of $\xi$ from the open ball of radius $R$, with the understanding $\tau_R = T$ if $|\xi_t| < R$ for every $t \in [u, T]$. Then

$$\begin{aligned}\xi_R(t) &= \eta + \int_u^{t\wedge\tau_R} b(\xi_r, r)\,dr + \int_u^{t\wedge\tau_R} \sigma(\xi_r, r)\,dB_r\\ &= \eta + \int_u^t b(\xi_r, r)1_{\{r<\tau_R\}}\,dr + \int_u^t \sigma(\xi_r, r)1_{\{r<\tau_R\}}\,dB_r\\ &= \eta + \int_u^t b(\xi_R(r), r)1_{\{r<\tau_R\}}\,dr + \int_u^t \sigma(\xi_R(r), r)1_{\{r<\tau_R\}}\,dB_r\,.\end{aligned}\tag{9.14}$$

Taking the modulus, the $p^{th}$ power and then the expectation we find, using (9.11),

$$\begin{aligned}&\mathrm{E}\Big[\sup_{u\le s\le t}|\xi_R(s)|^p\Big]\\ &\le 3^{p-1}\mathrm{E}[|\eta|^p] + 3^{p-1}\mathrm{E}\Big[\sup_{u\le s\le t}\Big|\int_u^s b(\xi_R(r), r)1_{\{r<\tau_R\}}\,dr\Big|^p\Big]\\ &\quad+3^{p-1}\mathrm{E}\Big(\sup_{u\le s\le t}\Big|\int_u^s \sigma(\xi_R(r), r)1_{\{r<\tau_R\}}\,dB_r\Big|^p\Big)\,.\end{aligned}\tag{9.15}$$

Now, by Hölder's inequality and (9.9),

$$\begin{aligned}\mathrm{E}\Big[\sup_{u\le s\le t}\Big|\int_u^s b(\xi_R(r), r)1_{\{r\le\tau_R\}}\,dr\Big|^p\Big] &\le T^{p-1}\mathrm{E}\Big[\int_u^t |b(\xi_R(r), r)|^p 1_{\{r<\tau_R\}}\,dr\Big]\\ &\le T^{p-1}M^p E\Big[\int_u^t (1+|\xi_R(r)|)^p\,dr\Big]\,.\end{aligned}$$

By the $L^p$ inequalities for stochastic integrals (Proposition 8.4) we have

$$\begin{aligned}\mathrm{E}\Big[\sup_{u\le s\le t}\Big|\int_u^s \sigma(\xi_R(r), r)1_{\{r<\tau_R\}}\,dB_r\Big|^p\Big] &\le c_p|t-u|^{\frac{p-2}{2}}\mathrm{E}\Big(\int_u^t |\sigma(\xi_R(r), r)|^p\,dr\Big)\\ &\le c_pT^{\frac{p-2}{2}}M^p\mathrm{E}\Big(\int_u^t (1+|\xi_R(r)|)^p\,dr\Big)\,.\end{aligned}$$

Hence, plugging these two inequalities into (9.15), we find

$$\begin{aligned}&\mathrm{E}\Big[\sup_{u\le s\le t}|\xi_R(s)|^p\Big]\\ &\le 3^{p-1}\mathrm{E}[|\eta|^p] + 3^{p-1}M^p\big(T^{p-1}+c_pT^{\frac{p-2}{2}}\big)\mathrm{E}\Big(\int_u^t |1+\xi_R(r)|^p\,dr\Big)\end{aligned}$$

$$
\begin{aligned}
&\le 3^{p-1}\mathrm{E}[|\eta|^p] + 3^{p-1}M^p\big(T^{p-1} + c_pT^{\frac{p-2}{2}}\big)2^{p-1}\mathrm{E}\Big(T^p + \int_u^t |\xi_R(r)|^p\,dr\Big)\\
&\qquad \le c_1(p,T,M)\big(1+\mathrm{E}|\eta|^p\big) + c_2(p,T,M)\int_u^t \mathrm{E}\big[|\xi_R(r)|^p\big]\,dr\,.
\end{aligned}
$$

Let now $v(t) = \mathrm{E}\big[\sup_{u\le s\le t}|\xi_R(s)|^p\big]$: from the previous inequality we have

$$
v(t) \le c_1(p,T,M)(1+\mathrm{E}|\eta|^p) + c_2(p,T,M)\int_u^t v(r)\,dr\,.
$$

Now $|\xi_R(t)| = |\eta|$ if $|\eta| \ge R$ and $|\xi_R(t)| \le R$ otherwise. Hence $|\xi_R(t)| \le R \vee |\eta|$ and $v(t) \le \mathrm{E}[R^p \vee |\eta|^p] < +\infty$. $v$ is therefore bounded and thanks to Gronwall's inequality

$$
\begin{aligned}
v(T) = \mathrm{E}\Big[\sup_{u\le s\le T}|\xi_R(s)|^p\Big] &\le c_1(p,T,M)\big(1+\mathrm{E}[|\eta|^p]\big)\mathrm{e}^{Tc_2(p,T,M)}\\
&= c(p,T,M)\big(1+\mathrm{E}[|\eta|^p]\big)\,.
\end{aligned}
$$

Note that the right-hand side above does not depend on $R$. Let us now send $R \to \infty$. The first thing to observe is that $\tau_R \to_{R\to\infty} = T$: as $\xi$ is continuous we have $\sup_{u\le t\le\tau_R}|\xi_t|^p = R^p$ on $\{\tau_R < T\}$ and therefore

$$
\mathrm{E}\Big[\sup_{u\le t\le\tau_R}|\xi_t|^p\Big] \ge R^p\,\mathrm{P}(\tau_R < T)
$$

so that

$$
\mathrm{P}(\tau_R < T) \le \frac{1}{R^p}\,\mathrm{E}\Big[\sup_{u\le t\le\tau_R}|\xi_t|^p\Big] \le \frac{k(T,M)(1+\mathrm{E}(|\eta|^p))}{R^p}
$$

and therefore $\mathrm{P}(\tau_R < T) \to 0$ as $R \to +\infty$. As $R \to \tau_R$ is increasing, $\lim_{R\to+\infty}\tau_R = T$ a.s. and

$$
\sup_{u\le s\le T}|\xi_R(s)|^p \quad \underset{R\to+\infty}{\to} \quad \sup_{u\le s\le T}|\xi_s|^p \qquad \text{a.s.}
$$

and by Fatou's lemma (or Beppo Levi's theorem) we have proved (9.12). As for (9.13) we have

$$
\begin{aligned}
&\sup_{u\le s\le t}|\xi_s - \eta|^p\\
&\le 2^{p-1}\sup_{u\le s\le t}\Big|\int_u^s b(\xi_r,r)\,dr\Big|^p + 2^{p-1}\sup_{u\le s\le t}\Big|\int_u^s \sigma(\xi_r,r)\,dB_r\Big|^p\,.
\end{aligned}
\tag{9.16}
$$

Now, by Hölder's inequality and (9.12),

$$\mathrm{E}\Big[\sup_{u\le s\le t}\Big|\int_u^s b(\xi_r,r)\,dr\Big|^p\Big]\le (t-u)^{p-1}\mathrm{E}\Big[\int_u^t |b(\xi_r,r)|^p\,dr\Big]$$

$$\le (t-u)^{p-1}M^p\mathrm{E}\Big[\int_u^t (1+|\xi_r|)^p\,dr\Big]\le 2^{p-1}(t-u)^{p-1}M^p\mathrm{E}\Big[\int_u^t (1+|\xi_r|^p)\,dr\Big]$$

$$\le c_1(p,T,m)(t-u)^p(1+\mathrm{E}[|\eta|^p])\ .$$

Similarly, using again Proposition 8.4,

$$\mathrm{E}\Big[\sup_{u\le s\le t}\Big(\int_u^s \sigma(\xi_r,r)\,dB_r\Big)^p\Big]\le c(t-u)^{\frac{p-2}{2}}\mathrm{E}\Big[\int_u^t |\sigma(\xi_r,r)|^p\,dr\Big]$$

$$\le c(t-u)^{\frac{p-2}{2}}M^p\mathrm{E}\Big[\int_u^t (1+|\xi_r|)^p\,dr\Big]\le c2^{p-1}(t-u)^{\frac{p-2}{2}}M^p\mathrm{E}\Big[\int_u^t (1+|\xi_r|^p)\,dr\Big]$$

$$\le c_2(p,T,m)(t-u)^{p/2}(1+\mathrm{E}[|\eta|^p])$$

and (9.13) follows from (9.16). □

*Remark 9.2* Note again that we have proved (9.12) before we knew of the existence of a solution: (9.12) is an *a priori* bound of the solutions.

As a first consequence, for $p = 2$, under assumption (9.9) (sublinearity of the coefficients) if $\eta \in L^2$ every solution $\xi$ of the SDE (9.1) is a process belonging to $M^2$. This implies also that the stochastic component

$$t\mapsto \int_0^t \sigma(\xi_s,s)\,dB_s$$

is a square integrable martingale. Actually the process $s\mapsto\sigma(\xi_s,s)$ itself belongs to $M^2([u,t])$: as $|\sigma(\xi_s,s)|\le M(1+|\xi_s|)$ we have $|\sigma(\xi_s,s)|^2\le 2M^2(1+|\xi_s|^2)$. Hence, as the stochastic integral $\int_u^t \sigma(\xi_s,s)\,dB_s$ has expectation equal to 0,

$$\mathrm{E}[\xi_t]=\mathrm{E}[\eta]+\mathrm{E}\Big[\int_u^t b(\xi_s,s)\,ds\Big]\ . \tag{9.17}$$

Therefore, intuitively, the coefficient $b$ *has the meaning of a trend*, i.e. in the average the process follows the direction of $b$. In dimension 1, in the average, the process increases in regions where $b$ is positive and decreases otherwise.

Conversely, $\sigma$ *determines zero-mean oscillations*. If the process is one-dimensional then, recalling Theorem 8.4, i.e. the fact that a stochastic integral is a time-changed Brownian motion, regions where $\sigma$ is large in absolute value will be regions where the process undergoes oscillations with high intensity

(continued)

*Remark 9.2* (continued)
and regions where $\sigma$ is small in absolute value will be regions where the oscillations have a lower intensity.

This is useful information when dealing with the problem of modeling a given phenomenon with an appropriate SDE.

*Remark 9.3* It is useful to point out a by-product of the proof of Theorem 9.1. If $\tau_R$ is the exit time from the sphere of radius $R$, then, for every $\eta \in L^p$,

$$\mathrm{P}(\tau_R < T) = \mathrm{P}\Big(\sup_{u \le t \le T} |\xi_t| \ge R\Big) \le \frac{k(T,M)(1+\mathrm{E}(|\eta|^p))}{R^p} \cdot$$

Hence $\lim_{R\to+\infty} \mathrm{P}(\tau_R < T) = 0$ and this rate of convergence is moreover uniform for $\eta$ in a bounded set of $L^p$ and does not depend on the Lipschitz constant of the coefficients.

## 9.4 Existence for Lipschitz continuous coefficients

In this section we prove existence and uniqueness of the solutions of an SDE under suitable hypotheses on the coefficients $b$ and $\sigma$.

Let us assume by now that (9.9) (sublinear growth) holds. Then, if $Y \in M^2([0,T])$, as $|b(x,s)|^2 \le M^2(1+|x|)^2 \le 2M^2(1+|x|^2)$,

$$\begin{aligned}\mathrm{E}\Big[\int_0^t |b(Y_s,s)|^2\,ds\Big] &\le 2M^2\mathrm{E}\Big[\int_0^t (1+|Y_s|^2)\,ds\Big]\\ &\le 2M^2\Big(t+\mathrm{E}\Big[\int_0^t |Y_s|^2\,ds\Big]\Big)\end{aligned} \tag{9.18}$$

whereas, with a similar argument and using Doob's maximal inequality (the second of the maximal inequalities (7.23)),

$$\begin{aligned}\mathrm{E}\Big[\sup_{0\le u\le t}\Big|\int_0^u \sigma(Y_s,s)\,dB_s\Big|^2\Big] &\le 4\mathrm{E}\Big[\int_0^t |\sigma(Y_s,s)|^2\,ds\Big]\\ &\le 8\,M^2\Big(t+\mathrm{E}\Big[\int_0^t |Y_s|^2\,ds\Big]\Big)\,.\end{aligned} \tag{9.19}$$

We are now able to prove that for an SDE under Assumption (A) there exist strong solutions and that pathwise uniqueness holds.

**Theorem 9.2** Let $u \ge 0$ and let $\eta$ be an $\mathbb{R}^m$-valued r.v., $\mathscr{F}_u$-measurable and square integrable. Then under Assumption (A) there exists a $\xi \in M^2([u, T])$ such that

$$\xi_t = \eta + \int_u^t b(\xi_s, s)\, ds + \int_u^t \sigma(\xi_s, s)\, dB_s\ . \tag{9.20}$$

Moreover, we have pathwise uniqueness, i.e. if $\xi'$ is another solution of (9.20) then

$$\mathrm{P}(\xi_t = \xi'_t \text{ for every } t \in [u, T]) = 1\ . \tag{9.21}$$

The proof is very similar to that of comparable theorems for ordinary equations (the method of successive approximations). Before going into it, let us point out that to ask for $\eta$ to be $\mathscr{F}_u$-measurable is equivalent to requiring that $\eta$ is independent of $\sigma(B_{t+u} - B_u, t \ge 0)$, i.e., intuitively, the initial position is assumed to be independent of the subsequent random evolution.

Actually, if $\eta$ is $\mathscr{F}_u$-measurable, then it is necessarily independent of $\sigma(B_{t+u} - B_u, t \ge 0)$, as developed in Exercise 3.4.

Conversely, assume $u = 0$. If $\eta$ is independent of $\sigma(B_t, t \ge 0)$, then, if $\mathscr{F}'_t = \mathscr{F}_t \vee \sigma(\eta)$, $B$ is also an $(\mathscr{F}'_t)_t$-standard Brownian motion (see Exercise 3.5) and now $\eta$ is $\mathscr{F}'_0$-measurable.

*Proof* We shall assume $u = 0$ for simplicity. Let us define by recurrence a sequence of processes by $\xi_0(t) \equiv \eta$ and

$$\xi_{m+1}(t) = \eta + \int_0^t b(\xi_m(s), s)\, ds + \int_0^t \sigma(\xi_m(s), s)\, dB_s\ . \tag{9.22}$$

The idea of the proof of existence is to show that the processes $(\xi_m)_m$ converge uniformly on the time interval $[0, T]$ to a process $\xi$ that will turn out to be a solution.

Let us first prove, by induction, that

$$\mathrm{E}\Big[\sup_{0\le u\le t} |\xi_{m+1}(u) - \xi_m(u)|^2\Big] \le \frac{(Rt)^{m+1}}{(m+1)!}\ , \tag{9.23}$$

where $R$ is a positive constant. For $m = 0$, thanks to Hölder's inequality,

$$\begin{aligned}\sup_{0\le u\le t} |\xi_1(u) - \eta|^2 &\le 2 \sup_{0\le u\le t} \Big|\int_0^u b(\eta, s)\, ds\Big|^2 + 2 \sup_{0\le u\le t} \Big|\int_0^u \sigma(\eta, s)\, dB_s\Big|^2\\ &\le 2t \int_0^t |b(\eta, s)|^2\, ds + 2 \sup_{0\le u\le t} \Big|\int_0^u \sigma(\eta, s)\, dB_s\Big|^2 .\end{aligned}$$

By (9.18) and (9.19),

$$\mathrm{E}\Big(\sup_{0\le u\le t}|\xi_1(u)-\eta|^2\Big)\le 4M^2t\big(t+t\mathrm{E}[|\eta|^2]\big)+16M^2\big(t+t\mathrm{E}[|\eta|^2]\big)\le Rt\,,$$

where $R=16M^2(T+1)(1+\mathrm{E}(|\eta|^2))$. Let us assume (9.23) true for $m-1$, then

$$\begin{aligned}\sup_{0\le u\le t}|\xi_{m+1}(u)-\xi_m(u)|^2\le 2\sup_{0\le u\le t}\Big|\int_0^u b(\xi_m(s),s)-b(\xi_{m-1}(s),s)\,ds\Big|^2\\ +2\sup_{0\le u\le t}\Big|\int_0^u \sigma(\xi_m(s),s)-\sigma(\xi_{m-1}(s),s)\,dB_s\Big|^2\,.\end{aligned}$$

Now, by Hölder's inequality,

$$\begin{aligned}&\sup_{0\le u\le t}\Big|\int_0^u b(\xi_m(s),s)-b(\xi_{m-1}(s),s)\,ds\Big|^2\\ &\le\sup_{0\le u\le t}u\int_0^u|b(\xi_m(s),s)-b(\xi_{m-1}(s),s)|^2\,ds\le tL^2\int_0^t|\xi_m(s)-\xi_{m-1}(s)|^2\,ds\end{aligned}$$

whereas by Doob's inequality

$$\begin{aligned}&\mathrm{E}\Big(\sup_{0\le u\le t}\Big|\int_0^u \sigma(\xi_m(s),s)-\sigma(\xi_{m-1}(s),s)\,dB_s\Big|^2\Big)\\ &\le 4\mathrm{E}\Big(\int_0^t|\sigma(\xi_m(s),s)-\sigma(\xi_{m-1}(s),s)|^2\,ds\Big)\le 4L^2\int_0^t\mathrm{E}(|\xi_m(s)-\xi_{m-1}(s)|^2)\,ds\end{aligned}$$

whence, with a possibly larger value for $R$,

$$\begin{aligned}&\mathrm{E}\Big(\sup_{0\le u\le t}|\xi_{m+1}(u)-\xi_m(u)|^2\Big)\\ &\le 2L^2t\mathrm{E}\Big(\int_0^t|\xi_m(s)-\xi_{m-1}(s)|^2\,ds\Big)+8L^2\mathrm{E}\Big(\int_0^t|\xi_m(s)-\xi_{m-1}(s)|^2\,ds\Big)\\ &\le R\int_0^t\mathrm{E}(|\xi_m(s)-\xi_{m-1}(s)|^2)\,ds\le R\int_0^t\frac{(Rs)^m}{m!}\,ds\\ &=\frac{(Rt)^{m+1}}{(m+1)!}\end{aligned}$$

and (9.23) is proved. Markov's inequality now gives

$$\begin{aligned}\mathrm{P}\Big(\sup_{0\le t\le T}|\xi_{m+1}(t)-\xi_m(t)|>\frac{1}{2^m}\Big)&\le 2^{2m}\mathrm{E}\Big[\sup_{0\le t\le T}|\xi_{m+1}(t)-\xi_m(t)|^2\Big]\\ &\le 2^{2m}\frac{(RT)^{m+1}}{(m+1)!}\,\cdot\end{aligned}$$

As the left-hand side is summable, by the Borel–Cantelli lemma,

$$\mathrm{P}\Big(\sup_{0\le t\le T}|\xi_{m+1}(t)-\xi_m(t)|>\frac{1}{2^m}\text{ for infinitely many indices }m\Big)=0\ ,$$

i.e. for almost every $\omega$ we have eventually

$$\sup_{0\le t\le T}|\xi_{m+1}(t)-\xi_m(t)|\le\frac{1}{2^m}$$

and therefore for fixed $\omega$ the series

$$\eta+\sum_{k=0}^{m-1}[\xi_{k+1}(t)-\xi_k(t)]=\xi_m(t)$$

converges uniformly on $[0,T]$ a.s. Let $\xi_t=\lim_{m\to\infty}\xi_m(t)$. Then $\xi$ is continuous, being the uniform limit of continuous processes, and therefore $\xi\in M^2_{loc}([0,T])$. Let us prove that it is a solution of (9.20). Recall that we have the relation

$$\xi_{m+1}(t)=\eta+\int_0^t b(\xi_m(s),s)\,ds+\int_0^t\sigma(\xi_m(s),s)\,dB_s\ . \tag{9.24}$$

Of course the left-hand side converges uniformly to $\xi$. Moreover, as $b$ and $\sigma$ are Lipschitz continuous (constant $L$),

$$\sup_{0\le t\le T}|b(\xi_m(t),t)-b(\xi_t,t)|\le L\sup_{0\le t\le T}|\xi_m(t)-\xi_t|\ .$$

This and the analogous inequality for $\sigma$ imply that, uniformly on $[0,T]$ a.s.,

$$\lim_{m\to\infty}b(\xi_m(t),t)=b(\xi_t,t),\qquad\lim_{m\to\infty}\sigma(\xi_m(t),t)=\sigma(\xi_t,t)\ .$$

Therefore

$$\int_0^t b(\xi_m(s),s)\,ds\ \underset{m\to\infty}{\to}\ \int_0^t b(\xi_s,s)\,ds\quad\text{a.s.}$$

and

$$\int_0^T|\sigma(\xi_m(t),t)-\sigma(\xi_t,t)|^2\,dt\ \underset{m\to\infty}{\to}\ 0\quad\text{a.s.}$$

and by Theorem 7.3 (a.s. convergence implies convergence in probability)

$$\int_0^t\sigma(\xi_m(s),s)\,dB_s\ \overset{\mathrm{P}}{\underset{m\to\infty}{\to}}\int_0^t\sigma(\xi_s,s)\,dB_s$$

so that we can take the limit in probability in (9.24) and obtain the relation

$$\xi_t = \eta + \int_0^t b(\xi_s, s)\, ds + \int_0^t \sigma(\xi_s, s)\, dB_s \ ,$$

so that $\xi$ is a solution. Of course, $\xi \in M^2([0,T])$ thanks to Theorem 9.1.

Let us turn to uniqueness. Let $\xi_1, \xi_2$ be two solutions, then

$$\begin{aligned} &|\xi_1(t) - \xi_2(t)| \\ &\le \Big| \int_0^t \big(b(\xi_1(s), s) - b(\xi_2(s), s)\big)\, ds\Big| + \Big| \int_0^t \big(\sigma(\xi_1(s), s) - \sigma(\xi_2(s), s)\big)\, dB_s\Big| \ . \end{aligned}$$

Using (9.18) and (9.19) and the Lipschitz continuity of the coefficients,

$$\begin{aligned} &\mathrm{E}\Big[\sup_{0\le u\le t} |\xi_1(u) - \xi_2(u)|^2\Big] \\ &\le 2\mathrm{E}\Big[\sup_{0\le u\le t} \Big| \int_0^u [b(\xi_1(s), s) - b(\xi_2(s), s)]\, ds\Big|^2\Big] \\ &\qquad +2\mathrm{E}\Big[\sup_{0\le u\le t} \Big| \int_0^u [\sigma(\xi_1(s), s) - \sigma(\xi_2(s), s)]\, dB_s\Big|^2\Big] \\ &\le 2t\mathrm{E}\Big[\int_0^t |b(\xi_1(s), s) - b(\xi_2(s), s)|^2\, ds\Big] \\ &\qquad +8\mathrm{E}\Big[\int_0^t |\sigma(\xi_1(s), s) - \sigma(\xi_2(s), s)|^2\, ds\Big] \\ &\le L^2[2T + 8] \int_0^t \mathrm{E}\big[|\xi_1(s) - \xi_2(s)|^2\big]\, ds \ . \end{aligned} \tag{9.25}$$

Therefore, if $v(t) = \mathrm{E}[\sup_{0\le u\le t} |\xi_1(u) - \xi_2(u)|^2]$, $v$ is bounded by Theorem 9.1, and satisfies the relation

$$v(t) \le c \int_0^t v(s)\, ds \qquad\qquad 0 \le t \le T$$

with $c = L^2(2T + 8)$; by Gronwall's inequality $v \equiv 0$ on $[0,T]$, i.e. (9.21).

□

*Remark 9.4* Let us assume that the initial condition $\eta$ is deterministic, i.e. $\eta \equiv x \in \mathbb{R}^m$ and that the starting time is $u = 0$. Let us consider the filtration $\mathscr{H}_t = \sigma(\xi_s, s \le t)$ generated by the solution $\xi$ of (9.1). Then we have $\mathscr{H}_t \subset \overline{\mathscr{G}}_t$ (recall that $(\overline{\mathscr{G}}_t)_t$ denotes the augmented natural filtration of the Brownian motion).

(continued)

*Remark 9.4* (continued)
Actually, it is immediate to check by induction that all the approximants $\xi_n$ are such that $\xi_n(t)$ is $\overline{\mathscr{G}}_t$-measurable. Hence the same holds for $\xi_t$, which is the a.s. limit of $(\xi_n(t))_n$.

The opposite inclusion is not always true. Consider, for example, the case where the last $k$ columns of the diffusion coefficient $\sigma$ vanish. In this case the solution $\xi$ would be independent of the Brownian motions $B_{d-k+1}, \dots, B_d$, which cannot be adapted to the filtration $(\mathscr{H}_t)_t$.

We shall see later that the question of whether the two filtrations $(\mathscr{H}_t)_t$ and $(\overline{\mathscr{G}}_t)_t$ coincide is of some interest.

## 9.5 Localization and existence for locally Lipschitz coefficients

In this section we prove existence and uniqueness under weaker assumptions than Lipschitz continuity of the coefficients $b$ and $\sigma$. The idea of this extension is contained in the following result, which is of great importance by itself. It states that as far as the solution $\xi$ remains inside an open set $D$, its behavior depends only on the values of the coefficients inside $D$.

**Theorem 9.3 (Localization)** Let $b_i, \sigma_i, i = 1, 2$, be measurable functions on $\mathbb{R}^m \times [u, T]$. Let $\xi_i, i = 1, 2$, be solutions of the SDE

$$\begin{aligned} d\xi_i(s) &= b_i(\xi_i(s), s)\, ds + \sigma_i(\xi_i(s), s)\, dB_s \\ \xi_i(u) &= \eta \end{aligned}$$

where $\eta$ is a square integrable, $\mathscr{F}_u$-measurable r.v. Let $D \subset \mathbb{R}^m$ be an open set such that, on $D \times [u, T]$, $b_1 = b_2$, $\sigma_1 = \sigma_2$ and, for every $x, y \in D$, $u \le t \le T$,

$$|b_i(x,t) - b_i(y,t)| \le L|x-y|, \qquad |\sigma_i(x,t) - \sigma_i(y,t)| \le L|x-y|.$$

Then, if $\tau_i$ denotes the exit time of $\xi_i$ from $D$,

$$\tau_1 \wedge T = \tau_2 \wedge T \text{ q.c.} \quad \text{and} \quad \mathrm{P}\big(\xi_1(t) = \xi_2(t) \text{ for every } u \le t \le \tau_1 \wedge T\big) = 1\ .$$

*Proof* Let, for $t \le T$,

$$\widetilde{b}(x,t) = b_1(x,t)1_D(x) = b_2(x,t)1_D(x), \ \ \widetilde{\sigma}(x,t) = \sigma_1(x,t)1_D(x) = \sigma_2(x,t)1_D(x).$$

Then it is easy to see that, as in the proof of Theorem 9.1,

$$\begin{aligned}\xi_i(t\wedge\tau_i) &= \eta + \int_0^{t\wedge\tau_i} b(\xi_i(s),s)\,ds + \int_0^{t\wedge\tau_i} \sigma(\xi_i(s),s)\,dB_s\\ &= \eta + \int_0^t \widetilde{b}(\xi_i(s\wedge\tau_i),s)\,ds + \int_0^t \widetilde{\sigma}(\xi_i(s\wedge\tau_i),s)\,dB_s\end{aligned}$$

and therefore, for $t \le T$,

$$\begin{aligned}&\sup_{0\le r\le t}|\xi_1(r\wedge\tau_1)-\xi_2(r\wedge\tau_2)|^2\\ &\le 2\sup_{0\le r\le t}\Big|\int_0^r \big(\widetilde{b}(\xi_1(s\wedge\tau_1),s)-\widetilde{b}(\xi_2(s\wedge\tau_2),s)\big)\,ds\Big|^2\\ &\qquad +2\sup_{0\le r\le t}\Big|\int_0^r \big(\widetilde{\sigma}(\xi_1(s\wedge\tau_1),s)-\widetilde{\sigma}(\xi_2(s\wedge\tau_2),s)\big)\,dB_s\Big|^2 .\end{aligned}$$

A repetition of the arguments that led us to (9.25) (Doob's inequality for the stochastic integral and Hölder's inequality for the ordinary one) gives

$$\begin{aligned}\mathrm{E}\Big[\sup_{0\le r\le t}|\xi_1(r\wedge\tau_1)-\xi_2(r\wedge\tau_2)|^2\Big] &\le L^2(2T+8)\mathrm{E}\Big(\int_0^t |\xi_1(s\wedge\tau_1)-\xi_2(s\wedge\tau_2)|^2\,ds\Big)\\ &\le L^2(2T+8)\mathrm{E}\Big(\int_0^t \sup_{0\le r\le s}|\xi_1(r\wedge\tau_1)-\xi_2(r\wedge\tau_2)|^2\,dr\Big) ,\end{aligned}$$

i.e., if we set $v(t) = \mathrm{E}\big[\sup_{0\le r\le t}|\xi_1(r\wedge\tau_1)-\xi_2(r\wedge\tau_2)|^2\big]$, $v$ satisfies the relation

$$v(t) \le L^2(2T+8)\int_0^t v(s)\,ds .$$

Now, $v$ is bounded by Theorem 9.1 and Gronwall's inequality gives $v(t) \equiv 0$, so that

$$v(T) = \mathrm{E}\Big[\sup_{0\le r\le T}|\xi_1(r\wedge\tau_1)-\xi_2(r\wedge\tau_2)|^2\Big] = 0 ,$$

whence the two statements follow simultaneously.

□

**Assumption (A')** We say that $b$ and $\sigma$ satisfy Assumption (A') if they are measurable in $(x,t)$ and

(continued)

(continued)
i) have sublinear growth (i.e. satisfy (9.9));
ii) are locally Lipschitz continuous in $x$, i.e. for every $N > 0$ there exists an $L_N > 0$ such that, if $x, y \in \mathbb{R}^m$, $|x| \le N$, $|y| \le N$, $t \in [0, T]$,

$$|b(x,t) - b(y,t)| \le L_N|x - y|, \qquad |\sigma(x,t) - \sigma(y,t)| \le L_N|x - y| ,$$

i.e. their restriction to every bounded set is Lipschitz continuous.

**Theorem 9.4** Under Assumption (A'), let $\eta$ be a square integrable $\mathscr{F}_u$-measurable r.v. Then there exists a process $\xi$ such that

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\, dt + \sigma(\xi_t, t)\, dB_t \\ \xi_u &= \eta\ . \end{aligned}$$

Moreover, $\xi \in M^2([0, T])$ and pathwise uniqueness holds.

*Proof* Existence. We shall assume $u = 0$. For $N > 0$ let $\Phi_N \in C_K^\infty(\mathbb{R}^m)$ such that $0 \le \Phi_N \le 1$ and

$$\Phi_N(x) = \begin{cases} 1 & \text{if } |x| \le N \\ 0 & \text{if } |x| \ge N + 1\ . \end{cases}$$

Let

$$b_N(x,t) = b(x,t)\Phi_N(x), \qquad \sigma_N(x,t) = \sigma(x,t)\Phi_N(x)\ .$$

$\sigma_N$ and $b_N$ therefore satisfy Assumption (A) and there exists a $\xi_N \in M^2([0, T])$ such that

$$\xi_N(t) = \eta + \int_0^t b_N(\xi_N(s), s)\, ds + \int_0^t \sigma_N(\xi_N(s), s)\, dB_s\ .$$

Let $\tau_N$ be the exit time of $\xi_N$ from the ball of radius $N$. If $N' > N$, on $\{|x| \le N\} \times [0, T]$ we have $b_N = b_{N'}$ and $\sigma_N = \sigma_{N'}$. Therefore, by the localization Theorem 9.3, the two processes $\xi_{N'}$ and $\xi_N$ coincide for $t \le T$ until their exit from the ball of radius $N$, i.e. on the event $\{\tau_N > T\}$ a.s. By Remark 9.3, moreover,

$$\mathrm{P}(\tau_N > T) = \mathrm{P}\Big(\sup_{0 \le t \le T} |\xi_N(t)| \le N\Big) \underset{N\to\infty}{\to} 1$$

and therefore $\{\tau_N > T\} \nearrow \Omega$ a.s. We can therefore define $\xi_t = \xi_N(t)$ on $\{\tau_N > T\}$: this is a well defined since if $N' > N$ then we know that $\xi_{N'}(t) = \xi_N(t)$ on $\{\tau_N > T\}$; by Theorem 7.2 on the event $\{\tau_N > T\}$ we have

$$\xi_t = \xi_N(t) = \eta + \int_0^t b_N(\xi_N(s), s)\, ds + \int_0^t \sigma_N(\xi_N(s), s)\, dB_s$$
$$= \eta + \int_0^t b(\xi_s, s)\, ds + \int_0^t \sigma(\xi_s, s)\, dB_s\ ,$$

hence, by the arbitrariness of $N$, $\xi$ is a solution. Of course, $\xi \in M^2([0, T])$ thanks to Theorem 9.1.

Uniqueness. Let $\xi_1, \xi_2$ be two solutions and let, for $i = 1, 2$, $\tau_i(N) = \inf\{t \le T; |\xi_i(t)| > N\}$. Then, by Theorem 9.3, $\tau_1(N) = \tau_2(N)$ a.s. and $\xi_1$ and $\xi_2$ coincide a.s. on $\{\tau_1(N) > T\}$. As $\{\tau_1(N) > T\} \nearrow \Omega$ a.s., $\xi_1$ and $\xi_2$ coincide for every $t \in [0, T]$ a.s.

□

*Remark 9.5* With the notations of the previous proof $\xi_N \to \xi$ uniformly for almost every $\omega$ ($\xi_N(\omega)$ and $\xi(\omega)$ even coincide a.s. on $[0, T]$ for $N$ large enough). Moreover, if $\eta \equiv x$, the event $\{\tau_N \le T\}$ on which $\xi_N$ and $\xi$ are different has a probability that goes to 0 as $N \to \infty$ uniformly for $u \in [0, T]$ and $x$ in a compact set.

*Remark 9.6* A careful look at the proofs of Theorems 9.1, 9.2, 9.3, 9.4 shows that, in analogy with ordinary differential equations, hypotheses of Lipschitz continuity of the coefficients are needed in order to guarantee local existence and uniqueness (thanks to Gronwall's inequality) whereas hypotheses of sublinear growth guarantee global existence. Exercise 9.23 presents an example of an SDE whose coefficients do not satisfy (9.9) and admits a solution that is defined only on a time interval $[0, \zeta(\omega)[$ with $\zeta < +\infty$ a.s.

## 9.6 Uniqueness in law

In this section we prove uniqueness in law of the solutions under Assumption (A').

**Theorem 9.5** Let $B_i = (\Omega^i, \mathscr{F}^i, (\mathscr{F}_t^i)_t, (B_t^i)_t, \mathrm{P}^i)$, $i = 1, 2$, be $d$-dimensional standard Brownian motions, $\eta_i, i = 1, 2$, $m$-dimensional r.v.'s in $L^2(\Omega^i, \mathscr{F}_u^i, \mathrm{P}^i)$, respectively, and having same law. Let us assume that $b$ and $\sigma$ satisfy Assumption (A') and let $\xi^i$ be the solutions of

(continued)

**Theorem 9.5** (continued)

$$\xi^i(t) = \eta_i + \int_u^t b(\xi^i(s), s)\, ds + \int_u^t \sigma(\xi^i(s), s)\, dB^i_s.$$

Then the processes $(\xi^i, B^i)$, $i = 1, 2$, have the same law.

*Proof* We assume $u = 0$. Let us first assume Assumption (A). The idea of the proof consists in proving by induction that, if $\xi^i_n$ are the approximants defined in (9.22), then the processes $(\xi^i_n, B^i)$, $i = 1, 2$, have the same law for every $n$.

This is certainly true for $n = 0$ as $\xi^i_0 \equiv \eta_i$ is independent of $B^i$ and $\eta_1$ and $\eta_2$ have the same law. Let us assume then that $(\xi^i_{n-1}, B^i)$, $i = 1, 2$, have the same law and let us prove that the same holds for $(\xi^i_n, B^i)$. This means showing that the finite-dimensional distributions of $(\xi^i_n, B^i)$ coincide. This is a consequence of the following lemma.

**Lemma 9.2** Let $B_i = (\Omega^i, \mathscr{F}^i, (\mathscr{F}^i_t)_t, (B^i_t)_t, \mathrm{P}^i)$, $i = 1, 2$, be $d$-dimensional standard Brownian motions, $\eta_i, i = 1, 2$, $m$-dimensional r.v.'s in $L^2(\Omega^i, \mathscr{F}^i_u, \mathrm{P}^i)$, respectively, and having same law. Let us assume that $b$ and $\sigma$ satisfy Assumption (A) and let $\xi^1, \xi^2$ be processes in $M^2_{loc,B^1}([0, T])$, $M^2_{loc,B^2}([0, T])$ respectively such that $(\eta_i, \xi^i, B^i)$, $i = 1, 2$, have the same law. Then, if

$$Y_i = \eta_i + \int_0^t b(\xi^i_s, s)\, ds + \int_0^t \sigma(\xi^i_s, s)\, dB^i_s\,,$$

the processes $(\eta_i, Y^i, B^i)$ have the same finite-dimensional distributions.

*Proof* If the $\xi^i$'s are elementary processes and $b$ and $\sigma$ are linear combinations of functions of the form $g(x)1_{[u,v[}(t)$ this is immediate as in this case $b(\xi^i_s, s)$ and $\sigma(\xi^i_s, s)$ are still elementary processes and by Definition 7.15 we have directly that the finite-dimensional distributions of $(\eta_1, Y^1, B^1)$ and $(\eta_2, Y^2, B^2)$ coincide. The passage to the general case is done by first approximating the $\xi^i$'s with elementary processes and then $b$ and $\sigma$ with functions as above and using the fact that a.s. convergence implies convergence in law. The details are omitted. □

*End of the Proof of Theorem 9.5* As the equality of the finite-dimensional distributions implies equality of the laws, the processes $(\xi^i_n, B^i)$ have the same law. As $(\xi^i_n, B^i)$ converges to $(\xi^i, B^i)$ uniformly in $t$ and therefore in the topology of $\mathscr{C}([0, T], \mathbb{R}^{m+d})$ (see the proof of Theorem 9.2) and the a.s. convergence implies convergence of the laws, the theorem is proved under Assumption (A).

Let us assume now Assumption (A'). By the first part of this proof, if $\xi^1_N, \xi^2_N$ are the processes as in the proof of Theorem 9.4, $(\xi^1_N, B^1)$ and $(\xi^2_N, B^2)$ have the same law. As $\xi^i_N \to \xi^i$ uniformly a.s. (see Remark 9.5), the laws of $(\xi^i_N, B^i)$, $i = 1, 2$, converge as $N \to \infty$ to those of $(\xi^i, B^i)$, which therefore coincide. □

## 9.7 The Markov property

In this section we prove that, under Assumption (A'), the solution of an SDE is a Markov process and actually a diffusion associated to a generator that the reader can already imagine. In Sect. 9.9 we shall consider the converse problem, i.e. whether for a given differential operator there exists a diffusion process that is associated to it and whether it is unique in law.

Let, for an $\mathscr{F}_s$-measurable r.v. $\eta$, $\xi^{\eta,s}_t$ be the solution of

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\, dt + \sigma(\xi_t, t)\, dB_t \\ \xi_s &= \eta \end{aligned} \tag{9.26}$$

and let us prove first the Markov property, i.e. that, for every $\Gamma \in \mathscr{B}(\mathbb{R}^m)$ and $s \le u \le t$,

$$\mathrm{P}\big(\xi^{\eta,s}_t \in \Gamma \,|\, \mathscr{F}^s_u\big) = p(u, t, \xi^{\eta,s}_u, \Gamma) \tag{9.27}$$

for some transition function $p$. We know that the transition function $p(u, t, x, \cdot)$ is the law at time $t$ of the process starting at $x$ at time $u$, i.e. if $\xi^{x,u}_t$ is the solution of

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\, dt + \sigma(\xi_t, t)\, dB_t \\ \xi_u &= x \end{aligned} \tag{9.28}$$

then the candidate transition function is, for $\Gamma \in \mathscr{B}(\mathbb{R}^m)$,

$$p(u, t, x, \Gamma) = \mathrm{P}(\xi^{x,u}_t \in \Gamma)\,. \tag{9.29}$$

In order to prove the Markov property (9.27) we note that

$$\begin{aligned} \xi^{\eta,s}_t &= \eta + \int_s^t b(\xi^{\eta,s}_v, v)\, dv + \int_s^t \sigma(\xi^{\eta,s}_v, v)\, dB_v \\ &= \underbrace{\eta + \int_s^u b(\xi^{\eta,s}_v, v)\, dv + \int_s^u \sigma(\xi^{\eta,s}_v, v)\, dB_v}_{=\xi^{\eta,s}_u} \\ &\qquad + \int_u^t b(\xi^{\eta,s}_v, v)\, dv + \int_u^t \sigma(\xi^{\eta,s}_v, v)\, dB_v\,, \end{aligned} \tag{9.30}$$

i.e. the value at time $t$ of the solution starting at $\eta$ at time $s$ is the same as the value of the solution starting at $\xi_u^{\eta,s}$ at time $u$, i.e.

$$\xi_t^{\eta,s} = \xi_t^{\xi_u^{\eta,s},u} . \tag{9.31}$$

Let us define

$$\psi(x,\omega) = 1_\Gamma(\xi_t^{x,u}(\omega))$$

and observe that

$$\phi(x) := \mathrm{E}[\psi(x,\cdot)] = \mathrm{E}[1_\Gamma(\xi_t^{x,u})] = \mathrm{P}(\xi_t^{x,u} \in \Gamma) = p(u,x,t,\Gamma) . \tag{9.32}$$

By (9.31)

$$\psi(\xi_u^{\eta,s},\omega) = 1_\Gamma(\xi_t^{\xi_u^{\eta,s},u}(\omega))$$

and, going back to (9.27),

$$\mathrm{P}\big(\xi_t^{\eta,s} \in \Gamma \,|\, \mathscr{F}_u^s\big) = \mathrm{P}\big(\xi_t^{\xi_u^{\eta,s},u} \in \Gamma \,|\, \mathscr{F}_u^s\big) = \mathrm{E}\big(\psi(\xi_u^{\eta,s},\cdot) \,|\, \mathscr{F}_u^s\big) .$$

The r.v. $\xi_u^{\eta,s}$ is $\mathscr{F}_u^s$-measurable. If we prove that the r.v. $\xi_t^{x,u}$ is independent of $\mathscr{F}_u^s$, then $\psi(x,\cdot)$ is independent of $\mathscr{F}_u^s$ and we can apply the freezing lemma (Lemma 4.1) and obtain

$$\mathrm{P}\big(\xi_t^{\eta,s} \in \Gamma \,|\, \mathscr{F}_u^s\big) = \mathrm{E}\big[\psi(\xi_u^{\eta,s},\cdot) \,|\, \mathscr{F}_u^s\big] = \phi(\xi_u^{\eta,s}) = p(u,\xi_u^{\eta,s},t,\Gamma)$$

($\phi$ is defined above in (9.32)), which proves the Markov property.

Hence in order to complete this proof we still have to

a) prove that $\xi_t^{x,u}$ is a r.v. that is independent of $\mathscr{F}_u^s$, and, unfortunately, also that
b) the map $(x,u) \mapsto \xi_t^{x,u}$ is measurable for every $\omega$ or at least that there is a measurable version of it. This is the object of the next section, where (Theorem 9.9) it is shown that, under Assumption (A), there exists a version of $(x,u,t) \mapsto \xi_t^{x,u}$ that is continuous in the three arguments $x, u, t$ (it will be another application of Theorem 2.1, the continuity Kolmogorov theorem).

Fact a) above follows easily from the following lemma. Let $\mathscr{H}_u^0 = \sigma(B_v - B_u, v \geq u)$ be the $\sigma$-algebra generated by the increments of $B$ after time $u$ and $\mathscr{H}_u$ the $\sigma$-algebra obtained by adding to $\mathscr{H}_u^0$ the negligible sets of $\mathscr{F}_\infty$. Of course (Exercise 3.4) $\mathscr{H}_u$ is independent of $\mathscr{F}_u$.

**Lemma 9.3** Let $Y \in M^2$ be a process such that $Y_s$ is $\mathscr{H}_u$-measurable for every $s \geq u$. Then the r.v.

$$\int_u^t Y_s \, dB_s \tag{9.33}$$

is also $\mathscr{H}_u$-measurable for every $t \geq u$ and therefore independent of $\mathscr{F}_u$.

*Proof* The statement is immediate if $Y = \sum_{i=1}^m X_i 1_{[t_i, t_{i+1}[}$ is an elementary process. In this case the r.v.'s $X_i$ are $\mathscr{H}_u$-measurable and in the explicit expression (7.7) only the increments of $B$ after time $u$ and the values of the $X_i$ appear. It is also immediate that in the approximation procedure of a general integrand with elementary processes described at the beginning of Sect. 7.3 if $Y$ is $\mathscr{H}_u$-measurable, then the approximating elementary processes are also $\mathscr{H}_u$-measurable. Then just observe that the stochastic integral in (9.33) is the a.s. limit of the integrals of the approximating elementary processes.

□

Let us assume Assumption (A). If we go back to the successive approximations scheme that is the key argument in the proof of existence in Theorem 9.2, we see that $(\xi_t^{x,u})_{t \geq u}$ is the limit of the approximating processes $\xi_m$, where $\xi_0(t) \equiv x$ and

$$\xi_{m+1}(t) = x + \int_u^t b(\xi_m(v), v) \, dv + \int_u^t \sigma(\xi_m(v), v) \, dB_v \; .$$

The deterministic process $\xi_0 \equiv x$ is obviously $\mathscr{H}_u$-measurable and by recurrence, using Lemma 9.3, the processes $\xi_m$ are such that $\xi_m(t)$ also is $\mathscr{H}_u$-measurable, as well as $(\xi_t^{x,u})_{t \geq u}$ which is their a.s. limit.

Hence, given as granted b) above, we have proved the Markov property under Assumption (A).

**Theorem 9.6** Under Assumption (A'), $p$ defined in (9.29) is a transition function and the process $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (\xi_t^{x,s})_t, \mathrm{P})$ is a Markov process starting at time $s$, with initial distribution $\delta_x$ and associated to the transition function $p$.

*Proof* Under Assumption (A) the Markov property has already been proved. Under Assumption (A') we shall approximate the solution with processes that are solutions of an SDE satisfying Assumption (A): let $\xi_N$ and $\tau_N$ be as in the proof of Theorem 9.4 and let

$$p_N(s, t, x, \Gamma) := \mathrm{P}(\xi_N^{x,s}(t) \in \Gamma) \; .$$

$\xi_N$ is the solution of an SDE with coefficients satisfying Assumption (A), and therefore enjoying the Markov property, i.e.

$$\mathrm{P}(\xi_N^{x,s}(t) \in \Gamma \,|\, \mathscr{F}_u^s) = p_N(u, t, \xi_N^{x,s}(u), \Gamma) \ .$$

We now have to pass to the limit as $N \to \infty$ in the previous equation. As $\xi_t^{x,s}$ and $\xi_N^{x,s}(t)$ coincide on $\{\tau_N > t\}$ and $\mathrm{P}(\tau_N \le t) \to 0$ as $N \to \infty$, we have, using the fact that probabilities pass to the limit with a monotone sequences of events,

$$\begin{aligned} p_N(s,t,x,\Gamma) &= \mathrm{P}(\xi_N^{x,s}(t) \in \Gamma) \\ = \mathrm{P}(\xi_t^{x,s} \in \Gamma, \tau_N > t) + \mathrm{P}(\xi_N^{x,s}(t) \in \Gamma, \tau_N \le t) &\underset{N\to\infty}{\to} \mathrm{P}(\xi_t^{x,s} \in \Gamma) \\ &= p(s,t,x,\Gamma) \ . \end{aligned} \tag{9.34}$$

From this relation we derive that

$$p_N(u, t, \xi_N^{x,s}(u), \Gamma) \underset{N\to\infty}{\to} p(u, t, \xi^{x,s}(u), \Gamma)$$

as $\xi_N^{x,s}(u) = \xi^{x,s}(u)$ for $N$ large enough. Moreover,

$$\begin{aligned} \mathrm{P}\big(\xi_t^{x,s} \in \Gamma \,|\, \mathscr{F}_u^s\big) &= \mathrm{E}\big[1_\Gamma(\xi_t^{x,s}) 1_{\{\tau_N > t\}} \,|\, \mathscr{F}_u^s\big] + \mathrm{E}\big[1_\Gamma(\xi_t^{x,s}) 1_{\{\tau_N \le t\}} \,|\, \mathscr{F}_u^s\big] \\ &= \mathrm{E}\big[1_\Gamma(\xi_N^{x,s}(t)) 1_{\{\tau_N > t\}} \,|\, \mathscr{F}_u^s\big] + \mathrm{E}\big[1_\Gamma(\xi_t^{x,s}) 1_{\{\tau_N \le t\}} \,|\, \mathscr{F}_u^s\big] \\ &= \mathrm{E}\big[1_\Gamma(\xi_N^{x,s}(t)) \,|\, \mathscr{F}_u^s\big] + \mathrm{E}\big[(1_\Gamma(\xi_t^{x,s}) - 1_\Gamma(\xi_N^{x,s}(t))) 1_{\{\tau_N \le t\}} \,|\, \mathscr{F}_u^s\big] \ . \end{aligned}$$

The second term on the right-hand side tends to zero a.s., as $\mathrm{P}(\tau_N \le t) \searrow 0$ as $N \to \infty$, hence

$$\lim_{n\to\infty} \mathrm{P}(\xi_N^{x,s}(t) \in \Gamma \,|\, \mathscr{F}_u^s) = \lim_{n\to\infty} \mathrm{E}\big[1_\Gamma(\xi_N^{x,s}(t)) \,|\, \mathscr{F}_u^s\big] = \mathrm{P}\big(\xi_t^{x,s} \in \Gamma \,|\, \mathscr{F}_u^s\big) \qquad a.s.$$

and putting things together

$$\begin{aligned} \mathrm{P}(\xi_t^{x,s} \in \Gamma \,|\, \mathscr{F}_u^s) &= \lim_{N\to\infty} \mathrm{E}[1_\Gamma(\xi_N^{x,s}(t)) \,|\, \mathscr{F}_u^s] \\ = \lim_{N\to\infty} p_N(u, t, \xi_N^{x,s}(u), \Gamma) &= p(u, t, \xi^{x,s}(u), \Gamma) \ . \end{aligned}$$

The Markov property is therefore proved under Assumption (A'). We still have to prove that $p$ satisfies the Chapman–Kolmogorov equation; this is a consequence of the argument described in Remark 6.2: as $p(s,t,x,\cdot)$ is the law of $\xi_t^{x,s}$

$$\begin{aligned} p(s,t,x,\Gamma) = \mathrm{P}(\xi_t^{x,s} \in \Gamma) = \mathrm{E}[\mathrm{P}(\xi_t^{x,s} \in \Gamma \,|\, \mathscr{F}_u^s)] &= \mathrm{E}[p(u,t,\xi_u^{x,s},\Gamma)] \\ = \int p(u,t,y,\Gamma) p(s,u,x,dy) \ . \end{aligned}$$

□

We can now construct a realization of the Markov process associated to the solution of an SDE. In order to do this, let $\mathscr{C} = \mathscr{C}([0,T], \mathbb{R}^m)$ and let $X_t : \mathscr{C} \to \mathbb{R}^m$ be the applications defined by $X_t(\gamma) = \gamma(t)$. Let $\mathscr{M} = \mathscr{B}(\mathscr{C})$, $\mathscr{M}_t^s = \sigma(X_u, s \le u \le t)$, $\mathscr{M}_\infty^s = \sigma(X_u, u \ge s)$. On $\mathscr{C}$ let $\mathrm{P}^{x,s}$ be the law of the process $\xi^{x,s}$ (Sect. 3.2) and denote by $\mathrm{E}^{x,s}$ the expectation computed with respect to $\mathrm{P}^{x,s}$. Of course, if $p$ is the transition function defined in (9.29), we have $p(s,t,x,\Gamma) = \mathrm{P}^{x,s}(X_t \in \Gamma)$.

As, by definition, the finite-dimensional distributions of $(X_t)_{t\ge s}$ with respect to $\mathrm{P}^{x,s}$ coincide with those of $\xi^{x,s}$ with respect to P, the following theorem is obvious.

**Theorem 9.7** Under Assumption (A'), $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t^s)_{t\ge s}, (X_t)_{t\ge s}, (\mathrm{P}^{x,s})_{x,s})$ is a realization of the Markov process associated to the transition function $p$ defined in (9.29).

Thanks to Theorem 9.5, which guarantees the uniqueness in law, the probability $\mathrm{P}^{x,s}$ does not depend on the Brownian motion that is used to construct the solution $\xi^{x,s}$. Therefore the realization $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t^s)_{t\ge s}, (X_t)_{t\ge s}, (\mathrm{P}^{x,s})_{x,s})$ is well defined. We shall call this family of processes the *canonical Markov process associated to the SDE* (9.1).

**Theorem 9.8** Under Assumption (A') the canonical Markov process associated to (9.1) is a diffusion with generator

$$L_t = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x,t)\frac{\partial}{\partial x_i},$$

where $a(x,t) = \sigma(x,t)\sigma(x,t)^*$.

*Proof* Let us prove first the Feller property which will imply strong Markovianity thanks to Theorem 6.1. We must prove that, for every $h > 0$ and for every bounded continuous function $f$, the map

$$(x,t) \mapsto \int f(y)p(t,t+h,x,dy) = \mathrm{E}^{x,t}[f(X_{t+h})] = \mathrm{E}[f(\xi_{t+h}^{x,t})]$$

is continuous. Under Assumption (A) this is immediate, using the fact that $\xi_t^{x,s}(\omega)$ is a continuous function of $x, s, t, t \ge s$ thanks to the forthcoming Theorem 9.9. Under Assumption (A') if $\xi_N^{x,s}$ and $\tau_N$ are as in the proof of Theorem 9.4 then as $\xi^{x,s}$ and $\xi_N^{x,s}$ coincide on $\{\tau_N > T\}$,

$$\mathrm{E}[f(\xi_{t+h}^{x,t})] = \mathrm{E}[f(\xi_N^{x,t}(t+h))] + \mathrm{E}\big[(f(\xi_{t+h}^{x,t}) - f(\xi_N^{x,t}(t+h)))1_{\{\tau_N \le T\}}\big].$$

Now it is easy to deduce that $(x,t) \mapsto \mathrm{E}[f(\xi^{x,t}_{t+h})]$ is continuous, as the first term on the right-hand side is continuous in $(x,t)$ whereas the second one is majorized by $2\|f\|_\infty \mathrm{P}(\tau_N \le T)$ and can be made small uniformly for $(x,t)$ in a compact set, thanks to Remark 9.3.

Let us prove that $L_t$ is the generator. By Ito's formula, as explained in Remark 9.1, if $f \in C^2_K$,

$$T_{s,t}f(x) = \mathrm{E}[f(\xi^{x,s}_t)]$$
$$= f(x) + \mathrm{E}\Big[\int_s^t L_u f(\xi^{x,s}_u)\,du\Big] + \mathrm{E}\Big[\int_s^t \sum_{i=1}^m \sum_{j=1}^d \sigma_{ij}(\xi^{x,s}_u, u)\frac{\partial f}{\partial x_i}(\xi^{x,s}_u)\,dB_j(u)\Big]\,.$$

The last integrand is a process of $M^2([s,t])$, as $\xi^{x,s} \in M^2([s,t])$ and $\sigma$ has a sublinear growth (the derivatives of $f$ are bounded), therefore the expectation of the stochastic integral vanishes and we have

$$T_{s,t}f(x) = f(x) + \int_s^t \mathrm{E}[L_u f(\xi^{x,s}_u)]\,du = f(x) + \int_s^t T_{s,u}(L_u f)(x)\,du\,.$$

□

*Example 9.3* What is the transition function $p(t,x,\cdot)$ of the Ornstein–Uhlenbeck process, i.e. the Markov process that is the solution of

$$d\xi_t = -\lambda \xi_t\,dt + \sigma\,dB_t\ ? \tag{9.35}$$

By (9.29) $p(t,x,\cdot)$ is the law of $\xi^x_t$, i.e. of the position at time $t$ of the solution of (9.35) starting at $x$. We have found in Sect. 9.2 that

$$\xi_t = \mathrm{e}^{-\lambda t}x + \mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{\lambda s}\sigma\,dB_s\,.$$

Now simply observe that $\xi_t$ has a distribution that is Gaussian with mean $\mathrm{e}^{-\lambda t}x$ and variance

$$\sigma^2_t := \int_0^t \mathrm{e}^{-2\lambda(t-s)}\sigma^2\,ds = \frac{\sigma^2}{2\lambda}\,(1-\mathrm{e}^{-2\lambda t})\,.$$

This observation can be useful if you want to simulate the process $\xi$: just fix a discretization step $h$. Then choose at random a number $z$ with distribution $N(\mathrm{e}^{-\lambda h}x, \sigma^2_h)$ and set $\xi_h = z$. Then choose at random a number $z$ with distribution $N(\mathrm{e}^{-\lambda h}\xi_h, \sigma^2_h)$ and set $\xi_{2h} = z$ and so on: the position $\xi_{(k+1)h}$ will be obtained by sampling a number with distribution $N(\mathrm{e}^{-\lambda h}\xi_{kh}, \sigma^2_h)$.

(continued)

*Example 9.3* (continued)
This procedure allows us to simulate subsequent positions of an Ornstein–Uhlenbeck process *exactly*, which means that the positions $(x, \xi_h, \dots, \xi_{mh})$ have exactly the same joint distributions as the positions of the Markov process which is the solution of (9.35) at the times $0, h, \dots, mh$.

This procedure can be easily extended to the case of a multidimensional Ornstein–Uhlenbeck process, the only difference being that in that case the transition function is a multivariate Gaussian distribution.

Chapter 11 is devoted to the simulation problems of a more general class of diffusion processes.

*Example 9.4* Let $B$ be a real Brownian motion and $X$ the solution of the two-dimensional SDE

$$\begin{aligned} d\xi_1(t) &= b_1(\xi_1(t))\,dt + \xi_1(t)\,dB_t \\ d\xi_2(t) &= b_2(\xi_2(t))\,dt + \xi_2(t)\,dB_t\ . \end{aligned}$$

What is the differential generator of $\xi$?

This equation can be written as

$$d\xi_t = b(\xi_t)\,dt + \sigma(\xi_t)\,dB_t\ ,$$

where $b(x_1, x_2) = \begin{pmatrix} b_1(x_1) \\ b_2(x_2) \end{pmatrix}$ and $\sigma(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$: here the Brownian motion $B$ is one-dimensional.

In order to obtain the generator, the only thing to be computed is the matrix $a = \sigma\sigma^*$ of the second-order coefficients. We have

$$a(x) = \sigma(x)\sigma^*(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} (x_1\ x_2) = \begin{pmatrix} x_1^2 & x_1x_2 \\ x_1x_2 & x_2^2 \end{pmatrix}\ .$$

Therefore

$$L = \frac{1}{2}\Big(x_1^2 \frac{\partial^2}{\partial x_1^2} + x_2^2 \frac{\partial^2}{\partial x_2^2} + 2x_1x_2 \frac{\partial^2}{\partial x_1 \partial x_2}\Big) + b_1(x_1)\frac{\partial}{\partial x_1} + b_2(x_2)\frac{\partial}{\partial x_2}\ \cdot$$

And if it was

$$\begin{aligned} d\xi_1(t) &= b_1(\xi_1(t))\,dt + \xi_1(t)\,dB_1(t) \\ d\xi_2(t) &= b_2(\xi_2(t))\,dt + \xi_2(t)\,dB_2(t)\ , \end{aligned}$$

(continued)

*Example 9.4* (continued)
where $B = (B_1, B_2)$ is a two-dimensional Brownian motion?

Now

$$\sigma(x) = \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}$$

and

$$a(x) = \sigma(x)\sigma^*(x) = \begin{pmatrix} x_1^2 & 0 \\ 0 & x_2^2 \end{pmatrix}$$

so that

$$L = \frac{1}{2}\Big(x_1^2 \frac{\partial^2}{\partial x_1^2} + x_2^2 \frac{\partial^2}{\partial x_2^2}\Big) + b_1(x_1)\frac{\partial}{\partial x_1} + b_2(x_2)\frac{\partial}{\partial x_2} \cdot$$

## 9.8 $L^p$ bounds and dependence on the initial data

Let us see how the solution depends on the initial value $\eta$ and initial time $u$. What if one changes the starting position or the starting time "just a little"? Does the solution change "just a little" too? Or, to be precise, is it possible to construct solutions so that there is continuity with respect to initial data?

In some situations, when we are able to construct explicit solutions, the answer to this question is immediate. For instance, if $\xi_s^{x,t}$ is the position at time $s$ of a Brownian motion starting at $x$ at time $t$, we can write

$$\xi_s^{x,t} = x + (B_t - B_s) ,$$

where $B$ is a Brownian motion. It is immediate that this is, for every $\omega$, a continuous function of the starting position $x$, of the starting time $t$ and of the actual time $s$. Similar arguments can be developed for other processes for which we have explicit formulas, such as the Ornstein–Uhlenbeck process or the geometric Brownian motion of Sect. 9.2.

The aim of this chapter is to give an answer to this question in more general situations. It should be no surprise that the main tool will be in the end Kolmogorov's continuity Theorem 2.1.

**Proposition 9.1** Under Assumption (A), let $\xi_1$ and $\xi_2$ be solutions of (9.20) with initial conditions $\eta_1, u$ and $\eta_2, v$, respectively. Let us assume $u \le v$. Then, if $p \ge 2$,

$$
\begin{aligned}
&\mathrm{E}\Big[\sup_{v\le t\le T} |\xi_1(t)-\xi_2(t)|^p\Big] \\
&\le c(L,T,M,p)\big(1+\mathrm{E}(|\eta_1|^p)\big)\big(|u-v|^{p/2}+\mathrm{E}(|\eta_1-\eta_2|^p)\big)\ .
\end{aligned}
\tag{9.36}
$$

*Proof* For $t \ge v$

$$
\begin{aligned}
&\xi_1(t)-\xi_2(t) \\
&= \xi_1(v)-\eta_2+\int_v^t \big[b(\xi_1(r),r)-b(\xi_2(r),r)\big]\,dr+\int_v^t \big[\sigma(\xi_1(r),r)-\sigma(\xi_2(r),r)\big]\,dB_r
\end{aligned}
$$

and therefore, again using Hölder's inequality on the integral in $dt$ and Doob's inequality on the stochastic integral,

$$
\begin{aligned}
&\mathrm{E}\Big[\sup_{v\le t\le s} |\xi_1(t)-\xi_2(t)|^p\Big] \\
&\le 3^{p-1}\mathrm{E}\big[|\xi_1(v)-\eta_2|^p\big]+3^{p-1}T^{p-1}\mathrm{E}\Big[\int_v^s |b(\xi_1(r),r)-b(\xi_2(r),r)|^p\,dr\Big] \\
&\qquad +3^{p-1}\mathrm{E}\Big[\sup_{v\le t\le s}\Big|\int_v^t [\sigma(\xi_1(r),r)-\sigma(\xi_2(r),r)]\,dB_r\Big|^p\Big] \\
&\le 3^{p-1}\mathrm{E}\big[|\xi_1(v)-\eta_2|^p\big]+(L,T,M,p)\int_v^s \mathrm{E}\Big[\sup_{v\le t\le r}|\xi_1(t)-\xi_2(t)|^p\Big]\,dr\ .
\end{aligned}
$$

The function $s\mapsto \mathrm{E}[\sup_{u\le t\le s}|\xi_1(t)-\xi_2(t)|^p]$ is bounded thanks to (9.12). One can therefore apply Gronwall's inequality and get

$$
\mathrm{E}\Big[\sup_{v\le t\le T}|\xi_1(t)-\xi_2(t)|^p\Big]\le c(L,T,M,p)\big(1+\mathrm{E}\big[|\xi_1(v)-\eta_2|^p\big]\big)\ .
$$

Now, by (9.13),

$$
\begin{aligned}
\mathrm{E}\big[|\xi_1(v)-\eta_2|^p\big] &\le 2^{p-1}\big(\mathrm{E}\big[|\xi_1(v)-\eta_1|^p\big]+\mathrm{E}\big[|\eta_1-\eta_2|^p\big]\big) \\
&\le c(T,M,p)\big(1+\mathrm{E}\big[|\eta_1|^p\big]\big)|v-u|^{p/2}+\mathrm{E}\big[|\eta_1-\eta_2|^p\big] \\
&\le c'(T,M,p)\big(1+\mathrm{E}\big[|\eta_1|^p\big]\big)\big(|v-u|^{p/2}+\mathrm{E}\big[|\eta_1-\eta_2|^p\big]\big)\ ,
\end{aligned}
$$

which allows us to conclude the proof. □

Let us denote by $\xi_t^{x,s}$ the solution of

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\,dt + \sigma(\xi_t, t)\,dB_t \\ \xi_s &= x\,. \end{aligned} \tag{9.37}$$

A classical result of ordinary differential equations is that, under suitable conditions, the solution depends continuously on the initial data $x, s$. The corresponding result for SDEs is the following.

**Theorem 9.9** Under Assumption (A) there exists a family $(Z_{x,s}(t))_{x,s,t}$ of r.v.'s such that

a) the map $(x, s, t) \mapsto Z_{x,s}(t)$ is continuous for every $\omega$ for $x \in \mathbb{R}^m, s, t \in \mathbb{R}^+, s \le t$.
b) $Z_{x,s}(t) = \xi_t^{x,s}$ a.s. for every $x \in \mathbb{R}^m$, $s, t \in \mathbb{R}^+, s \le t$.

*Proof* This is an application of Kolmogorov's continuity Theorem 2.1. Note that, by (9.30), for every $s \ge u$, $\xi^{x,u}$ is a solution of

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\,dt + \sigma(\xi_t, t)\,dB_t \\ \xi_s &= \xi_s^{x,u}\,. \end{aligned}$$

Let us fix $T > 0$ and let $u, v, s, t$ be times with $u \le s$, $v \le t$. We want to prove, for $x, y$ with $|x|, |y| \le R$ the inequality

$$\mathrm{E}\big[|\xi_s^{x,u} - \xi_t^{y,v}|^p\big] \le c(L, M, T, p, R)\big(|x - y|^p + |t - s|^{p/2} + |u - v|^{p/2}\big)\,. \tag{9.38}$$

Assume first $u \le v \le s \le t$. We have

$$\mathrm{E}\big[|\xi_s^{x,u} - \xi_t^{y,v}|^p\big] \le 2^{p-1}\mathrm{E}\big[|\xi_s^{x,u} - \xi_s^{y,v}|^p\big] + 2^{p-1}\mathrm{E}\big[|\xi_s^{y,v} - \xi_t^{y,v}|^p\big]\,.$$

Using Proposition 9.1 in order to bound the first term on the right-hand side, we have

$$\mathrm{E}\big[|\xi_s^{x,u} - \xi_s^{y,v}|^p\big] \le c(L, T, M, p)\big(1 + |x|^p\big)\big(|u - v|^{p/2} + |x - y|^p\big)\,,$$

whereas for the second one, recalling that $\xi_t^{y,v}$ is the solution at time $t$ of the SDE with initial condition $\xi_s^{y,v}$ at time $s$, (9.12) and (9.13) give

$$\begin{aligned} \mathrm{E}\big[|\xi_s^{y,v} - \xi_t^{y,v}|^p\big] &\le c(p, T, M)|t - s|^{p/2}(1 + \mathrm{E}[|\xi_s^{y,v}|^p]) \\ &\le c(p, T, M)|t - s|^{p/2}(1 + |y|^p)\,, \end{aligned}$$

which together give (9.38). The same argument proves (9.38) in the case $u \le v \le t \le s$.

If, instead, $u \le s \le v \le t$, we must argue differently (in this case $\xi_s^{y,v}$ is not defined); we have

$$\begin{aligned}
&\mathrm{E}\big[|\xi_s^{x,u} - \xi_t^{y,v}|^p\big] \\
&\le 3^{p-1}\mathrm{E}\big[|\xi_s^{x,u} - \xi_s^{y,u}|^p\big] + 3^{p-1}\mathrm{E}\big[|\xi_s^{y,u} - \xi_t^{y,u}|^p\big] + 3^{p-1}\mathrm{E}\big[|\xi_t^{y,u} - \xi_t^{y,v}|^p\big]
\end{aligned}$$

and again Proposition 9.1, (9.12) and (9.13) give (9.38). In conclusion, for $p$ large enough,

$$\mathrm{E}\big[|\xi_s^{x,u} - \xi_t^{y,v}|^p\big] \le c\big(|x-y| + |t-s| + |u-v|\big)^{m+2+\alpha}$$

for some $\alpha > 0$. Theorem 2.1 guarantees therefore the existence of a continuous version of $\xi_t^{x,s}$ in the three variables $x, s, t, s \le t$ for $|x| \le R$. $R$ being arbitrary, the statement is proved.

□

Note that in the previous proof we can choose $\alpha = \frac{p}{2} - m - 2$ and a more precise application of Theorem 2.1 allows us to state that the paths of $\xi^{x,u}$ are Hölder continuous with exponent $\gamma < \frac{1}{2p}(p - 2m) = \frac{1}{2}(1 - \frac{m}{p})$ and therefore, by the arbitrariness of $p$, Hölder continuous with exponent $\gamma$ for every $\gamma < \frac{1}{2}$, as for the Brownian motion.

## 9.9 The square root of a matrix field and the problem of diffusions

Given the differential operator

$$L_t = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x,t)\frac{\partial}{\partial x_i}\,, \tag{9.39}$$

where the matrix $a(x,t)$ is positive semidefinite, two questions naturally arise:

- under which conditions does a diffusion associated to $L_t$ exist?
- what about uniqueness?

From the results of the first sections of this chapter it follows that the answer to the first question is positive *provided* there exists a matrix field $\sigma(x,t)$ such that $a(x,t) = \sigma(x,t)\sigma(x,t)^*$ and that $\sigma$ and $b$ satisfy Assumption (A') (local Lipschitz continuity and sublinear growth in the $x$ variable, in addition to joint measurability).

In general, as $a(x,t)$ is positive semidefinite, for fixed $x, t$ there always exists an $m \times m$ matrix $\sigma(x,t)$ such that $\sigma(x,t)\sigma(x,t)^* = a(x,t)$. Moreover, it is unique under the additional assumption that it is symmetric. We shall denote this symmetric matrix field by $\sigma$, so that $\sigma(x,t)^2 = a(x,t)$.

Let us now investigate the regularity of $\sigma$. Is it possible to take such a square root in such a way that $(t,x) \mapsto \sigma(x,t)$ satisfies Assumptions (A) or (A')? Note also that, due to the lack of uniqueness of this square root $\sigma$, one will be led to enquire whether two different square root fields $\sigma(x,t)$ might produce different diffusions and therefore if uniqueness is preserved. We shall mention the following results. See Friedman (1975, p. 128), Priouret (1974, p. 81), or Stroock and Varadhan (1979, p. 131) for proofs and other details.

**Proposition 9.2** Let $D \subset \mathbb{R}^m$ be an open set and let us assume $a(x,t)$ is positive definite for every $(x,t) \in D \times [0,T]$. Then if $a$ is measurable in $(x,t)$, the same is true for $\sigma$. If $a$ is locally Lipschitz continuous in $x$ then this is also true for $\sigma$. If $a$ is Lipschitz continuous in $x$ and uniformly positive definite, $\sigma$ is Lipschitz continuous in $x$.

To be precise, let us recall that "$a(x,t)$ positive definite" means that

$$\langle a(x,t)z, z\rangle = \sum_{i,j=1}^m a_{ij}(x,t) z_i z_j > 0 \tag{9.40}$$

for every $z \in \mathbb{R}^m$, $|z| > 0$, or, equivalently, that, for every $x, t$, the smallest eigenvalue of $a(x,t)$ is $> 0$. "$a$ uniformly positive definite" means

$$\langle a(x,t)z, z\rangle = \sum_{i,j=1}^m a_{ij}(x,t) z_i z_j > \lambda |z|^2 \tag{9.41}$$

for every $z \in \mathbb{R}^m$ and some $\lambda > 0$ or, equivalently, that there exists a $\lambda > 0$ such that the smallest eigenvalue of $a(x,t)$ is $> \lambda$ for every $x, t$.

**Definition 9.5** We say that the matrix field $a$ is *elliptic* if $a(x,t)$ is positive definite for every $x, t$ and that it is *uniformly elliptic* if $a(x,t)$ is uniformly positive definite.

*Remark 9.7* Note that

$$\langle a(x,t)z, z\rangle = \langle \sigma(x,t)\sigma^*(x,t)z, z\rangle = \langle \sigma^*(x,t)z, \sigma^*(x,t)z\rangle = |\sigma^*(x,t)z|^2$$

(continued)

*Remark 9.7* (continued)
hence (9.40) is equivalent to requiring that

$$|\sigma^*(x,t)z|^2 > 0$$

for every $z \in \mathbb{R}^m$, $z \neq 0$, i.e. the matrix $\sigma(x,t)^* : \mathbb{R}^m \to \mathbb{R}^d$ must be injective. As $\sigma(x,t)$ is an $m \times d$ matrix, (9.40) can hold only if $d \geq m$, otherwise $\sigma^*(x,t)$ would have a kernel of dimension $\geq 1$ (recall that $m$ is the dimension of the diffusion process and $d$ the dimension of the driving Brownian motion).

Of course (9.40) implies that $a(x,t)$ is invertible for every $x, t$ and (9.41) that

$$\langle a(x,t)^{-1}w, w\rangle \leq \frac{1}{\lambda}\,|w|^2\,.$$

If $a$ is not positive definite, the square root $\sigma$ may even not be locally Lipschitz continuous, even if $a$ is Lipschitz continuous. Just consider, for $m = 1$, the case $a(x) = |x|$ and therefore $\sigma(x) = \sqrt{|x|}$. However, we have the following result.

**Proposition 9.3** Let $D \subset \mathbb{R}^m$ be an open set and assume that $a(x,t)$ is positive semidefinite for every $(x,t) \in D \times [0,T]$. If $a$ is of class $C^2$ in $x$ on $D \times [0,T]$, then $\sigma$ is locally Lipschitz continuous in $x$. If, moreover, the derivatives of order 2 of $a$ are bounded, $\sigma$ is Lipschitz continuous in $x$.

We can now state the main result of this section.

**Theorem 9.10** Let us assume that $b$ satisfies Assumption (A') and that $a(x,t)$ is measurable in $(x,t)$, positive semidefinite for every $(x,t)$ and of class $C^2$ in $x$ or positive definite for every $(x,t)$ and locally Lipschitz continuous. Let us assume, moreover, that, for every $t \in [0,T]$,

$$|a(x,t)| \leq M^2(1 + |x|^2)\,.$$

Then on the canonical space $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t^s)_{t\geq s}, (X_t)_{t\geq 0})$ there exists a unique family of probabilities $(\mathrm{P}^{x,s})_{x,s}$ such that $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t^s)_{t\geq s}, (X_t)_{t\geq s}, (\mathrm{P}^{x,s})_{x,s})$ is the realization of a diffusion process associated to $L_t$. Moreover, it is a Feller process.

*Proof* If $\sigma$ is the symmetric square root of the matrix field $a$, then the coefficients $b, \sigma$ satisfy Assumption (A') and if $\xi^{x,s}$ denotes the solution of

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\,dt + \sigma(\xi_t, t)\,dB_t \\ \xi_s &= x \end{aligned} \tag{9.42}$$

then the probabilities $\mathrm{P}^{x,s}$ on $(\mathscr{C}, \mathscr{M})$ introduced p. 279 satisfy the conditions required.

Uniqueness requires more attention. We must prove that if $\widetilde{\sigma}$ is any matrix field such that $a(x,t) = \widetilde{\sigma}(x,t)\widetilde{\sigma}(x,t)^*$ and $\widetilde{\xi}$ is a solution of

$$\begin{aligned} d\widetilde{\xi}_t &= b(\widetilde{\xi}_t, t)\,dt + \widetilde{\sigma}(\widetilde{\xi}_t, t)\,dB_t \\ \widetilde{\xi}_s &= x \end{aligned} \tag{9.43}$$

then the laws of $\xi^{x,s}$ and $\widetilde{\xi}$ coincide. If $a$ is elliptic, so that $a(x,t)$ is invertible for every $x, t$, and $\widetilde{\sigma}$ is another $m \times m$ matrix field such that $\widetilde{\sigma}(x,t)\widetilde{\sigma}(x,t)^* = a(x,t)$ for every $x, t$, then there exists an orthogonal matrix field $\rho(x,t)$, such that $\widetilde{\sigma} = \sigma\rho$ (just set $\rho = \sigma^{-1}\widetilde{\sigma}$ and use the fact that $a$ and its symmetric square root $\sigma$ commute). Hence $\widetilde{\xi}$ is a solution of

$$d\widetilde{\xi}_t = b(\widetilde{\xi}_t, t)\,dt + \sigma(\widetilde{\xi}_t, t)\rho(\widetilde{\xi}_t, t)\,dB_t = b(\widetilde{\xi}_t, t)\,dt + \sigma(\widetilde{\xi}_t, t)\,d\widetilde{B}_t\,,$$

where

$$\widetilde{B}_t = \int_0^t \rho(\widetilde{\xi}_s, s)\,dB_s$$

is once again a Brownian motion by Proposition 8.8. By Theorem 9.5 (uniqueness in law under Assumption (A')) $\xi$ and $\widetilde{\xi}$ have the same law.

This gives the main idea of the proof, but we must take care of other possible situations. What if $a$ is not elliptic? And if $\widetilde{\sigma}$ was an $m \times d$ matrix field with $d \neq m$? The first situation is taken care of by the following statement.

**Lemma 9.4** Let $A$ be an $m \times m$ semi-positive definite matrix and $S_1$, $S_2$ be two $m \times m$ matrices such that

$$A = S_i S_i^*\ .$$

Then there exists an orthogonal matrix $Q$ such that $S_1 = S_2 Q$.

*Proof* There exists an orthogonal matrix $O$ such that $A = O^* D O$, where $D$ is diagonal and with non-negative entries. Then if $R_i = OS_i$, we have

$$R_i R_i^* = O S_i S_i^* O^* = O O^* D O O^* = D\ . \tag{9.44}$$

Suppose $A$ has $m'$ non-zero eigenvalues, ordered in such a way that

$$D = \begin{pmatrix} \widetilde{D} & 0 \\ 0 & 0 \end{pmatrix}$$

for an $m' \times m'$ diagonal matrix $\widetilde{D}$ with strictly positive entries in its diagonal. Let $e_j$ be the vector having 1 as its $j$-th coordinate and 0 for the others. Then if $j > m'$

$$0 = \langle De_j, e_j \rangle = \langle R_i R_i^* e_j, e_j \rangle = |R_i^* e_j|^2 ,$$

which implies that the last $m - m'$ columns of $R_i^*$ vanish. Hence the last $m - m'$ rows of $R_i$ vanish and $R_i$ is of the form

$$R_i = \begin{pmatrix} R_i' \\ 0 \end{pmatrix} .$$

Clearly $R_i' R_i'^* = \widetilde{D}$ so that if $\widetilde{R}_i = \widetilde{D}^{-1/2} R_i'$ we have

$$\widetilde{R}_i \widetilde{R}_i^* = \widetilde{D}^{-1/2} R_i' R_i'^* \widetilde{D}^{-1/2} = I .$$

Hence the rows of $\widetilde{R}_i$ are orthonormal. Let $\widehat{R}_i$ be any $(m - m') \times m$ matrix whose rows complete to an orthonormal basis of $\mathbb{R}^m$. Then the matrices

$$\check{R}_i = \begin{pmatrix} \widetilde{R}_i \\ \widehat{R}_i \end{pmatrix}$$

are orthogonal and there exists an orthogonal matrix $Q$ such that

$$\check{R}_1 = \check{R}_2 Q .$$

From this relation, "going backwards" we obtain $\widetilde{R}_1 = \widetilde{R}_2 Q$, then $R_1 = R_2 Q$ and finally $S_1 = S_2 Q$. □

*End of the Proof of Theorem 9.10* Let us now consider the case of an $m \times d$-matrix field $\widetilde{\sigma}$ such that $\widetilde{\sigma}_2 \widetilde{\sigma}_2^* = a(x,t)$ for every $x, t$, hence consider in (9.43) a $d$-dimensional Brownian motion possibly with $d \neq m$. Also in this case we have a diffusion process associated to the generator $L$ and we must prove that it has the same law as the solution of (9.42). The argument is not really different from those developed above.

Let us assume first that $d > m$, i.e. the Brownian motion has a dimension that is strictly larger than the dimension of the diffusion $\xi$. Then $\widetilde{\sigma}(x,t)$ has a rank strictly smaller than $d$. Let us construct a $d \times d$ orthogonal matrix $\rho(x,t)$ by choosing its last $d - m$ columns to be orthogonal unitary vectors in $\ker \widetilde{\sigma}(x,t)$ and then completing the matrix by choosing the other columns so that the columns together form an orthogonal basis of $\mathbb{R}^d$. Then (9.43) can be written as

$$\begin{aligned} d\widetilde{\xi}_t &= b(\widetilde{\xi}_t, t)\,dt + \widetilde{\sigma}(\widetilde{\xi}_t, t)\rho(\widetilde{\xi}_t, t)\rho(\widetilde{\xi}_t, t)^{-1}\,dB_t \\ &= b(\widetilde{\xi}_t, t)\,dt + \widetilde{\sigma}_\rho(\xi_t, t)\,dW_t\ , \end{aligned}$$

where $\widetilde{\sigma}_\rho(x,t) = \widetilde{\sigma}(x,t)\rho(x,t)$ and $W_t = \int_0^t \rho(\widetilde{\xi}_s, s)^{-1}\,dB_s$ is a new $d$-dimensional Brownian motion. Now with the choice we have made of $\rho(x,t)$, it is clear that the last $d-m$ columns of $\sigma_\rho(x,t)$ vanish. Hence the previous equation can be written

$$d\widetilde{\xi}_t = b(\widetilde{\xi}_t, t)\,dt + \widetilde{\sigma}_2(\widetilde{\xi}_t, t)\,d\widetilde{W}_t\ ,$$

where $\widetilde{\sigma}_2(x,t)$ is the $m \times m$ matrix obtained by taking away from $\sigma_\rho(x,t)$ the last $d-m$ columns and $\widetilde{W}$ is the $m$-dimensional Brownian motion formed by the first $m$ components of $W$. It is immediate that $\underset{\sim}{\sigma_2}(x,t)\sigma_2(x,t)^* = a(x,t)$ and by the first part of the proof we know that the process $\underset{\sim}{\xi}$ has the same law as the solution of (9.42). It remains to consider the case where $\sigma_2$ is an $m \times d$ matrix with $d < m$ (i.e. the driving Brownian motion has a dimension that is smaller than the dimension of $\xi$), but this can be done easily using the same type of arguments developed so far, so we leave it to the reader.

Note that this proof is actually incomplete, as we should also prove that the orthogonal matrices $\rho(x,t)$ above can be chosen in such a way that the matrix field $(x,t) \to \rho(x,t)$ is locally Lipschitz continuous. □

## 9.10 Further reading

The theory of SDEs presented here does not account for other approaches to the question of existence and uniqueness.

Many deep results are known today concerning SDEs whose coefficients are not locally Lipschitz continuous. This is important because in the applications examples of SDEs of this type do arise.

An instance is the so-called square root process (or CIR model, as it is better known in applications to finance), which is the solution of the SDE

$$d\xi_t = (a - b\xi_t)\,dt + \sigma\sqrt{\xi_t}\,dB_t\ .$$

In Exercise 9.19 we see that if $a = \frac{\sigma^2}{4}$ then a solution exists and can be obtained as the square of an Ornstein–Uhlenbeck process. But what about uniqueness? And what about existence for general coefficients $a$, $b$ and $\sigma$?

For the problem of existence and uniqueness of the solutions of an SDE there are two approaches that the interested reader might look at.

The first one is the theory of linear (i.e. real-valued) diffusion as it is developed (among other places) in Revuz and Yor (1999, p. 300) or Ikeda and Watanabe (1981,

p. 446). The Revuz and Yor (1999) reference is also (very) good for many other advanced topics in stochastic calculus.

The second approach is the Strook–Varadhan theory that produces existence and uniqueness results under much weaker assumptions than local Lipschitz continuity and with a completely different and original approach. Good references are, without the pretension of being exhaustive, Stroock and Varadhan (1979), Priouret (1974) (in French), Rogers and Williams (2000) and Ikeda and Watanabe (1981).

## Exercises

**9.1** (p. 556) Let $\xi$ be the Ornstein–Uhlenbeck process solution of the SDE

$$\begin{aligned} d\xi_t &= -\lambda\xi_t\, dt + \sigma\, dB_t \\ \xi_0 &= x \end{aligned} \tag{9.45}$$

with $\lambda > 0$ (see Example 9.1).

a) Prove that, whatever the starting position $x$, the law of $\xi_t$ converges as $t \to +\infty$ to a probability $\mu$ to be determined.
b) Consider now the process $\xi$ but with a starting position $\xi_0 = \eta$, where $\eta$ is a square integrable r.v. having law $\mu$ and independent of the Brownian motion $B$. Show that $\xi_t$ has distribution $\mu$ for every $t$.

**9.2** (p. 557)

a) Let $B$ be a $d$-dimensional Brownian motion and let us consider the SDE, in dimension $m$,

$$\begin{aligned} d\xi_t &= b(t)\xi_t\, dt + \sigma(t)\, dB_t \\ \xi_0 &= x \end{aligned} \tag{9.46}$$

where $b$ and $\sigma$ are locally bounded measurable functions of the variable $t$ only, with values respectively in $M(m)$ and $M(m, d)$ ($m \times m$ and $m \times d$ matrices respectively). Find an explicit solution, show that it is a Gaussian process and compute its mean and covariance functions (see Example 2.5).
b) Let us consider, for $t < 1$, the SDE

$$\begin{aligned} d\xi_t &= -\frac{\xi_t}{1-t}\, dt + dB_t \\ \xi_0 &= x\,. \end{aligned} \tag{9.47}$$

Solve this equation explicitly. Compute the mean and covariance functions of $\xi$ (which is a Gaussian process by a)). Prove that if $x = 0$ then $\xi$ is a Brownian bridge.

**9.3** (p. 558) (The asymptotic of the Ornstein–Uhlenbeck processes) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a real Brownian motion and $\xi$ the Ornstein–Uhlenbeck process solution of the SDE, for $\lambda \neq 0, \sigma > 0$,

$$\begin{aligned} d\xi_t &= -\lambda \xi_t \, dt + \sigma \, dB_t \\ \xi_0 &= x \, . \end{aligned}$$

a) Let us assume $\lambda > 0$.

a1) Prove that

$$\varliminf_{t\to+\infty} \frac{\xi_t}{\sqrt{\log t}} = -\frac{\sigma}{\sqrt{\lambda}}, \qquad \varlimsup_{t\to+\infty} \frac{\xi_t}{\sqrt{\log t}} = \frac{\sigma}{\sqrt{\lambda}} \, \cdot$$

a2) What can be said of $\varliminf_{t\to+\infty} \xi_t$ and $\varlimsup_{t\to+\infty} \xi_t$?

b) Let $\lambda < 0$. Prove that there exists an event $A \in \mathscr{F}$ such that $\lim_{t\to+\infty} \xi_t = +\infty$ on $A$ and $\lim_{t\to+\infty} \xi_t = -\infty$ on $A^c$ with $\mathrm{P}(A) = \Phi(\frac{x\sqrt{-2\lambda}}{\sigma})$, where $\Phi$ denotes the partition function of an $N(0,1)$-distributed r.v. (in particular $\mathrm{P}(A) = \frac{1}{2}$ if $x = 0$).

**9.4** (p. 560) Let $\xi$ be the solution of the SDE, for $t < 1$,

$$\begin{aligned} d\xi_t &= -\frac{1}{2} \frac{\xi_t}{1-t} \, dt + \sqrt{1-t} \, dB_t \\ \xi_0 &= x \, . \end{aligned} \tag{9.48}$$

a) Find the solution of this equation and prove that it is a Gaussian process.

b) Compare the variance of $\xi_t$ with the corresponding variance of a Brownian bridge at time $t$. Is $\xi$ a Brownian bridge?

**9.5** (p. 561) Let $B$ be a real Brownian motion and $\xi$ the solution of the SDE

$$\begin{aligned} d\xi_t &= b(t)\xi_t \, dt + \sigma(t)\xi_t \, dB_t \\ \xi_0 &= x > 0 \, , \end{aligned} \tag{9.49}$$

where $b$ and $\sigma$ are measurable and locally bounded functions of the time only.

a) Find an explicit solution of (9.49).

b) Investigate the a.s. convergence of $\xi$ as $t \to +\infty$ when

$$\begin{array}{llll} (1) & \sigma(t) = \dfrac{1}{1+t} & b(t) = \dfrac{2+t}{2(1+t)^2} \\ (2) & \sigma(t) = \dfrac{1}{1+t} & b(t) = \dfrac{1}{3(1+t)^2} \\ (3) & \sigma(t) = \dfrac{1}{\sqrt{1+t}} & b(t) = \dfrac{1}{2(1+t)} \, \cdot \end{array}$$

**9.6** (p. 562) Let $b \in \mathbb{R}^m$, $\sigma$ an $m \times d$ matrix, $B$ a $d$-dimensional Brownian motion. Let us consider the SDE

$$\begin{aligned} d\xi_i(t) &= b_i\xi_i(t)\,dt + \xi_i(t)\sum_{j=1}^{d}\sigma_{ij}\,dB_j(t), \quad i = 1, \dots, m \\ \xi_0 &= x\,. \end{aligned} \tag{9.50}$$

a) Find an explicit solution. Prove that if $x_i > 0$ then $\mathrm{P}(\xi_i(t) > 0 \text{ for every } t > 0) = 1$ and compute $\mathrm{E}[\xi_i(t)]$.
b) Prove that the processes $t \mapsto \xi_i(t)\xi_j(t)$ are in $M^2$ and compute $\mathrm{E}[\xi_i(t)\xi_j(t))]$ and the covariances $\mathrm{Cov}(\xi_i(t), \xi_j(t))$.

**9.7** (p. 564) Let $B$ be a two-dimensional Brownian motion and let us consider the two processes

$$\begin{aligned} d\xi_1(t) &= r_1\xi_1(t)\,dt + \sigma_1\xi_1(t)\,dB_1(t) \\ d\xi_2(t) &= r_2\xi_2(t)\,dt + \sigma_2\xi_2(t)\,dB_2(t) \end{aligned}$$

where $r_1, r_2, \sigma_1, \sigma_2$ are real numbers.

a) Prove that both processes $X_t = \xi_1(t)\xi_2(t)$ and $Z_t = \sqrt{\xi_1(t)\xi_2(t)}$ are solutions of SDEs to be determined.
b) Answer the same questions as in a) if we had

$$\begin{aligned} d\xi_1(t) &= r_1\xi_1(t)\,dt + \sigma_1\xi_1(t)\,dB_1(t) \\ d\xi_2(t) &= r_2\xi_2(t)\,dt + \sigma_2\sqrt{1-\rho^2}\,\xi_2(t)\,dB_2(t) + \sigma_2\rho\xi_2(t)\,dB_1(t)\,, \end{aligned}$$

where $-1 \le \rho \le 1$.

**9.8** (p. 565) Let $\xi$ be the solution (geometric Brownian motion) of the SDE

$$\begin{aligned} d\xi_t &= b\xi_t\,dt + \sigma\xi_t\,dB_t \\ \xi_0 &= 1\,. \end{aligned}$$

a) Determine a real number $\alpha$ such that $(\xi_t^\alpha)_t$ is a martingale.
b) Let $\tau$ be the exit time of $\xi$ out of the interval $]\frac{1}{2}, 2[$. Compute $\mathrm{P}(\xi_\tau = 2)$.

**9.9** (p. 566) Let $B$ be a two-dimensional Brownian motion and, for $\nu \in \mathbb{R}$, let us consider the SDE

$$\begin{aligned} d\xi_t &= \eta_t\,dB_1(t) \\ d\eta_t &= \nu\eta_t\,dt + \eta_t\,dB_2(t) \\ \xi_0 &= x, \eta_0 = y > 0\,. \end{aligned}$$

a) Find an explicit expression for $(\eta_t)_t$ and $(\xi_t)_t$. Is $(\xi_t)_t$ a martingale?
b) Compute $\mathrm{E}(\xi_t)$ and $\mathrm{Var}(\xi_t)$.
c) Prove that, if $\nu < -\frac{1}{2}$, the limit

$$\lim_{t\to 0+} \xi_t$$

exists a.s. and in $L^2$.

**9.10** (p. 567) Let us consider the one-dimensional SDE

$$\begin{aligned} d\xi_t &= (a + b\xi_t)\, dt + (\lambda + \sigma\xi_t)\, dB_t \\ \xi_0 &= x\,. \end{aligned} \tag{9.51}$$

a) Let $v(t) = \mathrm{E}[\xi_t]$. Show that $v$ satisfies an ordinary differential equation and compute $\mathrm{E}[\xi_t]$.
b1) Prove that if $b < 0$, then $\lim_{t\to+\infty} \mathrm{E}[\xi_t] = -\frac{a}{b}$.
b2) Prove that if $x = -\frac{a}{b}$ then the expectation is constant and $\mathrm{E}[\xi_t] \equiv x$ whatever the value of $b$.
b3) Assume $b > 0$. Prove that

$$\lim_{t\to+\infty} \mathrm{E}[\xi_t] = \begin{cases} -\infty & \text{if } x_0 < -\frac{a}{b} \\ +\infty & \text{if } x_0 > -\frac{a}{b}\,. \end{cases}$$

**9.11** (p. 568) Let us consider the one-dimensional SDE

$$\begin{aligned} d\xi_t &= (a + b\xi_t)\, dt + (\lambda + \sigma\xi_t)\, dB_t \\ \xi_0 &= x\,. \end{aligned} \tag{9.52}$$

Recall (Example 9.2) that if $a = \lambda = 0$ then a solution is given by $\xi_0(t)x$, where $\xi_0(t) = \mathrm{e}^{(b-\frac{\sigma^2}{2})t+\sigma B_t}$.

a1) Determine a process $C$ such that $t \mapsto \xi_0(t)C_t$ is a solution of (9.52) and produce explicitly the solution of (9.52).
a2) Prove that if $\sigma = 0$ then $\xi$ is a Gaussian process and compute $\mathrm{E}[\xi_t]$, $\mathrm{Var}(\xi_t)$.
a3) Prove that, if $\sigma = 0$ and $b < 0$, $\xi$ converges in law as $t \to +\infty$ to a limit distribution to be determined.
b) We want to investigate the equation

$$\begin{aligned} dY_t &= (b + \theta \log Y_t) Y_t\, dt + \sigma Y_t\, dB_t \\ Y_0 &= y > 0\,. \end{aligned} \tag{9.53}$$

Show that $\xi_t = \log Y_t$ is the solution of an SDE to be determined. Show that (9.53) has a unique solution such that $Y_t > 0$ for every $t$ a.s. What is the law of $Y_t$? Prove that, if $\theta < 0$, $(Y_t)_t$ converges in law as $t \to +\infty$ and determine the limit distribution.

**9.12** (p. 570) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a one-dimensional Brownian motion and let $\xi$ be the Ornstein–Uhlenbeck process that is the solution of the SDE

$$\begin{aligned} d\xi_t &= \lambda \xi_t \, dt + \sigma dB_t \\ \xi_0 &= x \,, \end{aligned}$$

where $\lambda \neq 0$.

a) Show that

$$Z_t = \mathrm{e}^{-2\lambda t} \Big( \xi_t^2 + \frac{\sigma^2}{2\lambda} \Big)$$

is an $(\mathscr{F}_t)_t$-martingale.

b) Let $Y_t = \xi_t^2 + \frac{\sigma^2}{2\lambda}$. What is the value of $\mathrm{E}(Y_t)$? And of $\lim_{t\to\infty} \mathrm{E}(Y_t)$? (It will be useful to distinguish the cases $\lambda > 0$ and $\lambda < 0$.)

**9.13** (p. 571) Let us consider the one-dimensional SDE

$$\begin{aligned} d\xi_t &= \Big( \sqrt{1+\xi_t^2} + \frac{1}{2}\,\xi_t \Big)\, dt + \sqrt{1+\xi_t^2}\, dB_t \\ \xi_0 &= x \in \mathbb{R} \,. \end{aligned} \tag{9.54}$$

a) Does this equation admit strong solutions?

b1) Let

$$Y_t = \log \Big( \sqrt{1+\xi_t^2} + \xi_t \Big) \,.$$

What is the stochastic differential of $Y$?

b2) Deduce an explicit solution of (9.54).

Hint: $z \mapsto \log \big( \sqrt{1+z^2} + z \big)$ is the inverse function of $y \mapsto \sinh y$.

**9.14** (p. 572) Let $(\xi_t, \eta_t)$ be the solution of the SDE

$$\begin{aligned} d\xi_t &= -\lambda \xi_t \, dt + \sigma dB_1(t) \\ d\eta_t &= -\lambda \eta_t \, dt + \sigma\rho \, dB_1(t) + \sigma \sqrt{1-\rho^2}\, dB_2(t) \end{aligned}$$

with the initial conditions $(\xi_0, \eta_0) = (x, y)$, where $B = (B_1, B_2)$ is a two-dimensional Brownian motion and $-1 \le \rho \le 1$.

a) Compute the laws of $\xi_t$ and of $\eta_t$ and explicitly describe their dependence on the parameter $\rho$.
b) Compute the joint law of $(\xi_t, \eta_t)$. What is the value of $\mathrm{Cov}(\xi_t, \eta_t)$? For which values of $\rho$ is this covariance maximum? For which values of $\rho$ does the law of the pair $(\xi_t, \eta_t)$ have a density with respect to Lebesgue measure?
c) What is the differential generator of the diffusion $(\xi_t, \eta_t)$?

**9.15** (p. 573) Let $B$ be a Brownian motion. Let $\xi_t = (\xi_1(t), \xi_2(t))$ be the diffusion process that is the solution of the SDE

$$\begin{aligned} d\xi_1(t) &= -\tfrac{1}{2}\xi_1(t)\,dt - \xi_2(t)\,dB_t \\ d\xi_2(t) &= -\tfrac{1}{2}\xi_2(t)\,dt + \xi_1(t)\,dB_t \end{aligned}$$

with the initial condition $\xi_1(0) = 0, \xi_2(0) = 1$.

a) What is the generator of $(\xi_t)_t$? Is it elliptic? Uniformly elliptic?
b) Let $Y_t = \xi_1(t)^2 + \xi_2(t)^2$. Show that $(Y_t)_t$ satisfies an SDE and determine it. What can be said of the law of $Y_1$?

**9.16** (p. 574) Let $\mu \in \mathbb{R}$, $\sigma > 0$ and

$$\begin{aligned} X_t &= x + \int_0^t \mathrm{e}^{(\mu - \frac{\sigma^2}{2})u + \sigma B_u}\,du \\ Y_t &= \mathrm{e}^{(\mu - \frac{\sigma^2}{2})t + \sigma B_t}\;. \end{aligned}$$

a) Prove that the two-dimensional process $Z_t = (X_t, Y_t)$ is a diffusion and compute its generator.
b) Prove that

$$\xi_t = \frac{X_t}{Y_t}$$

is a (one-dimensional) diffusion which is the solution of an SDE to be determined.
c) Let $r \in \mathbb{R}$, $\sigma > 0$. Find an explicit solution of the SDE

$$\begin{aligned} dZ_t &= (1 + rZ_t)\,dt + \sigma Z_t\,dB_t \\ Z_0 &= z \in \mathbb{R} \end{aligned} \tag{9.55}$$

and compare with the one obtained in Exercise 9.11. Determine the value of $\mathrm{E}[Z_t]$ and its limit as $t \to +\infty$ according to the values of $r, \sigma, z$.

**9.17** (p. 576) (Do Exercise 9.11 first) Let us consider the SDE

$$\begin{aligned} d\xi_t &= \xi_t(a - b\xi_t)\,dt + \sigma\xi_t\,dB_t \\ \xi_0 &= x > 0\ , \end{aligned} \tag{9.56}$$

where $b > 0$. Note that the conditions for the existence of solutions are not satisfied, as the drift does not have a sublinear growth at infinity.

a) Compute, at least formally, the stochastic differential of $Z_t = \frac{1}{\xi_t}$ and determine an SDE satisfied by $Z$.
b) Write down the solution of the SDE for $Z$ (using Exercise 9.11) and derive a solution of (9.56). Prove that $\xi_t > 0$ for every $t \geq 0$.

**9.18** (p. 577) Let $\xi$ be the solution, if it exists, of the equation

$$\begin{aligned} d\xi_t &= \gamma\Big(\int_0^t \xi_s\,ds\Big)\,dt + \sigma\,dB_t \\ \xi_0 &= x\ , \end{aligned}$$

where $\gamma, x \in \mathbb{R}$, $\sigma > 0$. This is not an SDE, as the drift depends not only on the value of the position of $\xi$ at time $t$ but also on its entire past behavior.

a) Prove that if

$$\eta_t = \int_0^t \xi_s\,ds$$

then the two-dimensional process $Z_t = (\xi_t, \eta_t)$ is a diffusion. Write down its generator and prove that $\xi$ is a Gaussian process.
b) Prove (or take as granted) that if

$$M = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

then

$$e^M = \begin{pmatrix} \cosh\sqrt{ab} & \sqrt{\frac{a}{b}}\sinh\sqrt{ab} \\ \sqrt{\frac{b}{a}}\sinh\sqrt{ab} & \cosh\sqrt{ab} \end{pmatrix}$$

which of course means, if $ab < 0$,

$$e^M = \begin{pmatrix} \cos\sqrt{-ab} & -\sqrt{-\frac{a}{b}}\sin\sqrt{-ab} \\ -\sqrt{-\frac{b}{a}}\sin\sqrt{-ab} & \cos\sqrt{-ab} \end{pmatrix}\ . \tag{9.57}$$

c) Compute the mean and variance of the distribution of $\xi_t$ and determine their behavior as $t \to +\infty$, according to the values of $\gamma, \sigma$.

**9.19** (p. 579) (Have a look at Example 8.9 first) Let $B$ be a Brownian motion and $\xi$ the Ornstein–Uhlenbeck process solution of the SDE

$$\begin{aligned} d\xi_t &= b\xi_t\,dt + \sigma\,dB_t \\ \xi_0 &= x\,, \end{aligned}$$

where $b, \sigma \in \mathbb{R}$. Let $\eta_t = \xi_t^2$.

Prove that $\eta$ is the solution of an SDE to be determined.

**9.20** (p. 580) Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a function satisfying hypothesis (A'). Moreover, let us assume that $\sigma(x) \ge c > 0$ for every $x \in \mathbb{R}$ and let $\xi$ be the solution of the SDE

$$d\xi_t = \sigma(\xi_t)\,dB_t, \qquad \xi_0 = 0\,.$$

Let $I = [-a, b]$, where $a, b > 0$, and let $\tau$ be the exit time from $I$. Show that $\tau < +\infty$ a.s. and compute $\mathrm{P}(\xi_\tau = b)$.

**9.21** (p. 581) Let $a, b > 0$ and let $\xi$ be the Ornstein–Uhlenbeck process that is the solution of

$$\begin{aligned} d\xi_t &= -\lambda\xi_t\,dt + \sigma\,dB_t \\ \xi_0 &= x\,. \end{aligned}$$

a) Prove that the exit time, $\tau$, of $\xi$ from $]-a, b[$ is finite a.s. for every starting position $x$ and give an expression of $\mathrm{P}(\xi_\tau = b)$ as a function of $\sigma, \lambda, a, b$.

b1) Prove that if $a > b$ then

$$\lim_{\lambda \to +\infty} \frac{\displaystyle\int_0^a \mathrm{e}^{\lambda z^2}\,dz}{\displaystyle\int_0^b \mathrm{e}^{\lambda z^2}\,dz} = +\infty\,. \tag{9.58}$$

b2) Assume $a > b$. Prove that

$$\lim_{\lambda \to +\infty} \mathrm{P}^x(\xi_\tau = b) = 1$$

*whatever* the starting point $x \in ]-a, b[$.

b3) In b2) it appears that, even if $x$ is close to $-a$, the exit from $]-a, b[$ takes place mostly at $b$ if $\lambda$ is large. Would you be able to give an intuitive explanation of this phenomenon?

**9.22** (p. 582) Let $\xi^\varepsilon$ be the Ornstein–Uhlenbeck process solution of the SDE

$$\begin{aligned} d\xi_t^\varepsilon &= -\lambda \xi_t^\varepsilon \, dt + \varepsilon\sigma \, dB_t \\ \xi_0^\varepsilon &= x \, , \end{aligned}$$

where $\lambda \in \mathbb{R}$, $\sigma > 0$.

a) Prove that, for every $t \geq 0$

$$\xi_t^\varepsilon \quad \overset{\mathscr{L}}{\underset{\varepsilon\to 0}{\to}} \quad \mathrm{e}^{-\lambda t} x \, .$$

b) Prove that the laws of the processes $\xi^\varepsilon$ (which, remember, are probabilities on the space $\mathscr{C} = \mathscr{C}([0, T], \mathbb{R})$) converge in distribution to the Dirac mass concentrated on the path $x_0(t) = \mathrm{e}^{-\lambda t} x$ that is the solution of the ordinary equation

$$\begin{aligned} \dot{\xi}_t &= -\lambda \xi_t \\ \xi_0 &= x \, . \end{aligned} \tag{9.59}$$

In other words, the diffusion $\xi^\varepsilon$ can be seen as a small random perturbation of the ODE (9.59). See Exercise 9.28, where the general situation is considered.

**9.23** (p. 583) (Example of explosion). Let us consider the SDE, in dimension 1,

$$\begin{aligned} d\xi_t &= \xi_t^3 \, dt + \xi_t^2 \, dB_t \\ \xi_0 &= x \end{aligned} \tag{9.60}$$

and let $\tau_x = \inf\{t; B_t = \frac{1}{x}\}$ for $x \neq 0$, $\tau_0 = +\infty$. Prove that, for $t \in [0, \tau_x[$, a solution is given by

$$\xi_t = \frac{x}{1 - xB_t} \, \cdot$$

**9.24** (p. 584) Let $L$ be the operator in dimension 2

$$L = \frac{1}{2} \frac{\partial^2}{\partial x^2} + x \frac{\partial}{\partial y} \tag{9.61}$$

and $\xi$ the diffusion having $L$ as its generator.

a) Is $L$ elliptic? Uniformly elliptic? Compute the law of $\xi_t$ with the starting condition $\xi_0 = x$. Does it have a density with respect to the Lebesgue measure of $\mathbb{R}^2$?

b) Answer the same questions as in a) for the operator

$$L_2 = \frac{1}{2}\frac{\partial^2}{\partial x^2} + y\frac{\partial}{\partial y} \cdot \tag{9.62}$$

**9.25** (p. 585) (Continuation of Exercise 9.24) Let us consider the diffusion associated to the generator

$$L = \frac{1}{2}\sum_{i,j=1}^{m} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i,j=1}^{m} b_{ij}x_j\frac{\partial}{\partial x_i}\, , \tag{9.63}$$

where $a$ and $b$ are constant $m \times m$ matrices.

a) Compute the transition function and show that, if the matrix $a$ is positive definite, then the transition function has a density with respect to Lebesgue measure.
b) Show that for the operator $L$ in (9.63) the transition function has a density if and only if the kernel of $a$ does not contain subspaces different from $\{0\}$ that are invariant for the matrix $b^*$.

**9.26** (p. 586) Let $\xi$ be the solution of the SDE in dimension 1

$$\begin{aligned} d\xi_t &= b(\xi_t)\, dt + \sigma(\xi_t)\, dB_t \\ \xi_0 &= x\, . \end{aligned}$$

Assume that $b$ is locally Lipschitz continuous, $\sigma$ is $C^1$ and such that $\sigma(x) \geq \delta > 0$ for some $\delta$ and that $b$ and $\sigma$ have a sublinear growth at infinity.

a) Prove that there exists a strictly increasing function $f : \mathbb{R} \to \mathbb{R}$ such that the process $\eta_t = f(\xi_t)$ satisfies the SDE

$$d\eta_t = \widetilde{b}(\eta_t)\, dt + dB_t \tag{9.64}$$

for some new drift $\widetilde{b}$.
b) Prove that, under the assumptions above, the filtration $\mathscr{H}_t = \sigma(\xi_u, u \leq t)$ coincides with the natural filtration of the Brownian motion.

**9.27** (p. 587) (Doss 1977; Sussmann 1978). In this exercise we shall see that, in dimension 1, the solution of an SDE can be obtained by solving an ordinary differential equation with random coefficients. It is a method that can allow us to find solutions explicitly and to deduce useful properties of the solutions.

Let us assume $m = 1$. Let $b$ and $\sigma$ be Lipschitz continuous functions $\mathbb{R} \to \mathbb{R}$ and let us assume that $\sigma$ is twice differentiable and strictly positive. Let us denote by $h(x, y)$ the solution of

$$\begin{aligned} u'(y) &= \sigma(u(y)) \\ u(0) &= x\,. \end{aligned} \tag{9.65}$$

Let $t \mapsto D_t(\omega)$ be the solution of the ODE with random coefficients

$$D_t' = f(D_t, B_t(\omega)), \qquad D_0 = x\,,$$

where

$$f(x, z) = \Big[-\frac{1}{2}(\sigma'\sigma)(h(x, z)) + b(h(x, z))\Big] \exp\Big(-\int_0^z \sigma'(h(x, s))\, ds\Big)\,.$$

a) Prove that $\xi_t = h(D_t, B_t)$ is the solution of

$$\begin{aligned} d\xi_t &= b(\xi_t)\, dt + \sigma(\xi_t)\, dB_t \\ \xi_0 &= x\,. \end{aligned} \tag{9.66}$$

b) Let $\widetilde{b}$ be a Lipschitz continuous function such that $\widetilde{b}(x) \geq b(x)$ for every $x \in \mathbb{R}$ and let $\widetilde{\xi}$ be the solution of

$$\begin{aligned} d\widetilde{\xi}_t &= \widetilde{b}(\xi_t)\, dt + \sigma(\xi_t)\, dB_t \\ \widetilde{\xi}_0 &= \widetilde{x}\,. \end{aligned} \tag{9.67}$$

with $\widetilde{x} \geq x$. Note that the two SDEs (9.66) and (9.67) have the same diffusion coefficient and are with respect to the same Brownian motion. Show that $\widetilde{\xi}_t \geq \xi_t$ for every $t \geq 0$ a.s.

a) Thanks to differentiability results for the solutions of an ODE with respect to a parameter, $h$ is twice differentiable in every variable. Giving this fact as granted, prove that

$$\frac{\partial h}{\partial x}(x, y) = \exp\Big[\int_0^y \sigma'(h(x, s))\, ds\Big], \qquad \frac{\partial^2 h}{\partial y^2}(x, y) = \sigma'(h(x, y))\sigma(h(x, y)) \tag{9.68}$$

and apply Ito's formula to $t \mapsto h(D_t, B_t)$.

b) Recall the following comparison theorem for ODEs: if $g_1, g_2$ are Lipschitz continuous functions $\mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ such that $g_2(x, t) \geq g_1(x, t)$ for every $x, t$ and $\gamma_i, i = 1, 2$ are solutions of the ODEs

$$\gamma_i'(t) = g_i(\gamma_i(t), t)$$

$$\gamma_i(s) = x_i$$

with $x_2 \geq x_1$, then $\gamma_2(t) \geq \gamma_1(t)$ for every $t \geq 0$.

**9.28** (p. 588) Let $\xi^\varepsilon$ be the solution of the SDE, in dimension $m$,

$$\begin{aligned} d\xi_t^\varepsilon &= b(\xi_t^\varepsilon)\,dt + \varepsilon\sigma(\xi_t^\varepsilon)\,dB_t \\ \xi_0^\varepsilon &= x\,, \end{aligned}$$

where $b$ and $\sigma$ are Lipschitz continuous with Lipschitz constant $L$ and $\sigma$ is bounded. Intuitively, if $\varepsilon$ is small, this equation can be seen as a small random perturbation of the ordinary equation

$$\begin{aligned} \gamma_t' &= b(\gamma_t) \\ \gamma_0 &= x\,. \end{aligned} \tag{9.69}$$

In this exercise we see that $\xi^\varepsilon$ converges in probability as $\varepsilon \to 0$ to the path $\gamma$ that is the solution of (9.69) (and with an estimate of the speed of convergence).

a) If $\eta_t^\varepsilon = \xi_t^\varepsilon - \gamma_t$, show that $\eta^\varepsilon$ is a solution of

$$\begin{aligned} d\eta_t^\varepsilon &= [b(\eta_t^\varepsilon + \gamma_t) - b(\gamma_t)]\,dt + \varepsilon\sigma(\eta_t^\varepsilon + \gamma_t)\,dB_t \\ \eta_0^\varepsilon &= 0\,. \end{aligned}$$

b) Prove that, for $T > 0, \alpha > 0$,

$$\Big\{\sup_{0\le s\le T}\Big|\varepsilon\int_0^s \sigma(\eta_u^\varepsilon + \gamma_u)\,dB_u\Big| < \alpha\Big\} \subset \Big\{\sup_{0\le s\le T}|\eta_s^\varepsilon| < \alpha e^{LT}\Big\}$$

and therefore

$$\mathrm{P}\Big(\sup_{0\le s\le T}|\xi_s^\varepsilon - \gamma_s| > \alpha\Big) \le 2m\exp\Big[-\frac{1}{\varepsilon^2}\,\frac{\alpha^2 e^{-2LT}}{2mT\|\sigma\|_\infty^2}\Big]\,.$$

Deduce that $\xi^\varepsilon : \Omega \to \mathscr{C}([0,T],\mathbb{R}^m)$ converges in probability (and therefore in law) as $\varepsilon \to 0$ to the r.v. concentrated on the path $\gamma$ a.s.

c) Are you able to weaken the hypothesis and show that the same result holds if $b$ and $\sigma$ satisfy only Assumption (A')?

b) Use Gronwall's inequality and the exponential inequality (8.42). c) Localization...

**9.29** (p. 589) We have seen in Sect. 9.3 some $L^p$ estimates for the solutions of a SDE, under Assumption (A'). In this exercise we go deeper into the investigation of the *tail* of the law of the solutions. Let $B$ be a $d$-dimensional Brownian motion and, for a given process $Z$, let us denote $\sup_{0\le s\le t}|Z_s|$ by $Z_t^*$.

a) Let $X$ be the $m$-dimensional process

$$X_t = x + \int_0^t F_s\,ds + \int_0^t G_s\,dB_s$$

where $F$ and $G$ are processes in $M^1([0,T])$ and $M^2([0,T])$, respectively, that are $m$- and $m \times d$-dimensional, respectively; let us assume, moreover, that $G$ is bounded and that the inequality

$$|F_t| \le M(1 + |X_t|)$$

holds (which is satisfied, in particular, if $F$ is bounded). Then for every $T > 0$, $K > 0$, there exist $c = c_T > 0$ and $R_0 > 0$ such that, for $R > R_0$ and for every $x$ such that $|x| \le K$,

$$\mathrm{P}(X_T^* > R) \le \mathrm{e}^{-cR^2}\ . \tag{9.70}$$

b) Let us assume that $b$ and $\sigma$ satisfy Assumption (A') and that, moreover, the $m \times d$ matrix $\sigma$ is bounded. Let $\xi$ be the solution of

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\,dt + \sigma(\xi_t, t)\,dB_t \\ \xi_u &= x\ . \end{aligned}$$

Then for every $T > 0$ there exists a constant $c = c_T > 0$ such that for large $R$

$$\mathrm{P}(\xi_T^* > R) \le \mathrm{e}^{-cR^2}$$

uniformly for $x$ in a compact set.

c) Let us assume again that $b$ and $\sigma$ satisfy Assumption (A') and drop the assumption of boundedness for $\sigma$. Show that $Y_t = \log(1 + |\xi_t|^2)$ is an Ito process and compute its stochastic differential. Deduce that, for every $T > 0$, there exists a constant $c = c_T > 0$ such that for large $R$

$$\mathrm{P}(\xi_T^* > R) \le \frac{1}{R^{c\log R}}$$

uniformly for $x$ in a compact set.

d) Show that $\xi_T^* \in L^p$ for every $p$ (this we already know thanks to Theorem 9.1). Moreover, if $\sigma$ is bounded, all exponential moments $\xi_T^*$ are finite (i.e. $\mathrm{e}^{\lambda \xi_T^*}$ is integrable for every $\lambda \in \mathbb{R}$) and even

$$\mathrm{E}[\mathrm{e}^{\alpha \xi_T^{*2}}] < +\infty$$

for $\alpha < c^*$ for some constant $c^* > 0$.

a) Use the exponential inequality, Proposition 8.7, in order to estimate $\mathrm{P}(\sup_{0 \le t \le T} |\int_0^t G_s\,dB_s| > \rho)$ and then Gronwall's inequality. d) Recall the formula of Exercise 1.3.

# Chapter 10
# PDE Problems and Diffusions

In this chapter we see that the solutions of some PDE problems can be represented as expectations of functionals of diffusion process. These formulas are very useful from two points of view. First of all, for the investigation and a better understanding of the properties of the solutions of these PDEs. Moreover, in some situations, they allow to compute the solution of the PDE (through the explicit computation of the expectation of the corresponding functional) or the expectation of the functional (by solving the PDE explicitly). The exercises of this chapter and Exercise 12.8 provide some instances of this way of reasoning.

## 10.1 Representation of the solutions of a PDE problem

Let $D$ be a domain of $\mathbb{R}^m$ and $u$ a solution of the Dirichlet problem

$$\begin{cases} Lu = 0 & \text{on } D \\ u_{|\partial D} = \phi\ , \end{cases} \tag{10.1}$$

where $L$ is a second-order differential operator and $\phi$ a function $\partial D \to \mathbb{R}$. Without bothering about regularity assumptions, let $B$ be a Brownian motion and $\xi^x$ the diffusion associated to $L$ and starting at $x \in D$ obtained as seen in the previous chapter by solving an SDE with respect to the Brownian motion $B$. Then by Ito's formula applied to the function $u$ (that we assume to exist and to be regular enough) we have

$$u(\xi_t^x) = u(x) + \int_0^t Lu(\xi_s^x)\,ds + \int_0^t u'(\xi_s^x)\sigma(X_s)\,dB_s\ ,$$

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_10

$u'$ denoting the gradient of $u$ (recall Remark 9.1). If $\tau$ denotes the exit time of $\xi^x$ from $D$, then

$$u(\xi^x_{t\wedge\tau}) = u(x) + \int_0^{t\wedge\tau} \underbrace{Lu(\xi^x_s)}_{=0}\, ds + \int_0^{t\wedge\tau} u'(\xi^x_s)\sigma(\xi^x_s)\, dB_s$$

and if the process $t \mapsto u'(\xi^x_t)$ belongs to $M^2$, taking the expectation we find

$$\mathrm{E}[u(\xi^x_{t\wedge\tau})] = u(x)\ .$$

As $\xi^x_\tau \in \partial D$ and $u = \phi$ on $\partial D$, taking the limit as $t \to \infty$ we find finally, if $\tau$ is finite and $u$ bounded,

$$u(x) = \mathrm{E}[u(\xi^x_\tau)] = \mathrm{E}[\phi(\xi^x_\tau)]\ , \tag{10.2}$$

i.e. the value of $u$ at $x$ is equal to the mean of the boundary condition $\phi$ taken with respect to the exit distribution from the domain $D$ of the diffusion associated to $L$ and starting at $x$. Note that, denoting by $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_{t\geq 0}, (X_t)_{t\geq 0}, (\mathrm{P}^x)_x)$ the canonical diffusion associated to $L$, (10.2) can also be written as

$$u(x) = \mathrm{E}^x[\phi(X_\tau)]\ . \tag{10.3}$$

The idea is therefore really simple. There are, however, a few things to check in order to put together a rigorous proof. More precisely:

a) We need conditions guaranteeing that the exit time from $D$ is a.s. finite. We shall also need to know that it is integrable.
b) It is not allowed to apply Ito's formula to a function $u$ that is defined on $D$ only and not on the whole of $\mathbb{R}^d$. Moreover, it is not clear that the solution $u$ can be extended to the whole of $\mathbb{R}^d$ in a $C^2$ way. Actually, even the boundary condition $\phi$ itself might not be extendable to a $C^2$ function.
c) We must show that $t \mapsto u'(\xi^x_t)$ belongs to $M^2$.

In the next sections we deal rigorously with these questions and we shall find similar representation formulas for the solutions of other PDE problems.

Note that, as stated above, the previous argument leads to the representation of the solutions of problem (10.1), once we know already that this problem actually has a solution. The PDE theory is actually well developed and provides many existence results for these problems. Once the representation (10.2) is obtained we can, however, work the argument backwards: in a situation where the existence of a solution is not known we can consider the function defined in (10.2) and try to show that it is a solution. In other words, the representation formulas can serve as a starting point in order to prove the existence in situations in where this is not guaranteed *a priori*. This is what we shall do in Sects. 10.4 and 10.6.

## 10.2 The Dirichlet problem

We need a preliminary result. Let $\xi$ be the solution of the SDE

$$\xi_t = x + \int_s^t b(\xi_u, u)\, du + \int_s^t \sigma(\xi_u, u)\, dB_u \;,$$

where $b$ and $\sigma$ satisfy Assumption (A'). Let $D$ be a *bounded* open set containing $x$. We see now that, under suitable hypotheses, the exit time from $D$, $\tau = \inf\{t; t \geq s, \xi_t(\omega) \notin D\}$, is finite and actually integrable. Let, as usual,

$$L_t = \frac{1}{2} \sum_{i,j=1}^m a_{ij}(x,t)\, \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x,t)\, \frac{\partial}{\partial x_i},$$

where $a = \sigma\, \sigma^*$. Let us consider the following assumption.

Assumption $H_0$. There exists a $C^{2,1}(\mathbb{R}^m \times \mathbb{R}^+)$ function $\Phi$, positive on $D \times \mathbb{R}^+$ and such that, on $D \times \mathbb{R}^+$,

$$\frac{\partial \Phi}{\partial t} + L_t \Phi \leq -1.$$

Under Assumption $H_0$ the exit time from $D$ is finite and integrable.

**Proposition 10.1** Under Assumption $H_0$ and if $D$ is bounded, $\mathrm{E}(\tau) < +\infty$ for every $(x,s) \in D \times \mathbb{R}^+$.

*Proof* By Ito's formula, for every $t \geq s$ we have

$$d\Phi(\xi_t, t) = \Big(\frac{\partial \Phi}{\partial t} + L_t \Phi\Big)(\xi_t, t)\, dt + \sum_{i,j} \frac{\partial \Phi}{\partial x_i}(\xi_t, t) \sigma_{ij}(\xi_t, t)\, dB_j(t)$$

hence for $s < t$

$$\begin{aligned} &\Phi(\xi_{t\wedge\tau}, t \wedge \tau) - \Phi(x,s) \\ &= \int_s^{t\wedge\tau} \Big(\frac{\partial \Phi}{\partial u} + L_u \Phi\Big)(\xi_u, u)\, du + \int_s^{t\wedge\tau} \sum_{i,j} \frac{\partial \Phi}{\partial x_i}(\xi_u, u) \sigma_{ij}(\xi_u, u)\, dB_j(u) \;. \end{aligned}$$

As $\sigma$ and the first derivatives of $\Phi$ are bounded on $D \times [s, t]$, the last term is a martingale and has mean 0; hence for every $t > s$,

$$\mathrm{E}[\Phi(\xi_{t\wedge\tau}, t\wedge\tau)] - \Phi(x, s) = \mathrm{E}\Big(\int_s^{t\wedge\tau} \Big(\frac{\partial\Phi}{\partial u} + L_u\Phi\Big)(\xi_u, u)\, du\Big) \le -\mathrm{E}(t\wedge\tau) + s\,.$$

As $\Phi$ is positive on $D \times \mathbb{R}^+$, we obtain $\Phi(x, s) + s \ge \mathrm{E}(t\wedge\tau)$. Now just take the limit as $t \to +\infty$ and apply Fatou's lemma. □

The following two assumptions, which are easier to check, both imply that Assumption $H_0$ is satisfied.

Assumption $H_1$. There exist two constants $\lambda > 0$, $c > 0$ and an index $i$, $1 \le i \le m$, such that

$$a_{ii}(x, t) \ge \lambda\,, \qquad b_i(x, t) \ge -c$$

for every $x \in D$, $t \ge 0$.

Assumption $H_2$. On $D \times \mathbb{R}^+$ $b$ is bounded and the matrix field $a$ is uniformly elliptic, i.e. such that $\langle a(x, t)z, z\rangle \ge \lambda|z|^2$ for some $\lambda > 0$ and for every $(x, t) \in \overline{D} \times \mathbb{R}^+$, $z \in \mathbb{R}^m$.

**Proposition 10.2** $H_2 \Rightarrow H_1 \Rightarrow H_0$.

*Proof* $H_2 \Rightarrow H_1$: we have $a_{ii}(x, t) = \langle a(x, t)\, e_i, e_i\rangle \ge \lambda$, where $e_i$ denotes the unitary vector pointing in the $i$-th direction.

$H_1 \Rightarrow H_0$: let $\Phi(x, t) = \beta(\mathrm{e}^{\alpha R} - \mathrm{e}^{\alpha x_i})$, where $R$ is the radius of a sphere containing $D$, and $\alpha, \beta$ are positive numbers to be specified later ($\Phi$ therefore does not depend on $t$); then $\Phi > 0$ on $D$ and

$$\begin{aligned}\frac{\partial\Phi}{\partial t} + L_t\Phi &= -\beta\mathrm{e}^{\alpha x_i}\big(\tfrac{1}{2}\alpha^2 a_{ii}(x, t) + \alpha b_i(x, t)\big) \le -\beta\mathrm{e}^{\alpha x_i}\big(\tfrac{1}{2}\alpha^2\lambda - \alpha c\big)\\ &\le -\alpha\beta\mathrm{e}^{-\alpha R}\big(\tfrac{1}{2}\alpha\lambda - c\big)\end{aligned}$$

and we can make the last term $\le -1$ by first choosing $\alpha$ so that $\frac{1}{2}\alpha\lambda - c > 0$ and then $\beta$ large. □

In particular, the exit time from a bounded open set $D$ is integrable if the generator of the diffusion is elliptic.

Let us assume that the open bounded set $D$ is connected and with a $C^2$ boundary; let

$$L = \frac{1}{2} \sum_{i,j=1}^{m} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{m} b_i(x) \frac{\partial}{\partial x_i}$$

be a differential operator on $\overline{D}$ such that

a) $L$ is uniformly elliptic on $\overline{D}$, i.e. $\langle a(x)\, z, z\rangle \geq \lambda |z|^2$ for some $\lambda > 0$ and for every $x \in \overline{D}$, $z \in \mathbb{R}^m$;
b) $a$ and $b$ are Lipschitz continuous on $\overline{D}$.

The following existence and uniqueness theorem is classical and well-known (see Friedman 1964, 1975).

**Theorem 10.1** Let $\phi$ be a continuous function on $\partial D$ and $c$ and $f$ Hölder continuous functions $\overline{D} \to \mathbb{R}$ with $c \geq 0$. Then, if $L$ satisfies conditions a) and b) above, there exists a unique function $u \in C^2(D) \cap C(\overline{D})$ such that

$$\begin{cases} Lu - cu = f & \text{on } D \\ u_{|\partial D} = \phi\,. \end{cases} \tag{10.4}$$

We prove now that the solution $u$ of (10.4) can be represented as a functional of a suitable diffusion process, in particular giving a rigorous proof of (10.2).

By Proposition 9.2 there exists an $m \times m$ matrix field $\sigma$, Lipschitz continuous on $\overline{D}$, such that $a = \sigma\sigma^*$. As $\sigma$ and $b$ are assumed to be Lipschitz continuous on $\overline{D}$, they can be extended to the whole of $\mathbb{R}^m$ with Assumption (A) satisfied. Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion and $\xi^x$ the solution of the SDE

$$d\xi_t^x = x + \int_0^t b(\xi_s^x)\, ds + \int_0^t \sigma(\xi_s^x)\, dB_s\,.$$

Moreover, let $Z_t = \mathrm{e}^{-\int_0^t c(\xi_s^x)\, ds}$, so that $dZ_t = -c(\xi_t^x) Z_t\, dt$. Let, for $\varepsilon > 0$, $D_\varepsilon$ be a regular open set such that $\overline{D}_\varepsilon \subset D$ and $\mathrm{dist}(\partial D_\varepsilon, \partial D) \leq \varepsilon$; let $\tau_\varepsilon$ be the exit time from $D_\varepsilon$: obviously $\tau_\varepsilon \le \tau$ and, moreover, $\tau < +\infty$ by Propositions 10.1 and 10.2; as the paths are continuous, $\tau_\varepsilon \nearrow \tau$ a.s. as $\varepsilon \searrow 0$. Let $u_\varepsilon$ be a bounded $C^2(\mathbb{R}^m)$ function coinciding with $u$ on $D_\varepsilon$. Then Ito's formula gives

$$\begin{aligned} dZ_t u_\varepsilon(\xi_t^x) &= u_\varepsilon(\xi_t^x)\, dZ_t + Z_t\, du_\varepsilon(\xi_t^x) \\ &= -c(\xi_t^x) Z_t u_\varepsilon(\xi_t^x)\, dt + Z_t \Big( Lu_\varepsilon(\xi_t^x)\, dt + \sum_{i,j=1}^m \frac{\partial u_\varepsilon}{\partial x_i}(\xi_s^x) \sigma_{ij}(\xi_s^x)\, dB_j(s) \Big)\,. \end{aligned}$$

Therefore, P-a.s.,

$$\begin{aligned} Z_{t\wedge\tau_\varepsilon} u_\varepsilon(\xi_{t\wedge\tau_\varepsilon}^x) = u_\varepsilon(x) &+ \int_0^{t\wedge\tau_\varepsilon} \big( Lu_\varepsilon(\xi_s^x) - c(\xi_s^x) u_\varepsilon(\xi_s^x) \big) Z_s\, ds \\ &+ \int_0^{t\wedge\tau_\varepsilon} \sum_{i,j=1}^m \frac{\partial u_\varepsilon}{\partial x_i}(\xi_s^x) Z_s \sigma_{ij}(\xi_s^x)\, dB_j(s)\,. \end{aligned} \tag{10.5}$$

As the derivatives of $u$ on $\overline{D}_\varepsilon$ are bounded, the stochastic integral above has 0 mean and as $u$ coincides with $u_\varepsilon$ on $D_\varepsilon$ we get, taking the expectation for $x \in D_\varepsilon$,

$$\mathrm{E}[u(\xi_{t\wedge\tau_\varepsilon}^x) Z_{t\wedge\tau_\varepsilon}] = u(x) + \mathrm{E}\Big( \int_0^{t\wedge\tau_\varepsilon} f(\xi_s^x) Z_s\, ds \Big)\,.$$

Now, taking the limit as $\varepsilon \searrow 0$,

$$\mathrm{E}[u(\xi_{t\wedge\tau}^x) Z_{t\wedge\tau}] = u(x) + \mathrm{E}\Big[ \int_0^{t\wedge\tau} Z_s f(\xi_s^x)\, ds \Big]\,.$$

We now take the limit as $t \to +\infty$: we have $\mathrm{E}[u(\xi_{t\wedge\tau}^x) Z_{t\wedge\tau}] \to \mathrm{E}[u(\xi_\tau^x) Z_\tau]$ as $u$ is bounded (recall that $0 \le Z_t \le 1$). As for the right-hand side, since $Z_s \le 1$ for every $s$,

$$\Big| \int_0^{t\wedge\tau} Z_s f(\xi_s^x)\, ds \Big| \le \tau \|f\|_\infty\,.$$

As $\tau$ is integrable by Propositions 10.1 and 10.2, we can apply Lebesgue's theorem, which gives

$$\mathrm{E}[u(\xi_\tau^x) Z_\tau] = u(x) + \mathrm{E}\Big[ \int_0^\tau Z_s f(\xi_s^x)\, ds \Big]\,.$$

Finally, as $\xi_\tau^x \in \partial D$ a.s., $u(\xi_\tau^x) = \phi(\xi_\tau^x)$ and we obtain

$$u(x) = \mathrm{E}[\phi(\xi_\tau^x) Z_\tau] - \mathrm{E}\Big[ \int_0^\tau Z_s f(\xi_s^x)\, ds \Big]\,.$$

This formula can also be expressed in terms of the canonical realization:

**Theorem 10.2** Under the assumptions of Theorem 10.1, the solution of the PDE problem (10.4) is given by

$$u(x) = \mathrm{E}^x[\phi(X_\tau)Z_\tau] - \mathrm{E}^x\Big[\int_0^\tau Z_s f(X_s)\,ds\Big]\,, \tag{10.6}$$

where $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_{t\geq 0}, (X_t)_{t\geq 0}, (\mathrm{P}^x)_x)$ is the canonical diffusion associated to the infinitesimal generator $L$, $Z_t = \mathrm{e}^{-\int_0^t c(X_s)\,ds}$ and $\tau = \inf\{t; X_t \notin D\}$.

In the case $c = 0, f = 0$, (10.6) becomes simply

$$u(x) = \mathrm{E}^x[\phi(X_\tau)]\,, \tag{10.7}$$

i.e. the value at $x$ of the solution $u$ is the mean of the boundary value $\phi$ taken with respect to the law of $X_\tau$, which is the exit distribution of the diffusion starting at $x$.

As remarked at the beginning of this chapter, formulas such as (10.6) or (10.7) are interesting in two ways:

- for the computation of the means of functionals of diffusion processes, such as those appearing on the right-hand sides in (10.6) and (10.7), tracing them back to the computation of the solution of the problem (10.4),
- in order to obtain information about the solution $u$ of (10.4). For instance, let us remark that in the derivation of (10.6) we only used the existence of $u$. Our computation therefore provides a proof of the uniqueness of the solution in Theorem 10.1.

The *Poisson kernel* of the operator $L$ on $D$ is a family $(\Pi(x,\cdot))_{x\in D}$ of measures on $\partial D$ such that, for every continuous function $\phi$ on $\partial D$, the solution of

$$\begin{cases} Lu = 0 & \text{on } D \\ u_{|\partial D} = \phi \end{cases} \tag{10.8}$$

is given by

$$u(x) = \int_{\partial D} \phi(y)\,\Pi(x,dy)\,.$$

(10.7) states that, under the hypotheses of Theorem 10.1, the Poisson kernel always exists and $\Pi(x,\cdot)$ is the law of $X_\tau$ with respect to $\mathrm{P}^x$ (i.e. the exit distribution from $D$ of the diffusion $X$ starting at $x \in D$). This identification allows us to determine the exit distribution in those cases in which the Poisson kernel is known (see the next example and Exercise 10.4, for example).

*Example 10.1* Let $L = \frac{1}{2}\triangle$ and $B_R$ be the open sphere centered at 0 with radius $R$ in $\mathbb{R}^m$, $m \geq 1$; let $\sigma(dy)$ be the surface element of $\partial B_R$ and $\omega_m = \frac{2\pi^{m/2}}{\Gamma(m/2)}$ the measure of the surface of the sphere of radius 1 in $\mathbb{R}^m$. Then it is known that, for $y \in \partial B_R, x \in B_R$, if

$$N_R(x,y) = \frac{1}{R\omega_m}\,\frac{R^2 - |x|^2}{|x-y|^m}\;,$$

then the Poisson kernel of $\frac{1}{2}\triangle$ on $B_R$ is given by

$$\Pi(x,dy) = N_R(x,y)\,\sigma(dy)\;. \tag{10.9}$$

Therefore (10.9) gives the exit distribution from $B_R$ of a Brownian motion starting at a point $x \in B_R$. If $x =$the origin we find the uniform distribution on $\partial B_R$, as we already know from Exercise 3.18.

*Example 10.2* Let $(X_t)_t$ be a one-dimensional diffusion process associated to some SDE with Lipschitz continuous coefficients $b$, $\sigma$. Let $[a,b]$ be some finite interval and $x \in ]a,b[$ and assume that $\sigma > \alpha > 0$ on $[a,b]$. We know from Propositions 10.1 and 10.2 that the exit time, $\tau$, from $[a,b]$ is a.s. finite. We want to compute the probability

$$\mathrm{P}(X_\tau = a)\;.$$

We can use the representation formula (10.7) with $D = ]a,b[$ and $\phi$ given by $\phi(a) = 1$, $\phi(b) = 0$ (the boundary $\partial D$ here is reduced to $\{a,b\}$). By 10.7 therefore

$$u(x) = \mathrm{P}^x(X_\tau = a) = \mathrm{E}^x[\phi(X_\tau)]$$

is the solution of the ordinary problem

$$\begin{cases} bu' + \dfrac{\sigma^2}{2}u'' = 0 & \text{on } ]a,b[\;, \\ u(a) = 1,\; u(b) = 0\;. \end{cases}$$

(continued)

*Example 10.2* (continued)
Let us compute the solution: if $v = u'$ the ODE becomes $bv + \frac{1}{2}\sigma^2 v' = 0$, i.e.

$$\frac{v'}{v} = -\frac{2b}{\sigma^2}\,,$$

which gives $\log v(x) = -\int_a^x \frac{2b}{\sigma^2}(y)\,dy + c_1$ i.e.

$$v(x) = c_1 \mathrm{e}^{-\int_a^x \frac{2b(y)}{\sigma^2(y)}\,dy}$$

and therefore

$$u(x) = c_2 + c_1 \int_a^x \mathrm{e}^{-\int_a^z \frac{2b(y)}{\sigma^2(y)}\,dy}\,dz\,.$$

The condition $u(a) = 1$ gives at once $c_2 = 1$, from which also $c_1$ is easily computed giving

$$u(x) = 1 - \frac{\displaystyle\int_a^x \mathrm{e}^{-\int_a^z \frac{2b(y)}{\sigma^2(y)}\,dy}\,dz}{\displaystyle\int_a^b \mathrm{e}^{-\int_a^z \frac{2b(y)}{\sigma^2(y)}\,dy}\,dz}\,\cdot$$

It is useful to note that to solve the equation $bu' + \frac{\sigma^2}{2}u'' = 0$ is equivalent to finding a function $u$ such that $(u(X_t))_t$ is a martingale. We know already that problems of determining the exit distribution are often solved using an appropriate martingale.

In the following examples, conversely, formulas (10.6) and (10.7) are used in order to prove properties of the solutions of the Dirichlet problem (10.4).

*Remark 10.1 (The maximum principle)* Let $D$ be a bounded domain and

$$L = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x)\frac{\partial}{\partial x_i}$$

a uniformly elliptic differential operator on $\overline{D}$; let us assume $a$ and $b$ are Lipschitz continuous on $\overline{D}$. The *maximum principle* states that for a function $u$ which is in $C^2(D) \cap C(\overline{D})$ and such that $Lu = 0$, we have

$$\min_{\partial D} u \le u(x) \le \max_{\partial D} u\,.$$

(continued)

*Remark 10.1* (continued)
This is now an immediate consequence of (10.7), as it is immediate that

$$\min_{\partial D} u \le \mathrm{E}^x[u(X_\tau)] \le \max_{\partial D} u\ .$$

If $u$ satisfies the weaker condition $Lu = f \ge 0$ (but still $c = 0$) then it enjoys a weaker form of the maximum principle: starting from (10.6) we have

$$u(x) = \mathrm{E}^x[u(X_\tau)] - \mathrm{E}^x\Big[\int_0^\tau f(X_s)\,ds\Big] \le \mathrm{E}^x[u(X_\tau)] \le \max_{\partial D} u\ ,$$

($c = 0$ implies $Z_t \equiv 1$). Recall that we have already come across the maximum principle in Sect. 6.3 (Proposition 6.5).

*Remark 10.2* Let $D$ be a bounded domain and $u : D \to \mathbb{R}$ a harmonic function, i.e. such that $\frac{1}{2}\triangle u = 0$ on $D$. Then $u$ enjoys the *mean property*: if we denote by $B_R(x)$ the ball of radius $R$ and by $\sigma$ the normalized Lebesgue measure of the surface $\partial B_R(x)$, then

$$u(x) = \int_{\partial B_R(x)} u(y)\,d\sigma(y) \tag{10.10}$$

for every $x \in D$ and $R > 0$ such that $B_R(x) \subset D$.

Indeed, from (10.7), applied to $D = B_R(x)$, we have $u(x) = \mathrm{E}^x[u(X_{\tau_R})]$, where $\tau_R$ denotes the exit time from $B_R(x)$ and we know that the exit distribution of a Brownian motion from a ball centered at its starting point is the normalized Lebesgue measure of the surface of the ball (see Exercise 3.18).

Actually the mean property (10.10) characterizes the harmonic functions (see also Example 6.3)

Formula (10.6) also gives a numerical approach to the computation of the solution of a PDE problem as we shall see in Sect. 11.4.

## 10.3 Parabolic equations

A representation formula can also be obtained in quite a similar way for the Cauchy–Dirichlet problem.

Let $D \subset \mathbb{R}^m$ be a bounded connected open set having a $C^2$ boundary and let $Q = D \times [0, T[$ and

$$L_t = \frac{1}{2}\sum_{i,j=1}^{m} a_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{m} b_i(x,t)\frac{\partial}{\partial x_i} \tag{10.11}$$

be a differential operator on $\overline{Q}$ such that

a) $\langle a(x,t)\, z, z\rangle \ge \lambda |z|^2$, where $\lambda > 0$ for every $(x,t) \in Q$, $z \in \mathbb{R}^m$;
b) $a$ and $b$ are Lipschitz continuous in $\overline{Q}$.

Then (Friedman 1964, 1975) the following existence and uniqueness result holds. Note that assumption b) above requires that the coefficients are also Lipschitz continuous in the time variable.

**Theorem 10.3** Let $\phi$ be a continuous function on $\overline{D}$ and $g$ a continuous function on $\partial D \times [0, T]$ such that $g(x, T) = \phi(x)$ if $x \in \partial D$; let $c$ and $f$ be Hölder continuous functions $\overline{Q} \to \mathbb{R}$. Then, under the previous hypotheses a) and b), there exists a unique function $u \in C^{2,1}(Q) \cap C(\overline{Q})$ such that (Fig. 10.1)

$$\begin{cases} L_t u - cu + \dfrac{\partial u}{\partial t} = f & \text{on } Q \\ u(x, T) = \phi(x) & \text{on } D \\ u(x, t) = g(x, t) & \text{on } \partial D \times [0, T]\,. \end{cases} \tag{10.12}$$

By Proposition 9.2 there exists a Lipschitz continuous matrix field $\sigma$ on $\overline{Q}$ such that $\sigma\sigma^* = a$; moreover, the coefficients $\sigma$ and $b$ can be extended to

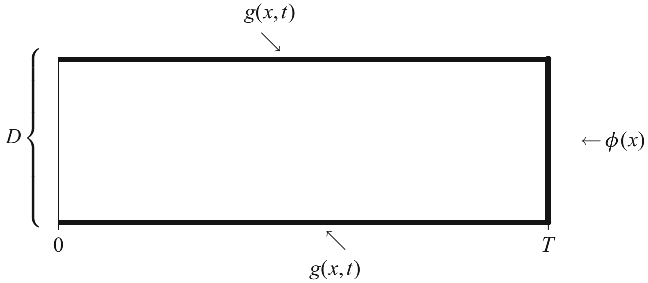


**Fig. 10.1** In the problem (10.12) the boundary values are given on the boundary of $D$ (i.e. on $\partial D \times [0, T]$) and at the final time $T$ (i.e. on $D \times \{T\}$). Note that the hypotheses of Theorem 10.3 require that on $\partial D \times \{T\}$, where these boundary conditions "meet", they must coincide

$\mathbb{R}^m \times [0, T]$ preserving Lipschitz continuity and boundedness. Hence $b$ and $\sigma$ satisfy Assumption (A) (see p. 260). Actually they are even Lipschitz continuous in the time variable, which is not required in Assumption (A).

**Theorem 10.4** Let $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t^s)_{t\ge s}, (X_t)_t, (\mathrm{P}^{x,s})_{x,s})$ be the canonical diffusion associated to the SDE with coefficients $b$ and $\sigma$. Then under the hypotheses of Theorem 10.3 we have the representation formula

$$\begin{aligned} u(x,t) &= \mathrm{E}^{x,t}\Big[g(X_\tau,\tau)\,\mathrm{e}^{-\int_t^\tau c(X_s,s)\,ds}1_{\{\tau<T\}}\Big] \\ &\quad + \mathrm{E}^{x,t}\Big[\phi(X_T)\,\mathrm{e}^{-\int_t^T c(X_s,s)\,ds}1_{\{\tau\ge T\}}\Big] \\ &\quad -\mathrm{E}^{x,t}\Big[\int_t^{\tau\wedge T} f(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_u,u)\,du}\,ds\Big], \end{aligned} \tag{10.13}$$

where $\tau$ is the exit time from $D$.

*Proof* The argument is similar to the one in the proof of (10.6): we must apply Ito's formula to the process $Y_s = u_\varepsilon(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_r,r)\,dr}$, where $u_\varepsilon$ approximates $u$ in a way similar to (10.6). □

In the particular case $g = 0$, $c = 0$, $f = 0$, (10.13) becomes

$$u(x,t) = \mathrm{E}^{x,t}[\phi(X_T)1_{\{\tau\ge T\}}]\;.$$

Note that this relation follows from Theorem 10.4 *if* we assume that $\phi$ vanishes on $\partial D$, otherwise the condition $g(x,T) = \phi(x)$ if $x \in \partial D$, which is required in Theorem 10.3, is not satisfied.

*Remark 10.3* If (Fig. 10.2) $b, \sigma, f, g, c$ do not depend on $t$ then if $v(x,t) = u(x,T-t)$, $v$ is a solution of

$$\begin{cases} Lv - \dfrac{\partial v}{\partial t} - cv = f(x) & \text{on } Q \\ v(x,0) = \phi(x) & \text{on } D \\ v(x,t) = g(x) & \text{on } \partial D\times[0,T]\;. \end{cases} \tag{10.14}$$

For the solution $v$ of (10.14) we have the representation (recall that now the diffusion associated to $L$ is time homogeneous, so that $\mathrm{E}^{x,t}[\phi(X_T)1_{\{\tau\ge T\}}]$ can be written as $\mathrm{E}^{x}[\phi(X_{T-t})1_{\{\tau\ge T-t\}}]$)

$$v(x,t) = u(x,T-t) = \mathrm{E}^{x}[\phi(X_t)1_{\{\tau\ge t\}}]\;.$$

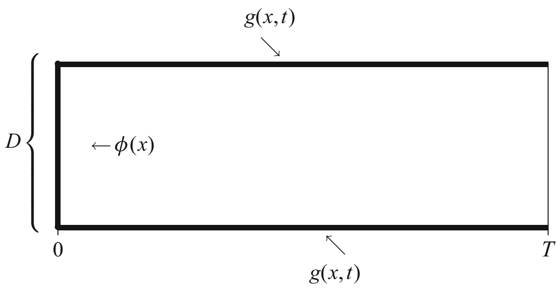


**Fig. 10.2** In the problem (10.14) the boundary values are given on $\partial D \times [0, T]$ and at time 0, i.e. it is an "initial value" problem. Note that, with respect to (10.12), the term in $\frac{\partial}{\partial t}$ has the $-$ sign

## 10.4 The Feynman–Kac formula

In this section we investigate representation formulas similar to the one of Theorem 10.4, for the domain $\mathbb{R}^m \times [0, T]$. We shall see applications of them in the next chapters. The main difficulties come from the fact that $\mathbb{R}^m$ is not bounded.

Let $\phi, f$ be two real continuous functions defined respectively on $\mathbb{R}^m$ and $\mathbb{R}^m \times [0, T]$ and let us make one of the assumptions

$$|\phi(x)| \le M(1 + |x|^\lambda), \quad |f(x,t)| \le M(1 + |x|^\lambda) \quad (x,t) \in \mathbb{R}^m \times [0, T] \tag{10.15}$$

for some $\lambda \ge 0$ (condition of polynomial growth), or

$$\phi \ge 0, \qquad f \ge 0\,. \tag{10.16}$$

Let $L_t$ be a differential operator on $\mathbb{R}^m \times [0, T[$, as in (10.11), and let us consider its associated canonical diffusion $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t^s)_{t \ge s}, (X_t)_t, (\mathrm{P}^{x,s})_{x,s})$.

**Theorem 10.5** Let $c : \mathbb{R}^m \times [0, T] \to \mathbb{R}$ be a continuous function bounded below (i.e. $c(x,t) \ge -K > -\infty$) and $w$ a $C^{2,1}(\mathbb{R}^m \times [0, T[)$ function, continuous on $\mathbb{R}^m \times [0, T]$ and a solution of the problem

$$\begin{cases} L_t w + \dfrac{\partial w}{\partial t} - cw = f & \text{on } \mathbb{R}^m \times [0, T[ \\ w(x, T) = \phi(x)\,. \end{cases} \tag{10.17}$$

(continued)

**Theorem 10.5** (continued)
Let us assume that the coefficients $a$ and $b$ are Lipschitz continuous on $[0,T]\times\mathbb{R}^m$ and that $a$ is uniformly elliptic. Assume also that the functions $\phi, f$, in addition to being continuous, satisfy one of the hypotheses (10.15) or (10.16). Assume, finally, that $w$ has a polynomial growth, i.e. there exists $M_1, \mu > 0$ such that

$$|w(x,t)| \le M_1(1+|x|^\mu) \tag{10.18}$$

for every $t \in [0,T]$, $x \in \mathbb{R}^m$. Then we have the representation formula

$$w(x,t) = \mathrm{E}^{x,t}\Big[\phi(X_T)\,\mathrm{e}^{-\int_t^T c(X_s,s)\,ds}\Big] - \mathrm{E}^{x,t}\Big[\int_t^T f(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_u,u)\,du}\,ds\Big]\,. \tag{10.19}$$

*Proof* Let $\tau_R$ be the exit time of $X$ from the sphere of radius $R$. Then, by Theorem 10.4 applied to $D = B_R$, we have, for every $(x,t) \in B_R \times [0,T[$,

$$\begin{aligned} w(X_t,t) &= -\mathrm{E}^{x,t}\Big[\int_t^{T\wedge\tau_R} f(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_u,u)\,du}\,ds\Big]\\ &\quad+\mathrm{E}^{x,t}\Big[w(X_{\tau_R},\tau_R)\,\mathrm{e}^{-\int_t^{\tau_R} c(X_u,u)\,du}1_{\{\tau_R<T\}}\Big]\\ &\quad+\mathrm{E}^{x,t}\Big[\phi(X_T)\,\mathrm{e}^{-\int_t^T c(X_u,u)\,du}1_{\{\tau_R\ge T\}}\Big]\\ &= I_1+I_2+I_3\,. \end{aligned} \tag{10.20}$$

Let us first assume that (10.15) holds. By Lemma 9.1 we know that

$$\mathrm{E}^{x,t}\Big[\max_{t\le u\le T}|X_u|^q\Big] \le C_q(1+|x|^q) \tag{10.21}$$

for every $q \ge 2$, where the constant $C_q$ does not depend on $x$. Then by the Markov inequality we have

$$\mathrm{P}^{x,t}(\tau_R \le T) = \mathrm{P}^{x,t}\Big(\max_{t\le u\le T}|X_u| \ge R\Big) \le C_q R^{-q}(1+|x|^q)\,, \tag{10.22}$$

hence $\mathrm{P}^{x,t}(\tau_R \le T) \to 0$ as $R \to +\infty$, faster than every polynomial. The map $R \to \tau_R$ is increasing and $\tau_R \wedge T \to T$ as $R \to +\infty$ for every $\omega$. Therefore if we let $R \to +\infty$ in the first term in (10.20) and apply Lebesgue's theorem, thanks to (10.21) and to the second relation in (10.15), we find that it converges to

$$-\mathrm{E}^{x,t}\Big[\int_t^T f(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_u,u)\,du}\,ds\Big]\,.$$

A similar argument works for $I_3$: as $R \to +\infty$ the r.v. inside the expectation converges to

$$\phi(X_T)\,\mathrm{e}^{-\int_t^T c(X_u,u)\,du}\,,$$

being bounded by $|\phi(X_T)|\,\mathrm{e}^{KT}$ which, thanks to (10.21) and to the first relation in (10.15), is an integrable r.v. We still have to prove that $I_2 \to 0$ as $R \to +\infty$. In fact

$$\Big|w(X_{\tau_R}, \tau_R)\,\mathrm{e}^{-\int_t^{\tau_R} c(X_u,u)\,du} 1_{\{\tau_R \le T\}}\Big| \le M_1 \mathrm{e}^{KT}(1 + R^\mu) 1_{\{\tau_R \le T\}}$$

and therefore

$$\mathrm{E}^{x,t}\big[w(X_{\tau_R}, \tau_R)\,\mathrm{e}^{-\int_t^{T} c(X_u,u)\,du} 1_{\{\tau_R > T\}}\big] \le M_1 \mathrm{e}^{KT}(1 + R^\mu)\,\mathrm{P}^{x,t}(\tau_R \le T)$$

and the left-hand side tends to 0 as $R \to +\infty$ as in (10.22) (just choose $q > \mu$).

If, conversely, (10.16) holds, then we argue similarly, the only difference being that now the convergence of $I_1$ and $I_3$ follows from Beppo Levi's Theorem. □

*Remark 10.4* The hypothesis of polynomial growth (10.18) can be weakened if, in addition, $\sigma$ is bounded. In this case the process $(X_t)_t$ enjoys a stronger integrability property than (10.21). In Exercise 9.29 we prove that if

$$\sup_{x\in\mathbb{R}^m} |\sigma(x)\sigma^*(x)| \le k^*, \qquad |b(x)| \le M_b(1 + |x|)\ ,$$

where $|\ |$ denotes the operator norm of the matrix, then

$$\mathrm{E}^{x,t}\Big[\exp\Big(\alpha \max_{t\le u\le T} |X_u|^2\Big)\Big] \le K$$

for every $\alpha < c = \mathrm{e}^{-2M_bT}(2Tmk^*)^{-1}$. It is easy then to see that, if we replace (10.18) by

$$|w(x,t)| \le M_1 \mathrm{e}^{\alpha|x|^2} \tag{10.23}$$

for some $\alpha < c$, then we can repeat the proof of Theorem 10.5 and the representation formula (10.19) still holds. In this case, i.e. if $\sigma$ is bounded, also the hypotheses of polynomial growth for $\phi$ and $f$ can be replaced by

$$|\phi(x)| \le M\mathrm{e}^{\alpha|x|^2}, \quad |f(x,t)| \le M\mathrm{e}^{\alpha|x|^2} \qquad (x,t) \in \mathbb{R}^m \times [0,T] \tag{10.24}$$

for some $\alpha < c = \mathrm{e}^{-2MT}(2Tmk^*)^{-1}$.

Theorem 10.5 gives an important representation formula for the solutions of the parabolic problem (10.17). We refer to it as the *Feynman–Kac formula*.

It would, however, be very useful to have conditions for the existence of a solution of (10.17) satisfying (10.18). This would allow us, for instance, to state that the function defined as the right-hand side of (10.19) is a solution of (10.17).

Existence results are available in the literature: see Friedman (1975, p. 147), for example, where, however, boundedness for the coefficients $a$ and $b$ is required, an assumption that is often too constraining and which will not be satisfied in some applications of Chap. 13.

Until now we have given *representation formulas* for the solutions of PDE problems as functionals of diffusion processes. Let us conversely try to *construct* a solution and therefore prove that a solution exists.

Let us assume $c$ and $f$ to be locally Hölder continuous and let, for $(x,t) \in \mathbb{R}^m \times [0,T]$,

$$u(x,t) = \underbrace{\mathrm{E}^{x,t}\Big[\phi(X_T)\,\mathrm{e}^{-\int_t^T c(X_s,s)\,ds}\Big]}_{:=u_1(x,t)} - \underbrace{\mathrm{E}^{x,t}\Big[\int_t^T f(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_u,u)\,du}\,ds\Big]}_{:=u_2(x,t)}, \tag{10.25}$$

where, as above, $L_t$ is a differential operator on $\mathbb{R}^m \times [0,T[$, as in (10.11), and $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t^s)_{t\ge s}, (X_t)_t, (\mathrm{P}^{x,s})_{x,s})$ is the associated canonical diffusion.

We shall now prove that such a function $u$ is continuous on $\mathbb{R}^m \times [0,T]$, that it is $C^{2,1}(\mathbb{R}^m \times [0,T[)$ and a solution of (10.17).

Let $B_R$ be the ball centered at the origin and of radius $R$ large enough so that $x \in B_R$. The idea is simply to show that, for every $R > 0$, $u$ satisfies a relation of the same kind as (10.13), with a function $g$ to be specified on $\partial B_R \times [0,T[$. Theorems 10.3 and 10.4 then allow us to state that $u$ has the required regularity and that it is a solution of (10.12) on $B_R \times [0,T[$. In particular, by the arbitrariness of $R$, we shall have

$$L_t u + \frac{\partial u}{\partial t} - cu = f \qquad \text{on } \mathbb{R}^m \times [0,T[\,. \tag{10.26}$$

The key tool will be the strong Markov property.

**Lemma 10.1** Let us assume that

- the coefficients $a$ and $b$ are Lipschitz continuous on $[0,T] \times \mathbb{R}^m$ and that $a$ is uniformly elliptic
- that the function $c$ is continuous and bounded below, on $\mathbb{R}^m \times [0,T]$, by a constant $-K$.

Then if $\phi$ is continuous and has polynomial growth, the function $u_1$ defined in (10.25) is continuous in $(x,t)$. The same holds for $u_2$ if $f$ is continuous and has polynomial growth.

If, moreover, the diffusion coefficient is bounded, it is sufficient to assume that $\phi$ and $f$ have an exponential growth, i.e. that there exist constants $k_1, k_2 \ge 0$ such that $|\phi(x)| \le k_1 \mathrm{e}^{k_2|x|}$, $|f(x,t)| \le k_1 \mathrm{e}^{k_2|x|}$ for every $0 \le t \le T$.

*Proof* This is an application of the continuity property of a diffusion with respect to the initial conditions (Sect. 9.8). The idea is simple but, to be rigorous, we need to be careful. One can write

$$u_1(x,t) = \mathrm{E}\Big[\phi(\xi_T^{x,t})\,\mathrm{e}^{-\int_t^T c(\xi_u^{x,t},u)\,du}\Big]\,,$$

where $\xi^{x,t}$ denotes, as in the previous chapter, the solution of

$$\begin{aligned} d\xi_s &= b(\xi_s,s)\,ds + \sigma(\xi_s,s)\,dB_s \\ \xi_t &= x \end{aligned}$$

with respect to a Brownian motion $(\Omega,\mathscr{F},(\mathscr{F}_t)_t,(B_t)_t,\mathrm{P})$.

By Theorem 9.9 we can assume that $(x,t,s)\mapsto \xi_s^{x,t}(\omega)$ is a continuous function for every $\omega\in\Omega$. If $(x_n,t_n)\to(x,t)$ as $n\to\infty$, let $R$ be a number large enough so that $x,x_n\in B_R$.

Let $\phi_R$ be a bounded continuous function coinciding with $\phi$ on $B_R$ and let us denote by $\tau$ and $\tau_n$, respectively, the exit times of $\xi^{x,t}$ and $\xi^{x_n,t_n}$ from $B_R$. Then

$$\begin{aligned} &|u_1(x_n,t_n)-u_1(x,t)| \\ &\le \mathrm{E}\Big[\Big|\phi(\xi_T^{x_n,t_n})\,\mathrm{e}^{-\int_{t_n}^T c(\xi_u^{x_n,t_n},u)\,du} - \phi(\xi_T^{x,t})\,\mathrm{e}^{-\int_t^T c(\xi_u^{x,t},u)\,du}\Big|\Big] \\ &\le \mathrm{E}\Big[\Big|\phi_R(\xi_T^{x_n,t_n})\,\mathrm{e}^{-\int_{t_n}^T c(\xi_u^{x_n,t_n},u)\,du} - \phi_R(\xi_T^{x,t})\,\mathrm{e}^{-\int_t^T c(\xi_u^{x,t},u)\,du}\Big|\Big] \\ &\quad +\mathrm{E}\Big[\big|\phi(\xi_T^{x_n,t_n})\big|\,\mathrm{e}^{-\int_{t_n}^T c(\xi_u^{x_n,t_n},u)\,du}1_{\{\tau_n\wedge\tau<T\}}\Big] \\ &\quad +\mathrm{E}\Big[\big|\phi(\xi_T^{x,t})\big|\,\mathrm{e}^{-\int_t^T c(\xi_u^{x,t},u)\,du}1_{\{\tau_n\wedge\tau<T\}}\Big]\,. \end{aligned} \tag{10.27}$$

Applying Lebesgue's theorem, we first have hat

$$\phi_R(\xi_T^{x_n,t_n})\,\mathrm{e}^{-\int_{t_n}^T c(\xi_u^{x_n,t_n},u)\,du} \underset{n\to\infty}{\to} \phi_R(\xi_T^{x,t})\,\mathrm{e}^{-\int_t^T c(\xi_u^{x,t},u)\,du}$$

and then (recall that $\phi_R$ is bounded and $c\ge -K$) that

$$\lim_{n\to\infty}\mathrm{E}\Big(\Big|\phi_R(\xi_T^{x_n,t_n})\,\mathrm{e}^{-\int_{t_n}^T c(\xi_u^{x_n,t_n},u)\,du} - \phi_R(\xi_T^{x,t})\,\mathrm{e}^{-\int_t^T c(\xi_u^{x,t},u)\,du}\Big|\Big)=0\,.$$

It remains to show that the last two terms in (10.27) can be made arbitrarily small uniformly in $n$ for large $R$. To this end let us observe that

$$1_{\{\tau_n\wedge\tau<T\}} \le 1_{\{\tau_n<T\}} + 1_{\{\tau<T\}} = 1_{\{\sup_{t\le u\le T}|\xi_u^{x,t}|>R\}} + 1_{\{\sup_{t_n\le u\le T}|\xi_u^{x_n,t_n}|>R\}}\cdot$$

Thanks to Hölder's inequality (1.3) and using the upper bound $|\phi(x)| \le M(1+|x|^\lambda)$ for some $M > 0, \lambda > 0$, we have for every $z$, $s \ge 0$

$$\begin{aligned}
&\mathrm{E}\Big[|\phi(\xi_T^{z,s})|\,\mathrm{e}^{-\int_s^T c(\xi_u^{z,s},u)\,du} 1_{\{\sup_{s\le u\le T}|\xi_u^{z,s}|>R\}}\Big] \\
&\le M\mathrm{e}^{KT}\mathrm{E}\Big[\Big(1+\sup_{s\le u\le T}|\xi_u^{z,s}|^\lambda\Big)1_{\{\sup_{s\le u\le T}|\xi_u^{z,s}|>R\}}\Big] \\
&\le M_2\mathrm{E}\Big[1+\sup_{s\le u\le T}|\xi_u^{z,s}|^{2\lambda}\Big]^{1/2}\mathrm{P}\Big(\sup_{s\le u\le T}|\xi_u^{z,s}|>R\Big)^{1/2}
\end{aligned}$$

and the last quantity tends to zero as $R \to +\infty$, uniformly for $z, s$ in a compact set. Actually thanks to Proposition 9.1

$$\mathrm{E}\Big[1+\sup_{s\le u\le T}|\xi_u^{z,s}|^{2\lambda}\Big] < +\infty$$

and thanks to Remark 9.3

$$\mathrm{P}\Big(\sup_{s\le u\le T}|\xi_u^{z,s}|>R\Big) \underset{R\to+\infty}{\to} 0\,.$$

Hence the last two terms in (10.27) can be made uniformly arbitrarily small for large $R$.

The proof of the continuity of $u_2$ follows the same pattern. □

**Lemma 10.2** Let $\sigma$ be a $t$-stopping time such that $\sigma \le T$ a.s. Then

$$\begin{aligned}
u_1(x,t) &= \mathrm{E}^{x,t}\Big[\mathrm{e}^{-\int_t^\sigma c(X_s,s)\,ds}u_1(X_\sigma,\sigma)\Big]\,,\\
u_2(x,t) &= \mathrm{E}^{x,t}\Big[\int_t^\sigma f(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_u,u)\,du}\,ds\Big] + \mathrm{E}^{x,t}\Big[\mathrm{e}^{-\int_t^\sigma c(X_u,u)\,du}\,u_2(X_\sigma,\sigma)\Big]
\end{aligned}$$

We prove Lemma 10.2 later.

Let $\tau_R = \inf\{s; s \ge t, X_s \notin B_R\}$ be the first exit time from $B_R$ after time $t$. This is of course a stopping $t$-time (see the definition p. 160). Let $\tau = \tau_R \wedge T$ and let us deal first with $u_1$.

Lemma 10.2 applied to the $t$-stopping time $\tau$ gives

$$\begin{aligned}
u(x,t) &= u_1(x,t) - u_2(x,t)\\
&= \mathrm{E}^{x,t}\Big[\mathrm{e}^{-\int_t^\tau c(X_s,s)\,ds}u(X_\tau,\tau)\Big] - \mathrm{E}^{x,t}\Big[\int_t^\tau f(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_u,u)\,du}\,ds\Big]
\end{aligned}$$

$$= \mathrm{E}^{x,t}\big[u(X_{\tau_R}, \tau_R)\, \mathrm{e}^{-\int_t^{\tau_R} c(X_u,u)\,du} 1_{\{\tau_R<T\}}\big] + \mathrm{E}^{x,t}\big[\phi(X_T)\, \mathrm{e}^{-\int_t^T c(X_s,s)\,ds} 1_{\{\tau_R\ge T\}}\big]$$
$$-\mathrm{E}^{x,t}\Big[\int_t^{\tau_R\wedge T} f(X_s, s)\, \mathrm{e}^{-\int_t^s c(X_u,u)\,du}\, ds\Big] .$$

Comparing with the representation formula of Theorem 10.4 with $g = u$ this allows us to state that if

- $f$ and $c$ are Hölder continuous and
- the function $u$ is continuous,

then $u$ coincides with the solution of

$$\begin{cases} L_t w + \dfrac{\partial w}{\partial t} - cw = f & \text{on } B_R \times [0, T[ \\ w(x, T) = \phi(x) & \text{on } B_R \\ w(x, t) = u(x, t) & \text{on } \partial B_R \times [0, T] \end{cases}$$

and, in particular, is of class $C^{2,1}$. By the arbitrariness of $R$, $u$ is of class $C^{2,1}$ and a solution of (10.26).

We have proved the following

**Theorem 10.6** Let us assume that

- the coefficients $a$ and $b$ of $L_t$ are Lipschitz continuous on $\mathbb{R}^m \times [0, T]$ and moreover that for every $R > 0$ there exists a $\lambda_R > 0$ such that $\langle a(x, t)z, z\rangle \ge \lambda_R|z|^2$ for every $(x, t)$, $|x| \le R, 0 \le t \le T$, $z \in \mathbb{R}^m$.
- $\phi : \mathbb{R}^m \to \mathbb{R}$ is continuous and has polynomial growth.
- the real functions $f$ and $c$, defined on $\mathbb{R}^m \times [0, T]$, are locally Hölder continuous; $c$ is bounded below, $f$ has polynomial growth.

Then there exists a function $u$, continuous on $\mathbb{R}^m \times [0, T]$ and $C^{2,1}(\mathbb{R}^m \times [0, T[)$, which is a solution of

$$\begin{cases} L_t u + \dfrac{\partial u}{\partial t} - cu = f & \text{on } \mathbb{R}^m \times [0, T[ \\ u(x, T) = \phi(x) . \end{cases} \tag{10.28}$$

It is given by

$$u(x, t) = \mathrm{E}^{x,t}\big[\phi(X_T)\, \mathrm{e}^{-\int_t^T c(X_s,s)\,ds}\big] - \mathrm{E}^{x,t}\Big[\int_t^T f(X_s, s)\, \mathrm{e}^{-\int_t^s c(X_v,v)\,dv}\, ds\Big] ,$$

(continued)

**Theorem 10.6** (continued)
where $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^{x,s})_{x,s})$ is the canonical diffusion associated to $L_t$. Moreover, $u$ is the unique solution with polynomial growth of (10.28).

If, moreover, the diffusion coefficient $\sigma$ is bounded, then if $\phi$ and $f$ have an exponential growth, $u$ is the unique solution with exponential growth.

*Proof of Lemma 10.2* The proof consists in a stronger version of the strong Markov property. We shall first make the assumption that $\sigma$ takes a discrete set of values.

Taking the conditional expectation with respect to $\mathscr{M}^t_\sigma$, as the r.v. $\int_t^\sigma c(X_s, s)\,ds$ is $\mathscr{M}^t_\sigma$-measurable,

$$u_1(x,t) = \mathrm{E}^{x,t}\Big[\mathrm{e}^{-\int_t^\sigma c(X_s,s)\,ds}\mathrm{E}^{x,t}\Big(\phi(X_T)\,\mathrm{e}^{-\int_\sigma^T c(X_s,s)\,ds}\,|\,\mathscr{M}^t_\sigma\Big)\Big]\,. \qquad (10.29)$$

Let us denote by $s_1,\dots,s_m$ the possible values of $\sigma$. Then if $C \in \mathscr{M}^t_\sigma$ we have

$$\begin{aligned}\mathrm{E}^{x,t}\Big[1_C\,\phi(X_T)\,\mathrm{e}^{-\int_\sigma^T c(X_s,s)\,ds}\Big] &= \mathrm{E}^{x,t}\Big[\sum_{k=1}^m 1_{C\cap\{\sigma=s_k\}}\phi(X_T)\,\mathrm{e}^{-\int_\sigma^T c(X_s,s)\,ds}\Big]\\ &= \mathrm{E}^{x,t}\Big[\sum_{k=1}^m 1_{C\cap\{\sigma=s_k\}}\phi(X_T)\,\mathrm{e}^{-\int_{s_k}^T c(X_s,s)\,ds}\Big]\\ &= \mathrm{E}^{x,t}\Big[\sum_{k=1}^m 1_{C\cap\{\sigma=s_k\}}\mathrm{E}^{x,t}\Big(\phi(X_T)\,\mathrm{e}^{-\int_{s_k}^T c(X_s,s)\,ds}\,|\,\mathscr{M}^t_{s_k}\Big)\Big]\,.\end{aligned}$$

Note now that the r.v. $\phi(X_T)\,\mathrm{e}^{-\int_{s_k}^T c(X_s,s)\,ds}$ is $\mathscr{M}^{s_k}_\infty$-measurable. Hence by Proposition 6.1

$$\mathrm{E}^{x,t}\Big(\phi(X_T)\,\mathrm{e}^{-\int_{s_k}^T c(X_s,s)\,ds}\,|\,\mathscr{M}^t_{s_k}\Big) = \mathrm{E}^{X_{s_k},s_k}\Big(\phi(X_T)\,\mathrm{e}^{-\int_{s_k}^T c(X_s,s)\,ds}\Big) \qquad \mathrm{P}^{x,t}\text{-a.s.}$$

Hence, as

$$u_1(X_\sigma,\sigma) = \sum_{k=1}^m 1_{\{\sigma=s_k\}}u_1(X_{s_k},s_k) = \sum_{k=1}^m 1_{\{\sigma=s_k\}}\mathrm{E}^{X_{s_k},s_k}\Big(\phi(X_T)\,\mathrm{e}^{-\int_{s_k}^T c(X_s,s)\,ds}\Big)\,,$$

we have

$$\begin{aligned}\mathrm{E}^{x,t}\Big[1_C\,\phi(X_T)\,\mathrm{e}^{-\int_\sigma^T c(X_s,s)\,ds}\Big] &= \mathrm{E}^{x,t}\Big[\sum_{k=1}^m 1_{C\cap\{\sigma=s_k\}}\mathrm{E}^{X_{s_k},s_k}\Big(\phi(X_T)\,\mathrm{e}^{-\int_{s_k}^T c(X_s,s)\,ds}\Big)\Big]\\ &= \mathrm{E}^{x,t}\Big[\sum_{k=1}^m 1_{C\cap\{\sigma=s_k\}}u_1(X_{s_k},s_k)\Big] = \mathrm{E}^{x,t}\big[1_C\,u_1(X_\sigma,\sigma)\big]\end{aligned}$$

and therefore

$$\mathrm{E}^{x,t}\big(\phi(X_T)\,\mathrm{e}^{-\int_\sigma^T c(X_s,s)\,ds}\,|\,\mathscr{M}_\sigma^t\big) = u_1(X_\sigma,\sigma)$$

and going back to (10.29) we see that the first equality of Lemma 10.2 is proved for a discrete stopping time $\sigma$. We obtain the result for a general stopping time $\sigma \le T$ with the usual argument: let $(\sigma_n)_n$ be a sequence of discrete stopping times decreasing to $\sigma$. We have proved that

$$\mathrm{E}^{x,t}\big[\phi(X_T)\,\mathrm{e}^{-\int_t^T c(X_s,s)\,ds}\,|\,\mathscr{M}_{\sigma_n}^t\big] = \mathrm{e}^{-\int_t^{\sigma_n} c(X_s,s)\,ds}u_1(X_{\sigma_n},\sigma_n)\ . \tag{10.30}$$

Conditioning both sides with respect to $\mathscr{M}_\sigma^t$, as $\mathscr{M}_\sigma^t \subset \mathscr{M}_{\sigma_n}^t$,

$$\mathrm{E}^{x,t}\big[\phi(X_T)\,\mathrm{e}^{-\int_t^T c(X_s,s)\,ds}\,|\,\mathscr{M}_\sigma^t\big] = \mathrm{E}^{x,t}\big[\mathrm{e}^{-\int_t^{\sigma_n} c(X_s,s)\,ds}u_1(X_{\sigma_n},\sigma_n)\,|\,\mathscr{M}_\sigma^t\big]\ .$$

Now clearly

$$\mathrm{e}^{-\int_t^{\sigma_n} c(X_s,s)\,ds}u_1(X_{\sigma_n},\sigma_n) \underset{n\to\infty}{\to} \mathrm{e}^{-\int_t^{\sigma} c(X_s,s)\,ds}u_1(X_{\sigma},\sigma)$$

and thanks to (10.30) the sequence on the left-hand side above is uniformly integrable so that, $\mathrm{P}^{x,t}$-a.s.,

$$\begin{aligned}\mathrm{E}^{x,t}\big[\mathrm{e}^{-\int_t^{\sigma_n} c(X_s,s)\,ds}u_1(X_{\sigma_n},\sigma_n)\,|\,\mathscr{M}_\sigma^t\big] &\underset{n\to\infty}{\to} \mathrm{E}^{x,t}\big[\mathrm{e}^{-\int_t^{\sigma} c(X_s,s)\,ds}u_1(X_{\sigma},\sigma)\,|\,\mathscr{M}_\sigma^t\big]\\ &= \mathrm{e}^{-\int_t^{\sigma} c(X_s,s)\,ds}u_1(X_{\sigma},\sigma)\end{aligned}$$

which proves the first equality of Lemma 10.2. The second one is proved along the same lines. □

*Example 10.3* Let us consider the problem

$$\begin{cases}\dfrac{1}{2}\triangle u(x,t) + \dfrac{\partial u}{\partial t}(x,t) - \lambda|x|^2 u(x,t) = 0 & \text{on } \mathbb{R}^m \times [0,T[\\ u(x,T) = 1\ . \end{cases}$$

By Theorem 10.6 a solution is therefore given by

$$u(x,t) = \mathrm{E}^{x,t}\big[\mathrm{e}^{-\lambda\int_t^T |X_s|^2\,ds}\big]$$

(continued)

*Example 10.3* (continued)
(here $\phi \equiv 1$, $f \equiv 0$), where $X = (\Omega, \mathscr{F}, (\mathscr{F}_t^s)_{t\geq s}, (X_t)_t, (\mathrm{P}^{x,t})_{x,t})$ is a realization of the diffusion associated at generator $\frac{1}{2}\triangle$, i.e.

$$u(x,t) = \mathrm{E}\Big[\mathrm{e}^{-\lambda \int_0^{T-t} |B_s+x|^2\, ds}\Big]\,, \tag{10.31}$$

where $B$ is a Brownian motion. This solution is unique among those with a polynomial growth. In Exercise 12.8 the expectation on the right-hand side of (10.31) will be computed giving the explicit solution

$$u(x,t) = \cosh\big(\sqrt{2\lambda}\,(T-t)\big)^{-m/2} \exp\Big[-\frac{\sqrt{2\lambda}\,|x|^2}{2}\tanh\big(\sqrt{2\lambda}\,(T-t)\big)\Big].$$

## 10.5 The density of the transition function, the backward equation

In the examples of diffusion processes we have met so far (all of them being $\mathbb{R}^m$-valued for some $m \geq 1$), the transition probability $p$ had a density, i.e. there existed a positive measurable function $q(s,t,x,y)$ such that

$$p(s,t,x,dy) = q(s,t,x,y)\,dy\,,$$

$dy$ denoting the Lebesgue measure of $\mathbb{R}^m$. We now look for general conditions ensuring the existence of a density.

**Definition 10.1** A *fundamental solution* of the Cauchy problem on $\mathbb{R}^m\times[0,T]$ associated to $L_t$ is a function $\Gamma(s,t,x,y)$ defined for $x, y \in \mathbb{R}^m$, $0 \leq s < t \leq T$, such that if, for every continuous compactly supported function $\phi$,

$$v(x,s) = \int \Gamma(s,t,x,y)\phi(y)\,dy\,,$$

then $v$ is $C^{2,1}(\mathbb{R}^m \times [0,t[)$, is bounded and is a solution of

$$\begin{cases} L_s v + \dfrac{\partial v}{\partial s} = 0 & \text{on } \mathbb{R}^m \times [0,t[ \\ \lim_{s\to t-} v(x,s) = \phi(x)\,. \end{cases} \tag{10.32}$$

Let us assume that the operator $L_t$ satisfies the assumptions of Theorem 10.6. If the transition function $p(s,t,x,dy)$ has a density $q(s,t,x,y)$, then this is a fundamental solution. Actually, by Theorem 10.6 there is a unique bounded solution of (10.32), given by

$$v(x,s) = \mathrm{E}^{x,s}[\phi(X_t)] = \int \phi(y)\,p(s,t,x,dy) = \int \phi(y)q(s,t,x,y)\,dy\;.$$

Therefore $\Gamma = q$ is a fundamental solution. Conversely, if a fundamental solution $\Gamma$ exists, by the same argument, necessarily

$$\int \phi(y)\,p(s,t,x,dy) = \int \phi(y)\Gamma(s,t,x,y)\,dy$$

for every compactly supported continuous function $\phi$. This implies that the transition function has density $\Gamma$.

We still have to investigate conditions under which the transition function has a density, or equivalently, the differential generator has a fundamental solution.

Let us make the following assumptions

- $\langle a(x,t)\,z,z\rangle \geq \lambda|z|^2$ for some $\lambda > 0$ and for every $z \in \mathbb{R}^m$, $(x,t) \in \mathbb{R}^m \times [0,T]$.
- $a$ and $b$ are continuous and bounded for $(x,t) \in \mathbb{R}^m \times [0,T]$. Moreover, $a$ is continuous in $t$ uniformly in $x$.
- $a$ and $b$ are Lipschitz continuous in $x$ uniformly for $t \in [0,T]$.

**Theorem 10.7** Under the assumptions above there exists a unique fundamental solution $\Gamma$ of the Cauchy problem on $\mathbb{R}^m \times [0,T]$ associated to $L_t$. Moreover, it satisfies the inequalities

$$\begin{aligned}|\Gamma(s,t,x,y)| &\leq C(t-s)^{-m/2}\exp\Big[-c\frac{|x-y|^2}{t-s}\Big]\\ \Big|\frac{\partial\Gamma}{\partial x_i}(s,t,x,y)\Big| &\leq C(t-s)^{-(m+1)/2}\exp\Big[-c\frac{|x-y|^2}{t-s}\Big]\,,\end{aligned} \tag{10.33}$$

where $C$ and $c$ are positive constants. Furthermore the derivative of $\Gamma$ with respect to $s$ and its second derivatives with respect to $x$ exist and are continuous for $t > s$. Finally, as a function of $(x,s)$, $s < t$, $\Gamma$ is a solution of

$$L_s\Gamma + \frac{\partial\Gamma}{\partial s} = 0\;. \tag{10.34}$$

Theorem 10.7 is a classical application of the parametrix method. See Levi (1907) and also Friedman (1964, Chapter 1), Friedman (1975, Chapter 6) or Chapter 4 of Azencott et al. (1981). Note, however, that Assumption (H) is only a sufficient condition. In Exercise 9.25 b) the transition density (and therefore the fundamental solution) is computed explicitly for an operator $L$ whose second-order matrix coefficient $a(x)$ is not elliptic for any value of $x$.

As we have identified the transition density $q(s,t,x,y)$ and the fundamental solution, $q$ satisfies, as a function of $s,x$ (the "backward variables"), the *backward equation*

$$L_s q + \frac{\partial q}{\partial s} = 0\,. \tag{10.35}$$

## 10.6 Construction of the solutions of the Dirichlet problem

In this section we construct a solution of the Dirichlet problem

$$\begin{cases} u \in C^2(D) \cap C(\overline{D}) & \\ \frac{1}{2}\triangle u = 0 & \text{on } D \\ u_{|\partial D} = \phi & \end{cases} \tag{10.36}$$

without using the existence Theorem 10.1. In particular, we shall point out the hypotheses that must be must satisfied by the boundary $\partial D$ in order to have a solution; the $C^2$ hypothesis for $\partial D$ will be much weakened.

Let $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ be the canonical realization of an $m$-dimensional Brownian motion, $D$ a connected bounded open set of $\mathbb{R}^m$, $\tau$ the exit time from $D$. Let

$$u(x) = \mathrm{E}^x[\phi(X_\tau)]\,. \tag{10.37}$$

Let us recall that in Example 6.3 we have proved the following result.

**Proposition 10.3** For every bounded Borel function $\phi$ on $\partial D$, $u$ is harmonic in $D$.

(10.37) therefore provides a candidate solution of the Dirichlet problem (10.36). We see now which conditions on $\phi$ and $\partial D$ are needed for $u$ to be continuous on $\overline{D}$ and verify the condition at the boundary. It is clear that, for $z \in \partial D$, $u(z) = \phi(z)$ and in order for $u$ to be continuous at $x$ it will be necessary to show that if $x \in D$ is a point near $z$, then the law of $X_\tau$ with respect to $\mathrm{P}^x$ is concentrated near $z$. The investigation of this type of property is the object of the remainder of this section. It is suggested at this time to go back and have a look at Example 6.4.

**Proposition 10.4** Let $g : \mathbb{R}^m \to \mathbb{R}$ be a bounded Borel function. Then the function $\psi(x) = \mathrm{E}^x[g(X_t)]$ is $C^\infty$, for every fixed $t > 0$.

Proposition 10.4 is just a repetition of Remark 6.4.

Let $\tau'(\omega) = \inf\{t > 0; X_t \notin D\}$. $\tau'$ is not the exit time from $D$, because of the $>$ instead of $\geq$. Of course, $\tau' = \tau$ if $X_0 \in D$ or $X_0 \in \overline{D}^c$; $\tau$ and $\tau'$ can be different if the initial point of the process lies on the boundary $\partial D$, as then $\tau = 0$ (as $\partial D \subset D^c$), whereas we may have $\tau' > 0$ (if the path enters $D$ immediately).

It is immediate that $\tau'$ is a stopping time of the filtration $(\mathscr{M}_{t+})_t$. Indeed

$$\{\tau' = 0\} = \{\text{there exists a sequence } (t_n)_n \text{ of times with } t_n > 0,\ t_n \searrow 0 \text{ and } X_{t_n} \notin \ \mathrm{D}\}$$

and this is an event of $\mathscr{M}_\varepsilon$ for every $\varepsilon > 0$. Therefore $\{\tau' = 0\} \in \mathscr{M}_{0+}$. Moreover, for $u > 0$,

$$\{\tau' \leq u\} = \{\tau \leq u\} \setminus \{\tau = 0\} \cup \{\tau' = 0\} \in \mathscr{M}_u \subset \mathscr{M}_{u+}.$$

**Definition 10.2** $x \in \partial D$ is said to be a regular point if $\mathrm{P}^x(\tau' = 0) = 1$.

It is immediate that, as Brownian motion is a continuous Feller process, by Blumenthal's 0–1 law (Proposition 6.23), $\mathrm{P}^x(\tau' = 0)$ can take the values 0 or 1 only.

**Lemma 10.3** For every $u > 0$ the function $x \mapsto \mathrm{P}^x(\tau' > u)$ is upper semicontinuous.

*Proof* Let $\tau_s(\omega) = \inf\{t \geq s; X_t \notin D\}$. Then $\tau_s \searrow \tau'$ as $s \searrow 0$, therefore $\mathrm{P}^x(\tau_s > u)$ decreases to $\mathrm{P}^x(\tau' > u)$ as $s \searrow 0$. We now just need to prove that $x \mapsto \mathrm{P}^x(\tau_s > u)$ is a continuous function, as the lower envelope of a family of continuous functions is known to be upper semicontinuous. Using translation operators we can write $\tau_s = s + \tau \circ \theta_s$: in fact $\tau \circ \theta_s(\omega)$ is the time between time $s$ and the first exit from $D$ of $\omega$

after time $s$. By the strong Markov property, Proposition 6.2,

$$\begin{aligned}\mathrm{P}^x(\tau_s > u) = \mathrm{P}^x(\tau\circ\theta_s > u-s) = \mathrm{E}^x[1_{\{\tau>u-s\}}\circ\theta_s]\\ = \mathrm{E}^x[\mathrm{E}(1_{\{\tau>u-s\}}\circ\theta_s\,|\,\mathscr{F}_s)] = \mathrm{E}^x\big[\mathrm{P}^{X_s}(\tau > u-s)\big] = \mathrm{E}^x[g(X_s)]\ ,\end{aligned}$$

where $g(x) = \mathrm{P}^x(\tau > u - s\}$. Therefore $x \mapsto \mathrm{P}^x(\tau_s > u\}$ is even $C^\infty$ by Proposition 10.4. □

**Lemma 10.4** If $z \in \partial D$ is a regular point and $u > 0$, then

$$\lim_{\substack{x\to z\\ x\in D}} \mathrm{P}^x(\tau > u) = 0\ .$$

*Proof* If $x \in D$, then $\mathrm{P}^x(\tau > u) = \mathrm{P}^x(\tau' > u)$ and by Lemma 10.3

$$0 \le \varlimsup_{\substack{x\to z\\ x\in D}} \mathrm{P}^x(\tau > u) = \varlimsup_{\substack{x\to z\\ x\in D}} \mathrm{P}^x(\tau' > u) \le \mathrm{P}^z(\tau' > u) = 0\ .$$

□

The following lemma states that if $z \in \partial D$ is a regular point, then, starting from a point $x \in D$ near to $z$, the Brownian motion exits from $D$ mainly at points near to $z$.

**Lemma 10.5** If $z \in \partial D$ is a regular point, then for every $\varepsilon > 0$ and for every neighborhood $V$ of $z$ there exists a neighborhood $W$ of $z$ such that, if $x \in D \cap W$, $\mathrm{P}^x(X_\tau \notin V \cap \partial D) \le \varepsilon$.

*Proof* By the Doob and Chebyshev inequalities

$$\mathrm{P}^x\Big(\sup_{0\le s\le t} |X_s - x| \ge \frac{\alpha}{2}\Big) \le \frac{4}{\alpha^2}\,\mathrm{E}^x\Big[\sup_{0\le s\le t} |X_s - x|^2\Big] \le \frac{16m}{\alpha^2}\,\mathrm{E}^x[|X_t - x|^2] = \frac{16tm}{\alpha^2}\,\cdot$$

Let $\alpha > 0$ be such that $B_\alpha(z) \subset V$; then if $|x - z| < \frac{\alpha}{2}$

$$\begin{aligned}\mathrm{P}^x(|X_\tau - z| \ge \alpha) \le \mathrm{P}^x(\tau > u) + \mathrm{P}^x(|X_\tau - x| \ge \tfrac{\alpha}{2}, \tau \le u)\\ \le \mathrm{P}^x(\tau > u) + \mathrm{P}^x\Big(\sup_{0\le s\le u} |X_s - x| \ge \frac{\alpha}{2}\Big) \le \mathrm{P}^x(\tau > u) + \frac{16um}{\alpha^2}\,\cdot\end{aligned}$$

Now just choose $u$ small so that $\frac{16um}{\alpha^2} < \frac{\varepsilon}{2}$ and apply Lemma 10.4. □

**Proposition 10.5** If $z \in \partial D$ is a regular point and $\phi : \partial D \to \mathbb{R}$ is continuous at $z$, then the function $u(x) = \mathrm{E}^x[\phi(X_\tau)]$ satisfies $\lim_{\substack{x \to z \\ x \in D}} u(x) = \phi(z)$.

*Proof* Let $\eta > 0$ be such that $|\phi(y) - \phi(z)| < \frac{\varepsilon}{2}$ for every $y \in \partial D$ with $|y - z| < \eta$; then

$$
\begin{aligned}
|u(x) - \phi(z)| &\le \mathrm{E}^x[|\phi(X_\tau) - \phi(z)|] \\
&\le \mathrm{E}^x\big[|\phi(X_\tau) - \phi(z)| 1_{\{|X_\tau - z| < \eta\}}\big] + \mathrm{E}^x\big[|\phi(X_\tau) - \phi(z)| 1_{\{|X_\tau - z| \ge \eta\}}\big] \\
&\le \frac{\varepsilon}{2} + 2\|\phi\|_\infty \mathrm{P}^x(|X_\tau - z| \ge \eta) \;.
\end{aligned}
$$

Now, by Lemma 10.5, there exists a neighborhood $V$ of $z$ such that if $x \in V$ then $\mathrm{P}^x(|X_\tau - z| \ge \eta) \le \varepsilon$, which allows us to conclude the proof. □

The following statement is now a consequence of Propositions 10.3 and 10.5.

**Theorem 10.8** Let $D \subset \mathbb{R}^m$ be a bounded open set such that every point in $\partial D$ is regular. Then if $\phi$ is a continuous function on $\partial D$, the function $u(x) = \mathrm{E}^x[\phi(X_\tau)]$ is in $C^2(D) \cap C(\overline{D})$ and is a solution of

$$
\begin{cases}
\frac{1}{2}\triangle u = 0 & \text{on } D \\
u_{|\partial D} = \phi \;.
\end{cases}
$$

It is now natural to investigate conditions for the regularity of points of $\partial D$.

**Proposition 10.6 (Cone property)** Let $z \in \partial D$ and let us assume that there exist an open cone $C$ with vertex $z$ and a neighborhood $V$ of $z$ such that $C \cap V \subset D^c$. Then $z$ is regular.

*Proof* Let $\alpha > 0$ be such that the ball $B_\alpha(z)$ of radius $\alpha$ and centered at $z$ is contained in $V$; then, as $\{\tau' > t\} \searrow \{\tau' > 0\}$, we have

$$
\mathrm{P}^z(\tau' > 0) = \lim_{t \to 0+} \mathrm{P}^z(\tau' > t)
$$

but for every $t > 0$

$$\mathrm{P}^z(\tau' > t) \le \mathrm{P}^z(X_t \notin C \cap B_\alpha(z)) = 1 - \mathrm{P}^z(X_t \in C \cap B_\alpha(z))$$

so that

$$\mathrm{P}^z(\tau' > 0) \le 1 - \varliminf_{t \to 0+} \mathrm{P}^z(X_t \in C \cap B_\alpha(z)) \ . \tag{10.38}$$

But if $C_1 = C - z$, $C_1$ is an open cone with vertex at the origin, which is invariant under multiplication by a positive constant, hence $C_1 = C_1/\sqrt{t}$ and

$$\begin{aligned} \mathrm{P}^z(X_t \in C \cap B_\alpha(z)) &= \mathrm{P}^0(X_t \in C_1 \cap B_\alpha) = \mathrm{P}^0(\sqrt{t}X_1 \in C_1 \cap B_\alpha) \\ &= \mathrm{P}^0(X_1 \in C_1 \cap B_{\alpha/\sqrt{t}}) = \frac{1}{(2\pi)^{m/2}} \int_{C_1 \cap B_{\alpha/\sqrt{t}}} \mathrm{e}^{-|y|^2/2}\, dy \ . \end{aligned}$$

Taking the limit as $t \to 0+$ in (10.38) we find

$$\mathrm{P}^z(\tau' > 0) \le 1 - \frac{1}{(2\pi)^{m/2}} \int_{C_1} \mathrm{e}^{-|y|^2/2}\, dy < 1 \ .$$

As $\mathrm{P}^z(\tau' > 0)$ can be equal to 0 or 1 only, we must have $\mathrm{P}^z(\tau' > 0) = 0$ and $z$ is a regular point. □

By Proposition 10.6 all the points of the boundary of an open convex set are therefore regular, as in this case we can choose a cone which is a whole half-plane. One easily sees that all the points at which the boundary is $C^2$ are regular. The cone condition requires that the boundary does not have inward pointing thorns, see Fig. 10.3.

The cone hypothesis of Proposition 10.6 is only a sufficient condition for the regularity of the boundary, as we see in the following example.


**Fig. 10.3** The domain on the left-hand side enjoys the cone property at $z$, the one on the right-hand side doesn't

*Example 10.4* Let $D \subset \mathbb{R}^2$ be a domain formed by a disc minus a segment $S$ as in Fig. 10.5. Let us prove that all the points of $\partial D$ are regular. This fact is obvious for the points of the circle so we need to consider only the points of the segment $S$. We can assume that $S$ belongs to the axis $y = 0$ and that one of its ends is the origin. Let us prove that the origin is a regular point.

Let $V$ be a ball centered at 0 and with a small enough radius and let

$$V_S = V \cap S, \qquad V_1 = V \cap (\{y = 0\} \backslash S), \qquad V_0 = V_S \cup V_1$$

(see Fig. 10.4). Consider the three stopping times

$$\tau'' = \inf\{t > 0; X_t \in V_S\}, \quad \tau_1 = \inf\{t > 0; X_t \in V_1\}, \quad \tau_0 = \inf\{t > 0; X_t \in V_0\}.$$

The idea of the proof is simple: starting at 0, by the Iterated Logarithm Law applied to the component $X_2$ of the Brownian motion $X$, there exists a sequence of times $(t_n)_n$ decreasing to 0 such that $X_2(t_n) = 0$ for every $n$. Now we just need to show that with positive probability, possibly passing to a subsequence, we have $X_1(t_n) \in V_S$ for every $n$. We have

1) $\{\tau'' = 0\} \subset \{\tau' = 0\}$.
2) $\{\tau_0 = 0\} = \{\tau_1 = 0\} \cup \{\tau'' = 0\}$. In one direction the inclusion is obvious. Conversely, if $\omega \in \{\tau_0 = 0\}$ then there exists a sequence $(t_n)_n$, decreasing to 0 and such that $X_{t_n} \in V_0$. As $V_0 = V_S \cup V_1$, there exists a subsequence $(t_{n_k})_k$ such that $X_{t_{n_k}} \in V_1$ for every $k$, or $X_{t_{n_k}} \in V_S$ for every $k$; therefore $\omega \in \{\tau_1 = 0\}$ or $\omega \in \{\tau'' = 0\}$, from which we derive the opposite inclusion.
3) $\mathrm{P}^0(\tau'' = 0) = \mathrm{P}^0(\tau_1 = 0)$, which is obvious by symmetry.
4) $\mathrm{P}^0(\tau' = 0) \geq \frac{1}{2}\,\mathrm{P}^0(\tau_0 = 0)$. From 2) and 3) we have

$$\mathrm{P}^0(\tau_0 = 0) \leq \mathrm{P}^0(\tau'' = 0) + \mathrm{P}^0(\tau_1 = 0) = 2\,\mathrm{P}^0(\tau'' = 0) \leq 2\,\mathrm{P}^0(\tau' = 0).$$

5) $\mathrm{P}^0(\tau_0 = 0) \geq \mathrm{P}^0(\sup_{0 \leq t \leq \alpha} |X_1(t)| < \eta) > 0$ where $\eta$ is the radius of $V$ and $\alpha$ is any number $> 0$. Indeed $\{\tau_0 = 0\}$ contains the event

$$\Big\{\overline{\lim_{t\to 0+}} \frac{X_2(t)}{\big(2t \log\log \frac{1}{t}\big)^{1/2}} = 1,\ \underline{\lim_{t\to 0+}} \frac{X_2(t)}{\big(2t \log\log \frac{1}{t}\big)^{1/2}} = -1,\ \sup_{0\leq t\leq \alpha} |X_1(t)| < \eta\Big\}$$

and this event is equal to $\big\{\sup_{0 \leq t \leq \alpha} |X_1(t)| < \eta\big\}$ by the Iterated Logarithm Law.

From 4) and 5), we have $\mathrm{P}^0(\tau' = 0) > 0$ and by Blumenthal's 0–1 law 0 is a regular point (Fig. 10.5).

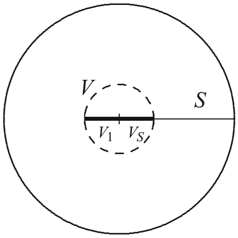


**Fig. 10.4** The different objects appearing in Example 10.4

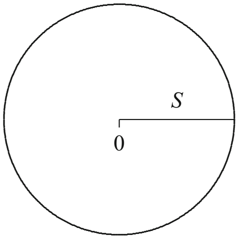


**Fig. 10.5** The domain $D$ is the interior of the disc minus the segment $S$

## Exercises

**10.1** (p. 591)

a) Let $D \subset \mathbb{R}^m$ be a bounded open set with a $C^2$ boundary. Let $L$ be the differential operator

$$L = \frac{1}{2} \sum_{i,j=1}^{m} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{m} b_i(x) \frac{\partial}{\partial x_i}$$

and assume that the coefficients $a$ and $b$ satisfy the conditions a) and b) stated p. 309 before Theorem 10.1. Let $u \in C^2(D) \cap C(\overline{D})$ be the solution of

$$\begin{cases} Lu = -1 & \text{on } D \\ u_{|\partial D} = 0\ . \end{cases} \tag{10.39}$$

Let $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ be the canonical diffusion associated to $L$ and $\tau$ the exit time from $D$. Prove that

$$u(x) = \mathrm{E}^x(\tau)\ .$$

b) Let $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ be the canonical realization of an $m$-dimensional Brownian motion. Let $\tau$ be the exit time from the sphere of radius 1.

Solve (10.39) explicitly and prove that

$$E^x(\tau) = \frac{1}{m}(1 - |x|^2) \cdot$$

Compare the result of b) with Exercise 5.10 c).

**10.2** (p. 592) Let $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (P^x)_x)$ be the canonical diffusion associated to the SDE in dimension 1

$$dX_t = b(X_t)\,dt + dB_t \,.$$

Let $0 < a < b$ and let us assume that $b(x) = \frac{\delta}{x}$ for $a \le x \le b$ and that, on $\mathbb{R}$, $b$ satisfies Assumption (A'). Let $\tau$ be the exit time of $X$ from $]a, b[$.

a) Show that $\tau < +\infty$ $P^x$-a.s. for every $x$.
b) Prove that, if $\delta \neq \frac{1}{2}$ and for $a < x < b$,

$$P^x(X_\tau = b) = \frac{1 - (\frac{a}{x})^\lambda}{1 - (\frac{a}{b})^\lambda} \tag{10.40}$$

with $\lambda = 2\delta - 1$.
c) Let $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (P^x)_x)$ be the canonical realization of an $m$-dimensional Brownian motion with $m \ge 3$ and let $D = \{a < |x| < b\}$. Let us denote by $\sigma$ the exit time of $X$ from the annulus $D$. What is the value of $P^x(|B_\sigma| = b)$ for $x \in D$? How does this probability behave as $m \to \infty$?

Go back to Exercise 8.24 for the SDE satisfied by the process $\xi_t = |X_t - x|$.

**10.3** (p. 593) Let $a, b > 0$, $\sigma > 0$, $\mu \in \mathbb{R}$.

a) Let $\xi^x$ be the solution of the SDE

$$\begin{aligned} d\xi_t &= -\mu\,dt + \sigma\,dB_t \\ \xi_0 &= x \end{aligned}$$

and $\tau$ the exit time of $\xi^x$ from the interval $]-a, b[$.
a1) Prove that $\tau < +\infty$ a.s. whatever the starting point $x$.
a1) What is the generator, $L$, of the diffusion $\xi$? Compute $P(\xi^0_\tau = b)$ ($\xi^0$ is the solution starting at $x = 0$).
b) Let $\eta^x$ be the solution of the SDE

$$\begin{aligned} d\eta_t &= -\frac{\mu}{1 + \eta_t^2}\,dt + \frac{\sigma}{\sqrt{1 + \eta_t^2}}\,dB_t \\ \eta_0 &= x \,. \end{aligned}$$

Let now $\tau$ be the exit time of $\eta^x$ from the interval $]-a, b[$.

b1) What is the generator of the diffusion $\eta$? Can you say that $\tau < +\infty$ a.s.? Whatever the starting point $x$?
b2) Compute $\mathrm{P}(\eta^0_\tau = b)$ ($\eta^0$ is the solution starting at $x = 0$).

Compare the result of a) with Exercise 5.20.

**10.4** (p. 594) Let $B$ be a two-dimensional Brownian motion, $\Gamma$ the circle of radius 1, $x = (\frac{1}{2}, 0)$ and $\tau$ the exit time of $B$ from the ball of radius 1. If $\Gamma' \subset \Gamma$ denotes the set of the points of the boundary with a positive abscissa, compute $\mathrm{P}^x(B_\tau \in \Gamma')$ (Fig. 10.6).

See Example 10.1.

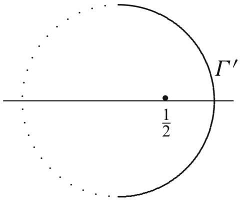


**Fig. 10.6** The starting point and the piece of boundary of Exercise 10.4

**10.5** (p. 595) (The Laplace transform of the exit time, to be compared with Exercise 5.32)

a) Let $X$ be the canonical diffusion on $\mathbb{R}^m$ associated to the generator

$$L = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x)\frac{\partial}{\partial x_i}$$

that we assume to be uniformly elliptic and satisfying Assumption (A'). If $D \subset \mathbb{R}^m$ is a bounded open set having a $C^2$ boundary and $\theta \in \mathbb{R}$, let $u$ (if it exists) be a solution of

$$\begin{cases} Lu + \theta u = 0 & \text{on } D \times [0, T[ \\ u_{\partial D} = 1\,. \end{cases} \tag{10.41}$$

We know that, if $\theta \le 0$, a unique solution exists by Theorem 10.1 and is given by $u(x) = \mathrm{E}^x[\mathrm{e}^{\theta\tau}]$. Let, for every $\varepsilon > 0$, $D_\varepsilon$ be an open set such that $\overline{D}_\varepsilon \subset D$ and $\operatorname{dist}(\partial D_\varepsilon, \partial D) \le \varepsilon$ and let us denote by $\tau, \tau_\varepsilon$ the respective exit times of $X$ from $D, D_\varepsilon$.
a1) Show that if $M_t = \mathrm{e}^{\theta t}u(X_t)$ then, for every $\varepsilon > 0$, $(M_{t\wedge\tau_\varepsilon})_t$ is a martingale.
a2) Prove that if $u \ge 0$ then

$$u(x) = \mathrm{E}^x[\mathrm{e}^{\theta\tau}]\,.$$

b) Let $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ be the canonical realization of a real Brownian motion and $D = ]-a, a[$, $a > 0$.

b1) Prove that

$$\mathrm{E}^x[\mathrm{e}^{\theta\tau}] = \frac{\cos(\sqrt{2\theta}x)}{\cos(\sqrt{2\theta}a)}, \qquad \theta < \frac{\pi^2}{8a^2}$$

and $\mathrm{E}^x[\mathrm{e}^{\theta\tau}] = +\infty$ for $\theta \geq \frac{\pi^2}{8a^2}$.

b2) Deduce that the r.v. $\tau$ *is not* bounded but that there exist numbers $\beta > 0$ such that $\mathrm{P}^x(\tau > R) \leq \mathrm{const} \cdot \mathrm{e}^{-\beta R}$ and determine them.

b3) Compute $\mathrm{E}^x[\tau]$.

b) Note that $u(x) = \mathrm{E}^x[M_{t \wedge \tau_\varepsilon}]$ for every $t$ and prove first, using Fatou's lemma, that the r.v. $\mathrm{e}^{\theta\tau}$ is $\mathrm{P}^x$-integrable.

**10.6** (p. 596) Let us consider the problem

$$\begin{cases} \dfrac{1}{2}\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial u}{\partial t} = 0 & \text{on } \mathbb{R} \times [0, T[ \\ u(x, T) = \phi(x)\ . \end{cases} \tag{10.42}$$

a) Find a solution for $\phi(x) = x^2$. What can be said about uniqueness?

b) Show that if $\phi$ is a polynomial, then the unique solution of (10.42) having polynomial growth is a polynomial in $x, t$.

**10.7** (p. 597)

a) Find a solution of the problem

$$\begin{cases} \dfrac{\sigma^2}{2} x^2 \dfrac{\partial^2 u}{\partial x^2}(x,t) + bx \dfrac{\partial u}{\partial x}(x,t) + \dfrac{\partial u}{\partial t}(x,t) = 0 & \text{on } \mathbb{R} \times [0, T[ \\ u(x, T) = x\ . \end{cases} \tag{10.43}$$

b) What if the boundary condition was replaced by $u(x, T) = x^2$?

**10.8** (p. 597) Let

$$L = \frac{\sigma^2}{2} \sum_{i=1}^m \frac{\partial^2}{\partial x_i^2} + \lambda \sum_{i=1}^m x_i \frac{\partial}{\partial x_i} \cdot$$

a) Find a solution of the problem

$$\begin{cases} Lu + \dfrac{\partial u}{\partial t} = 0 & \text{on } \mathbb{R}^m \times [0, T[ \\ u(x, T) = \cos(\langle \theta, x \rangle) , \end{cases} \tag{10.44}$$

where $\theta \in \mathbb{R}^m$.

b) What can be said of $x \mapsto u(x, 0)$ as $T \to +\infty$?

**10.9** (p. 598) Compute the fundamental solution of the Cauchy problem (in dimension 1) of the operator

$$L = \frac{1}{2} a x^2 \frac{\partial^2}{\partial x^2} + bx \frac{\partial}{\partial x} ,$$

where $a > 0, b \in \mathbb{R}$.

**10.10** (p. 599) Let $\xi^{x,s}$ be the solution of the SDE in dimension 1

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\xi_t \, dt + \sigma(\xi_t, t)\xi_t \, dB_t \\ \xi_s &= x , \end{aligned}$$

where $b$ and $\sigma$ are continuous functions in $x, s$, Lipschitz continuous in $x$ and *bounded*. Let us assume $\sigma(x, t) > 0$ for every $x, t$.

a) Assume $x > 0$. Show that the process $\eta_t^{x,s} = \log(\xi_t^{x,s})$ is defined for every $s > 0$ and is the solution of an SDE to be determined. Deduce that $\xi_t > 0$ for every $t$ a.s.

b) Let $\phi : \mathbb{R} \to \mathbb{R}$ be a continuous function having polynomial growth and $f, c$ functions $\mathbb{R}^m \times [0, T] \to \mathbb{R}$ measurable in $x, t$ and locally Hölder continuous in $x$, such that $f$ has a polynomial growth as a function of $x$ and $c$ is bounded below. Let

$$u(x, t) = \mathrm{E}\Big[\phi(\xi_T^{x,t}) \, \mathrm{e}^{-\int_t^T c(\xi_v^{x,t}, v) \, dv}\Big] - \mathrm{E}\Big[\int_t^T f(\xi_s^{x,t}, s) \, \mathrm{e}^{-\int_t^s c(\xi_v^{x,t}, v) \, dv} \, ds\Big]$$

and let $L_t$ be the differential operator

$$L_t = \frac{1}{2} \sigma(x, t)^2 x^2 \frac{\partial^2}{\partial x^2} + b(x, t) x \frac{d}{dx} \cdot$$

Show that $u$ is a solution of

$$\begin{cases} L_t u + \dfrac{\partial u}{\partial t} - cu = f & \text{on } \mathbb{R}^+ \times [0, T[ \\ u(x, T) = \phi(x) . \end{cases} \tag{10.45}$$

- Note that in b) Theorem 10.6 cannot be applied directly as $L_t$ does not satisfy all the required assumptions (the diffusion coefficient vanishes at 0).

b) Trace back to the diffusion $\eta$ introduced in a) and apply the Feynman–Kac formula.

**10.11** (p. 600) Let $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t^s)_t, (X_t)_t, (\mathrm{P}^{x,t})_{x,t})$ be the canonical realization of an $m$-dimensional Brownian motion and $\phi : \mathbb{R}^m \to \mathbb{R}$ a bounded Borel function. Let

$$u(x,t) = \mathrm{E}^{x,t}[\phi(X_T)] \ .$$

a) Show that $u$ is $C^\infty(\mathbb{R}^m \times [0,T[)$.
b) Show that $u(x,s) = \mathrm{E}^{x,s}[u(X_t,t)]$ for every $t, s \le t < T$.
c) Show that $u$ is a solution of

$$\frac{1}{2}\triangle u + \frac{\partial u}{\partial t} = 0 \qquad \text{on } \mathbb{R}^m \times [0,T[ \tag{10.46}$$

$$\lim_{t\to T-} u(x,t) = \phi(x) \qquad \text{for every } x \text{ of continuity for } \phi \ . \tag{10.47}$$

d) For $m = 1$, find a solution of (10.46) such that

$$\lim_{t\to T-} u(x,t) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0 \ . \end{cases}$$

What is the value of $\lim_{t\to T-} u(0,t)$?

**10.12** (p. 601) Let $X = (\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ be the canonical diffusion associated to the differential operator

$$L = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x)\frac{\partial}{\partial x_i} \tag{10.48}$$

that we assume to satisfy the same hypotheses as in Theorem 10.1. Let $D$ be an open set of $\mathbb{R}^m$, $x \in \partial D$; we shall say that $\partial D$ has a *local barrier* for $L$ at $x$ if there exists a function $w(y)$ defined and twice differentiable in a neighborhood $W$ of $x$ and such that $Lw \le -1$ on $W \cap D$, $w \ge 0$ on $W \cap \overline{D}$ and $w(x) = 0$.

Then

a) if $\partial D$ has a local barrier for $L$ at $x$, then $x$ is regular for the diffusion $X$.
b) (The sphere condition) Let $x \in \partial D$ and assume that there exists a ball $S \subset D^c$ such that $S \cap \overline{D} = \{x\}$. Then $x$ is regular for $X$.

a) Apply Ito's formula and compute the stochastic differential of $t \mapsto w(X_t)$. b) Construct a local barrier at $x$ of the form $w(y) = k[|x-z|^{-p} - |y-z|^{-p}]$, where $z$ is the center of $S$.

# Chapter 11
# *Simulation

Applications often require the computation of the expectation of a functional of a diffusion process. But for a few situations there is no closed formula in order to do this and one must recourse to approximations and numerical methods. We have seen in the previous chapter that sometimes such an expectation can be obtained by solving a PDE problem so that specific numerical methods for PDEs, such as finite elements, can be employed. Simulation of diffusion processes is another option which is explored in this chapter.

## 11.1 Numerical approximations of an SDE

Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a $d$-dimensional Brownian motion and $\xi$ the $m$-dimensional solution of the SDE

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\,dt + \sigma(\xi_t, t)\,dB_t \\ \xi_u &= \eta \end{aligned} \tag{11.1}$$

where

a) $b$ is an $m$-dimensional vector field and $\sigma$ a $m \times d$ matrix field satisfying assumptions to be made precise later;
b) $\eta$ is an $\mathscr{F}_u$-measurable square integrable r.v.

We have already discussed the question of the simulation of the paths of a process in the chapter about Brownian motion in Sect. 3.7. Of course, the method indicated there cannot be extended naturally to the case of a general diffusion unless the transition function $p$ of $\xi$ is known and easy to handle, as is the case, for instance, for the Ornstein–Uhlenbeck process, as explained in Example 9.3.

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_11

This is not, however, a common situation: the transition function is in most cases unknown explicitly or difficult to sample. Hence other methods are to be considered, possibly taking into account that we shall have only an approximate solution.

The simplest method in this direction is the so-called *Euler scheme*, which borrows the idea of the scheme that, with the same name, is used in order to solve numerically Ordinary Differential Equations. Sometimes it is called the Euler–Maruyama scheme, G. Maruyama (1955), being the first to apply it to SDEs.

The idea is to discretize the time interval $[u, T]$ into $n$ small intervals of length $h = \frac{1}{n}(T-u)$. Let $t_k = u+kh$, $k = 0, 1, \dots, n$. Then we consider the approximation

$$\begin{aligned}\xi_{t_k} &= \xi_{t_{k-1}} + \int_{t_{k-1}}^{t_k} b(\xi_s, s)\, ds + \int_{t_{k-1}}^{t_k} \sigma(\xi_s, s)\, dB_s \\ &\simeq \xi_{t_{k-1}} + b(\xi_{t_{k-1}}, t_{k-1})h + \sigma(\xi_{t_{k-1}}, t_{k-1})(B_{t_k} - B_{t_{k-1}})\ .\end{aligned} \tag{11.2}$$

Let $(Z_n)_n$ be a sequence of $d$-dimensional independent $N(0, I)$-distributed r.v.'s; then the r.v.'s $\sqrt{h}\, Z_k$ have the same joint distributions as the increments $B_{t_k} - B_{t_{k-1}}$ of the Brownian motion. We can construct the subsequent positions of an approximating process $\overline{\xi}^{(n)}$ by choosing the initial value $\overline{\xi}^{(n)}(u)$ by sampling with the same law as $\eta$ (and independently of the $Z_k$'s) and then through the iteration rule

$$\boxed{\overline{\xi}^{(n)}_{t_k} = \overline{\xi}^{(n)}_{t_{k-1}} + b(\overline{\xi}^{(n)}_{t_{k-1}}, t_{k-1})h + \sigma(\overline{\xi}^{(n)}_{t_{k-1}}, t_{k-1})\sqrt{h}\, Z_k\ . \qquad (11.3)}$$

Of course we need to prove that, as $h \to 0$, the process obtained in this way converges, in some appropriate way, to the solution of the SDE (11.1).

There are two points of view concerning the convergence. The first one is the so-called strong approximation. A result about strong approximation is a statement that gives an estimate of how close *the values* of the approximants $\overline{\xi}^{(n)}_{t_k}$ are to those of the solution $\xi_{t_k}$ when $Z_k = \frac{1}{\sqrt{h}}(B_{t_k} - B_{t_{k-1}})$ as in (11.2).

The second point of view, the weak approximation, concerns the estimation of how close *the law* of $\overline{\xi}^{(n)}$ is to the law of $\xi$. Typically it consists in estimates of how close the value of $\mathrm{E}[f(\overline{\xi}^{(n)}_T)]$ is to the "true" value $\mathrm{E}[f(\xi_T)]$ for a function $f$ that is sufficiently regular.

**Definition 11.1** An approximation scheme $\overline{\xi}^{(n)}$ is said to be *strongly convergent* of order $\beta$ if for every $k$, $1 \le k \le n$,

$$\mathrm{E}[|\overline{\xi}^{(n)}_{t_k} - \xi_{t_k}|^2]^{1/2} \le \text{const}\, h^\beta\ . \tag{11.4}$$

(continued)

**Definition 11.1** (continued)

Let $\mathscr{W}$ be a class of functions $\mathbb{R}^m \to \mathbb{R}$. An approximation scheme $\overline{\xi}^{(n)}$ is said to be *weakly convergent* of order $\beta$ for the class of functions $\mathscr{W}$ if for some $T > 0$

$$\left|\mathrm{E}[f(\xi_T^{(n)})] - \mathrm{E}[f(\xi_T)]\right| \le \text{const } h^\beta \tag{11.5}$$

for every function $f \in \mathscr{W}$.

Of course, for every Lipschitz continuous function $f$ with Lipschitz constant $L$ we have

$$\left|\mathrm{E}[f(\overline{\xi}_T^{(n)})] - \mathrm{E}[f(\xi_T)]\right| \le \mathrm{E}[|f(\overline{\xi}_T^{(n)}) - f(\xi_T)|] \le L\mathrm{E}[|\overline{\xi}_T^{(n)} - \xi_T|] \tag{11.6}$$

so that strong approximation gives information of weak type, at least for some class of functions $f$.

In the next sections we shall often drop the superscript $^{(n)}$. We shall use the notations

- $n$: number of subintervals into which the interval $[u, T]$ is partitioned;
- $h = \frac{1}{n}(T - u)$: length of each of these subintervals;
- $t_k = u + hk$, $k = 0, \dots, n$: endpoints of these subintervals.

## 11.2 Strong approximation

We shall make the following assumptions.

**Assumption (E)** We say that the coefficients $b$ and $\sigma$ satisfy Assumption (E) if there exist constants $L > 0, M > 0$ such that for every $x, y \in \mathbb{R}^m$, $t \in [0, T]$,

$$|b(x,t) - b(x,s)| \le L(1 + |x|)|t - s|^{1/2}, \tag{11.7}$$

$$|\sigma(x,t) - \sigma(x,s)| \le L(1 + |x|)|t - s|^{1/2}, \tag{11.8}$$

$$|b(x,t)| \le M(1 + |x|), \qquad |\sigma(x,t)| \le M(1 + |x|), \tag{11.9}$$

$$|b(x,t) - b(y,t)| \le L|x - y|, \qquad |\sigma(x,t) - \sigma(y,t)| \le L|x - y|\,. \tag{11.10}$$

In practice this is almost the same as Assumption (A) p. 260, but for the fact that we require more regularity in the time variable.

Let us consider an Euler scheme whose subsequent positions $\overline{\xi}_{t_0}, \dots, \overline{\xi}_{t_n}$ are defined in (11.3), with the choice $Z_k = \frac{1}{\sqrt{h}}\,(B_{t_k} - B_{t_{k-1}})$. It will be useful to define an approximating process $\widehat{\xi}$ interpolating the values of the Euler scheme between the times $t_k$. More precisely, for $t_{k-1} \le t \le t_k$, let

$$\widehat{\xi}_t = \overline{\xi}_{t_{k-1}} + b(\overline{\xi}_{t_{k-1}}, t_{k-1})(t - t_{k-1}) + \sigma(\overline{\xi}_{t_{k-1}}, t_{k-1})(B_t - B_{t_{k-1}})\ . \tag{11.11}$$

Let us define $j_n(s) = t_n$ for $t_n \le s < t_{n+1}$; then $\widehat{\xi}$ satisfies

$$\widehat{\xi}_t = \eta + \int_0^t b(\widehat{\xi}_{j_n(s)}, j_n(s))\, ds + \int_0^t \sigma(\widehat{\xi}_{j_n(s)}, j_n(s))\, dB_s\ . \tag{11.12}$$

In particular, $\widehat{\xi}$ is a Ito process.

Let us prove first that $\widehat{\xi}$ has finite moments of order $p$ for every $p \ge 1$.

**Lemma 11.1** If the coefficients $b, \sigma$ satisfy (11.9) (sublinear growth) and the initial value $\eta$ belongs to $L^p$, $p \ge 2$, then

$$\mathrm{E}\Big[\sup_{k=0,\dots,n} |\overline{\xi}_{t_k}|^p\Big] \le \mathrm{E}\Big[\sup_{u\le t\le T} |\widehat{\xi}_t|^p\Big] < +\infty\ .$$

*Proof* The proof follows almost exactly the steps of the proof of Theorem 9.1. For $R > 0$, let $\widehat{\xi}_R(t) = \widehat{\xi}(t \wedge \tau_R)$, where $\tau_R = \inf\{t; t \le T, |\widehat{\xi}_t| > R\}$ denotes the exit time of $\widehat{\xi}$ from the open ball of radius $R$. Then a repetition of the steps of the proof of Theorem 9.1 gives the inequality

$$\begin{aligned}\mathrm{E}\Big[\sup_{u\le s\le t} |\widehat{\xi}_R(s)|^p\Big] &\le c_1(p,T,M)\big(1 + \mathrm{E}|\eta|^p\big) + c_2(p,T,M) \int_u^t \mathrm{E}\big[|\widehat{\xi}_R(j_n(r))|^p\big]\, dr\\ &\le c_1(p,T,M)\big(1 + \mathrm{E}|\eta|^p\big) + c_2(p,T,M) \int_u^t \mathrm{E}\Big[\sup_{u\le s\le r} |\widehat{\xi}_R(s)|^p\Big]\, dr\ .\end{aligned}$$

Let now $v(t) = \mathrm{E}\big(\sup_{u\le s\le t} |\widehat{\xi}_R(s)|^p\big)$: from the previous inequality we have

$$v(t) \le c_1(p,T,M)(1 + \mathrm{E}|\eta|^p) + c_2(p,T,M) \int_u^s v(r)\, dr\ .$$

As $v(s) \le R$ we can apply Gronwall's inequality, which gives

$$v(T) \le c(p,T,M)\big(1 + \mathrm{E}[|\eta|^p]\big)\ ,$$

i.e.

$$\mathrm{E}\Big[\sup_{u\le s\le T\wedge\tau_R}|\widehat{\xi}_s|^p\Big]\le c(p,T,M)\big(1+\mathrm{E}[|\eta|^p]\big)\ .$$

We remark that the right-hand side does not depend on $R$ and we can conclude the proof by letting $R\to+\infty$ with the same argument as in the proof of Theorem 9.1.

□

The following theorem gives the main strong estimate.

**Theorem 11.1** Assume that $b$ and $\sigma$ satisfy Assumption (E). Let $B$ be a $d$-dimensional Brownian motion, $\eta\in L^p$, $p\ge 2$, an $m$-dimensional r.v. independent of $\mathscr{F}_u$ and $\xi$ the solution of

$$\begin{aligned}d\xi_t&=b(\xi_t,t)\,dt+\sigma(\xi_t,t)\,dB_t\\ \xi_u&=\eta\ .\end{aligned}$$

Let $\widehat{\xi}=\widehat{\xi}^{(n)}$ be the approximating process defined by the Euler scheme (11.11) and with initial condition $\widehat{\xi}_u=\eta$, $u\ge 0$. Then for every $p\ge 1$ and $T>u$, we have

$$\mathrm{E}\Big[\sup_{u\le t\le T}|\widehat{\xi}_t-\xi_t|^p\Big]\le\text{const}\ h^{p/2}\ , \tag{11.13}$$

where the constant $c$ depends only on $T$, $p$, $L$, $M$ and the $L^p$ norm of $\eta$.

*Proof* The idea of the proof is to find upper bounds allowing us to apply Gronwall's inequality. We shall assume for simplicity that $u=0$. Let $j_n(s)=t_n$ for $t_n\le s<t_{n+1}$, then, thanks to (11.12), we have

$$\begin{aligned}|\widehat{\xi}_t-\xi_t|^p\le 2^{p-1}&\Big|\int_0^t(b(\widehat{\xi}_{j_n(s)},j_n(s))-b(\xi_s,s))\,ds\Big|^p\\ &+2^{p-1}\Big|\int_0^t(\sigma(\widehat{\xi}_{j_n(s)},j_n(s))-\sigma(\xi_s,s))\,dB_s\Big|^p\end{aligned}$$

and, by Hölder's inequality and the $L^p$ bound of Proposition 8.9,

$$\begin{aligned}\mathrm{E}\Big[\sup_{0\le u\le t}|\widehat{\xi}_u-\xi_u|^p\Big]&\le 2^{p-1}T^p\mathrm{E}\Big[\int_0^t|b(\widehat{\xi}_{j_n(s)},j_n(s))-b(\xi_s,s)|^p\,ds\Big]\\ &+2^{p-1}c(p,m,d)T^{\frac{p-2}{2}}\mathrm{E}\Big[\int_0^t|\sigma(\widehat{\xi}_{j_n(s)},j_n(s))-\sigma(\xi_s,s)|^p\,ds\Big]\ .\end{aligned} \tag{11.14}$$

We have, for $t_n \le s \le t_{n+1}$,

$$
\begin{aligned}
&|b(\widehat{\xi}_{j_n(s)}, j_n(s)) - b(\xi_s, s)| = |b(\widehat{\xi}_{t_n}, t_n) - b(\xi_s, s)| \\
&\le |b(\xi_s, s) - b(\xi_{t_n}, s)| + |b(\xi_{t_n}, s) - b(\xi_{t_n}, t_n)| + |b(\xi_{t_n}, t_n) - b(\widehat{\xi}_{t_n}, t_n)|
\end{aligned}
$$

hence, for $t_n \le v < t_{n+1}$, under Assumption (E),

$$
\begin{aligned}
&\int_{t_n}^{v} |b(\widehat{\xi}_{j_n(s)}, j_n(s)) - b(\xi_s, s)|^p \, ds \\
&\le 3^{p-1} L^p \Big( \int_{t_n}^{v} |\xi_s - \xi_{t_n}|^p \, ds + (1 + |\xi_{t_n}|)^p \int_{t_n}^{v} |s - t_n|^{p/2} \, ds + \int_{t_n}^{v} |\xi_{t_n} - \widehat{\xi}_{t_n}|^p \, ds \Big) \, .
\end{aligned}
$$

The same arguments give

$$
\begin{aligned}
&\int_{t_n}^{v} |\sigma(\widehat{\xi}_{j_n(s)}, j_n(s)) - \sigma(\xi_s, s)|^p \, ds \\
&\le 3^{p-1} L^p \Big( \int_{t_n}^{v} |\xi_s - \xi_{t_n}|^p \, ds + (1 + |\xi_{t_n}|)^p \int_{t_n}^{v} |s - t_n|^{p/2} \, ds + \int_{t_n}^{v} |\xi_{t_n} - \widehat{\xi}_{t_n}|^p \, ds \Big) \, .
\end{aligned}
$$

Thanks to the $L^p$ estimates for the solution of an SDE with coefficients with a sublinear growth, (9.13) and (9.12), p. 261, we have for $t_n \le s < t_{n+1}$ and denoting by $c(p,T)$, $c(L,p,T)$ suitable constants,

$$
\mathrm{E}[|\xi_s - \xi_{t_n}|^p] \le c(p,T)(s - t_n)^{p/2}(1 + \mathrm{E}[|\xi_{t_n}|^p]) \le c(p,T)(s - t_n)^{p/2}(1 + \mathrm{E}[|\eta|^p]) \, .
$$

Substituting into (11.14) we find

$$
\begin{aligned}
\mathrm{E}\Big[\sup_{0 \le u \le t} |\widehat{\xi}_u - \xi_u|^p\Big] \le c(L,T,p) \Big\{ &(1 + \mathrm{E}[|\eta|^p]) \int_0^T |s - j_n(s)|^{p/2} \, ds \\
&+ \int_0^t \mathrm{E}[|\xi_{j_n(s)} - \widehat{\xi}_{j_n(s)}|^p] \, , ds \Big\} \, .
\end{aligned}
\tag{11.15}
$$

As

$$
\int_0^T |s - j_n(s)|^{p/2} \, ds = \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} (s - t_k)^{p/2} = \frac{n}{\frac{p}{2} + 1} \, h^{1+p/2} = \frac{T}{\frac{p}{2} + 1} \, h^{p/2}
$$

and

$$
\int_0^t \mathrm{E}[|\xi_{j_n(s)} - \widehat{\xi}_{j_n(s)}|^p] \, ds \le \int_0^t \mathrm{E}\Big[\sup_{0 \le u \le s} |\xi_u - \widehat{\xi}_u|^p\Big] \, ds \, ,
$$

for the function $v(t) = \mathrm{E}[\sup_{0\le u\le t} |\overline{\xi}_u - \xi_u|^p]$ we have the inequality

$$v(t) \le h^{p/2} c(L,T,p)(1+\mathrm{E}[|\eta|^p]) + c(L,T,p)\int_0^t v(s)\,ds\,.$$

As by Lemma 11.1 and Theorem 9.1 the function $v$ is bounded, we can apply Gronwall's inequality, which gives

$$\mathrm{E}\Big[\sup_{0\le u\le T} |\widehat{\xi}_u - \xi_u|^p\Big] \le c(L,T,p)\mathrm{e}^{Tc(L,T,p)}(1+\mathrm{E}[|\eta|^p])\,h^{p/2}$$

concluding the proof.

□

By Theorem 11.1 with $p = 2$ the Euler scheme is strongly convergent of order $\frac{1}{2}$.

Theorem 11.1 states that the $L^2$ difference between the r.v.'s $\overline{\xi}_{t_k}$ and $\xi_{t_k}$ tends to 0 as $h \to 0$. This is enough in many situations, typically if we need to estimate a functional of the type $\mathrm{E}[f(\xi_T)]$. Sometimes we might be concerned with more complicated functionals of the path of $\xi$, which leads to the following extension.

Let us define a process, which we shall denote again by $\overline{\xi}$, by setting, for $t_k \le t \le t_{k+1}$, $\overline{\xi}_t$ as the value that is obtained by interpolating linearly between the values $\overline{\xi}_{t_k}$ and $\overline{\xi}_{t_{k+1}}$. More precisely, if $t_k \le t \le t_{k+1}$, we have

$$t = \frac{1}{h}(t_{k+1}-t)t_k + \frac{1}{h}(t-t_k)t_{k+1}$$

and we define

$$\begin{aligned}\overline{\xi}_t &= \frac{1}{h}(t_{k+1}-t)\overline{\xi}_{t_k} + \frac{1}{h}(t-t_k)\overline{\xi}_{t_{k+1}} \\ &= \overline{\xi}_{t_k} + b(\overline{\xi}_{t_k},t_k)(t-t_k) + \frac{1}{h}\sigma(\overline{\xi}_{t_k},t_k)(t-t_k)(B_{t_{k+1}}-B_{t_k})\,.\end{aligned} \tag{11.16}$$

The processes $\overline{\xi}$ and $\widehat{\xi}$ coincide at the discretization times $t_k$ but differ between the times $t_k$ and $t_{k+1}$ because the stochastic components are different. $\overline{\xi}$ has the advantage that it can be numerically simulated.

**Corollary 11.1** Under the assumptions of Theorem 11.1, as $n \to \infty$ the Euler approximations $\overline{\xi}^{(n)}$ (11.16) converge in law to the law of $\xi$ on $\mathscr{C}([u,T],\mathbb{R}^m)$ for every $T > 0$ and in $\mathscr{C}(\mathbb{R}^+,\mathbb{R}^m)$.

*Proof* Again we shall assume $u = 0$. We shall prove that, for a fixed $T$, from every sequence $(n_j)_j$ converging to $+\infty$ we can extract a subsequence $(n'_j)_j$ such that $\overline{\xi}^{(n'_j)}$

converges to $\xi$ uniformly on $[0,T]$. As a.s. convergence implies convergence in law, this will conclude the proof. If we define $h_j = \frac{1}{n_j}$, then we have thanks to (11.13) and the Markov inequality, for every $\varepsilon > 0$,

$$P\Big(\sup_{k=0,\dots,n_j} |\overline{\xi}^{(n_j)}_{t_k} - \xi_{t_k}| \ge \varepsilon\Big) \le \frac{1}{\varepsilon^2} E\Big[\sup_{k=0,\dots,n_j} |\overline{\xi}^{(n_j)}_{t_k} - \xi_{t_k}|^2\Big] \le c\,\frac{h_j}{\varepsilon^2} = \frac{cT}{n_j\varepsilon^2}$$

hence certainly there exists a subsequence $(n'_j)_j$ such that

$$\sum_{j=1}^{\infty} P\Big(\sup_{k=0,\dots,n'_j} |\overline{\xi}^{(n'_j)}_{t_k} - \xi_{t_k}| \ge \varepsilon\Big) < +\infty \;.$$

By the Borel–Cantelli lemma, for every $\varepsilon > 0$ the event

$$\Big\{\sup_{k=0,\dots,n'_j} |\overline{\xi}^{(n'_j)}_{t_k} - \xi_{t_k}| \ge \varepsilon \text{ for infinitely many indices } j\Big\}$$

has probability 0. Let $\omega \in \Omega$. Then as the map $t \mapsto \xi_t(\omega)$ is continuous, for every $\varepsilon > 0$ there exists a $\delta_\varepsilon > 0$ such that if $|t-s| \le \delta_\varepsilon$ then $|\xi_t(\omega)-\xi_s(\omega)| \le \varepsilon$. Let us fix $\varepsilon > 0$ and let $j_0 > 0$ be such that $h'_j \le \varepsilon$, $h'_j \le \delta_\varepsilon$ and $\sup_{k=0,\dots,n'_j} |\overline{\xi}^{(n'_j)}_{t_k}(\omega) - \xi_{t_k}(\omega)| < \varepsilon$ for every $j \ge j_0$. Then we have, for $t_k \le t \le t_{k+1}$,

$$|\xi_t(\omega) - \xi_{t_{k+1}}(\omega)| \le \varepsilon \qquad \text{and} \qquad |\xi_t(\omega) - \xi_{t_k}(\omega)| \le \varepsilon \tag{11.17}$$

and also, for $j > j_0$,

$$|\xi_{t_k}(\omega) - \overline{\xi}^{(n'_j)}_{t_k}(\omega)| \le \varepsilon \qquad \text{and} \qquad |\xi_{t_{k+1}}(\omega) - \overline{\xi}^{(n'_j)}_{t_{k+1}}(\omega)| \le \varepsilon \;. \tag{11.18}$$

By (11.16) we have $\overline{\xi}^{(n'_j)}_t(\omega) = \alpha\overline{\xi}^{(n'_j)}_{t_k}(\omega) + (1-\alpha)\overline{\xi}^{(n'_j)}_{t_{k+1}}(\omega)$ with $\alpha = \frac{1}{h'_j}(t_{k+1} - t)$. Thanks to (11.17) and (11.18), we have by adding and subtracting the quantity $\alpha\xi_{t_k}(\omega) + (1-\alpha)\xi_{t_{k+1}}(\omega)$, if $t_k \le t \le t_{k+1}$,

$$\begin{aligned}
&|\xi_t(\omega) - \overline{\xi}^{(n'_j)}_t(\omega)| \\
&\le |\xi_t(\omega) - (\alpha\xi_{t_k}(\omega) + (1-\alpha)\xi_{t_{k+1}}(\omega))| \\
&\quad + |(\alpha\xi_{t_k}(\omega) + (1-\alpha)\xi_{t_{k+1}}(\omega)) - \underbrace{(\alpha\overline{\xi}^{(n'_j)}_{t_k}(\omega) + (1-\alpha)\overline{\xi}^{(n'_j)}_{t_{k+1}}(\omega)}_{=\overline{\xi}^{(n'_j)}_t(\omega)}| \\
&\le \alpha|\xi_t(\omega) - \xi_{t_k}(\omega)| + (1-\alpha)|\xi_t(\omega) - \xi_{t_{k+1}}(\omega)| \\
&\quad + \alpha|\xi_{t_k}(\omega) - \overline{\xi}^{(n'_j)}_{t_k}(\omega)| + (1-\alpha)|\xi_{t_{k+1}}(\omega) - \overline{\xi}^{(n'_j)}_{t_{k+1}}(\omega)| \\
&\le 2\varepsilon \;.
\end{aligned}$$

Hence we have proved that, for $j \geq j_0$, $\sup_{0 \leq t \leq T} |\xi_t - \overline{\xi}_t^{(n'_j)}| \leq 2\varepsilon$ hence that $(\overline{\xi}^{(n'_j)})_j$ converges a.s. to $\xi$ in $\mathscr{C}([0,T], \mathbb{R}^m)$.

The convergence in law of $\overline{\xi}^{(n)}$ to $\xi$ in $\mathscr{C}(\mathbb{R}^+, \mathbb{R}^m)$ is easily deduced recalling that convergence in $\mathscr{C}(\mathbb{R}^+, \mathbb{R}^m)$ means uniform convergence on $[0,T]$ for every $T$. The details are left to the reader.

□

We remarked above that Theorem11.1 and (11.6) guarantee that the Euler scheme is weakly convergent of order $\frac{1}{2}$ for the class of Lipschitz functions. In the next section we shall see that, for a class $\mathscr{W}$ of regular functions and under regularity assumptions on the coefficients $b$ and $\sigma$, the weak rate of convergence is of order 1.

## 11.3 Weak approximation

The following theorem gives an estimate concerning the weak convergence of the Euler scheme in the case of a time homogeneous diffusion. We shall skip the proof, putting the focus on the applications of the results.

**Theorem 11.2 (The Talay–Tubaro expansion)** Let us assume that

a) the coefficients $b$ and $\sigma$ are time homogeneous and differentiable infinitely many times with bounded derivatives of all orders ($b$ and $\sigma$ may be unbounded themselves);
b) the initial condition is $\eta \equiv x \in \mathbb{R}^m$;
c) $f : \mathbb{R}^m \to \mathbb{R}$ is differentiable infinitely many times and all its derivatives have polynomial growth, i.e. for every multi-index $\alpha = (\alpha_1, \dots, \alpha_m)$ there exist numbers $C_\alpha, p_\alpha > 0$ such that

$$\Big| \frac{\partial^{|\alpha|} f}{\partial x^\alpha} \Big| \leq C_\alpha (1 + |x|^{p_\alpha}) \, .$$

Then there exists a constant $c$ such that

$$\mathrm{E}[f(\xi_T)] - \mathrm{E}[f(\overline{\xi}_T)] = c\, h + O(h^2) \, . \tag{11.19}$$

The Talay–Tubaro theorem, Talay and Tubaro (1990) or Graham and Talay (2013, p. 180), actually gives more precision about the constant $c$. Note also that Theorem 11.2 does not just give a bound of the error as in Theorem 11.1, but gives an expansion of the error, which is a more precise statement. This fact will be of some importance in Sect. 11.6.

*Example 11.1* The usual weak convergence estimate, such as the one of Theorem 11.2, gives information concerning how much the expectation $\mathrm{E}[f(\overline{\xi}_T)]$ differs from the true value $\mathrm{E}[f(\xi_T)]$, but what about the discrepancy when considering the expectation of a more complicated function of the path of the diffusion $\xi$? For instance, if $\phi : \mathbb{R}^m \to \mathbb{R}$ is a regular function how do we estimate the difference between

$$\mathrm{E}\Big[\int_0^T \phi(\xi_s)\,ds\Big] \tag{11.20}$$

and, for instance, its Euler approximation

$$\mathrm{E}\Big[h\sum_{k=0}^{n-1}\phi(\overline{\xi}_{t_k})\Big]\ ? \tag{11.21}$$

A nice approach can be the following. Let us consider the $(m+1)$-dimensional diffusion that is the solution of the SDE

$$\begin{aligned} d\xi_t &= b(\xi_t)\,dt + \sigma(\xi_t)\,dB_t \\ d\eta_t &= \phi(\xi_t)\,dt\ . \end{aligned} \tag{11.22}$$

Hence $\eta_T = \int_0^T \phi(\xi_s)\,ds$ and we have written our functional of interest as a function of the terminal value of the diffusion $(\xi, \eta)$. The Euler scheme gives for the $\eta$ component the approximation

$$\overline{\eta}_T = h\sum_{k=0}^{n-1}\phi(\overline{\xi}_{t_k})\ . \tag{11.23}$$

If $\xi$ satisfies the assumptions of Theorem 11.2 and $\phi$ is differentiable infinitely many times with bounded derivatives, then (11.19) guarantees for the Euler approximation (11.23) a rate of convergence of order $h$.

## 11.4 Simulation and PDEs

The representation results of Chap. 10 allow us to compute numerically the solution of a PDE problem via the numerical computation of expectations of functionals of diffusion processes. Let us consider, as in Sect. 10.4, the differential operator

$$L_t = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x,t)\frac{\partial}{\partial x_i}$$

and the parabolic problem

$$\begin{cases} L_t u - cu + \dfrac{\partial u}{\partial t} = f & \text{on } \mathbb{R}^m \times [0, T[ \\ u(x, T) = \phi(x) & x \in \mathbb{R}^m . \end{cases}$$

We know that, under the hypotheses of Theorem 10.6, a solution is

$$u(x,t) = \mathrm{E}^{x,t}\Big[\phi(X_T)\,\mathrm{e}^{-\int_t^T c(X_s,s)\,ds}\Big] - \mathrm{E}^{x,t}\Big[\int_t^T f(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_v,v)\,dv}\,ds\Big], \qquad (11.24)$$

where $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^{x,s})_{x,s})$ is the canonical diffusion associated to $L_t$. Hence, thanks to Corollary 11.1, as the functionals inside the expectations in (11.24) are continuous functionals of the paths $(X_t)_t$, we have that, denoting by $\overline{\mathrm{P}}_n^{x,t}$ the law on $(\mathscr{C}, \mathscr{M})$ of the Euler approximation with time step $h = \frac{T}{n}$, if

$$\overline{u}_n(x,t) = \overline{\mathrm{E}}_n^{x,t}\Big[\phi(X_T)\,\mathrm{e}^{-\int_t^T c(X_s,s)\,ds}\Big] - \overline{\mathrm{E}}_n^{x,t}\Big[\int_t^T f(X_s,s)\,\mathrm{e}^{-\int_t^s c(X_v,v)\,dv}\,ds\Big]$$

then

$$\lim_{n\to\infty} \overline{u}_n(x,t) = u(x,t) .$$

In practice, an approximation of the value of the solution $u$ at $(x,t)$ can be obtained by simulating, with the Euler scheme, $N$ paths $\overline{\xi}_1, \dots, \overline{\xi}_N$ of the diffusion associated to the generator $L$ and with starting condition $x, t$ and considering the approximation

$$\overline{u}_{n,N}(x,t) = \frac{1}{N}\sum_{k=1}^N \overline{\Phi}_k ,$$

where

$$\overline{\Phi}_k = \phi(\overline{\xi}_k(T))\,\mathrm{e}^{-\int_t^T c(\overline{\xi}_k(s),s)\,ds} - \int_t^T f(\overline{\xi}_k(s),s)\,\mathrm{e}^{-\int_t^s c(\overline{\xi}_k(v),v)\,dv}\,ds .$$

In the simpler situation $f \equiv 0$, $c \equiv 0$ the approximation becomes

$$\overline{u}_{n,N}(x,t) = \frac{1}{N}\sum_{k=1}^N \phi(\overline{\xi}_k(T)) .$$

The solutions of the Dirichlet and Cauchy–Dirichlet problems of Sects. 10.2 and 10.3 can be numerically approximated in the same spirit. But then an additional problem appears. Let us consider the Dirichlet problem

$$\begin{cases} Lu - c = f & \text{on } D \\ u_{|\partial D} = \phi \, . \end{cases}$$

Theorem 10.2 states that, under suitable assumptions, the solution is given by

$$u(x) = \mathrm{E}^x[\phi(X_\tau)Z_\tau] - \mathrm{E}^x\Big[\int_0^\tau Z_s f(X_s)\,ds\Big]\,, \tag{11.25}$$

where $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ denotes the canonical realization of the diffusion associated to the generator $L$, $Z_t = \mathrm{e}^{-\int_0^t c(X_s)\,ds}$ and $\tau$ is the exit time from $D$.

We can think of simulating the diffusion with its Euler approximation in order to approximate $u(x)$ with

$$\overline{u}(x) = \overline{\mathrm{E}}^x[\phi(X_\tau)Z_\tau] - \overline{\mathrm{E}}^x\Big[\int_0^\tau Z_s f(X_s)\,ds\Big] \tag{11.26}$$

where by $\overline{\mathrm{E}}^x$ we denote the expectation with respect to the law of the Euler approximation with step $h$.

We know from Corollary 11.1 that under mild assumptions the laws of the Euler approximations $\overline{\mathrm{P}}^x$ converge weakly, as $h \to 0$, to $\mathrm{P}^x$. Unfortunately the map $\tau : \mathscr{C} \to \mathbb{R}^+$ is not continuous so we cannot immediately state that $\overline{u}(x)$ converges to $u(x)$ as $h \to 0$. In this section we prove that, in most cases, this is not a difficulty.

The idea is very simple. In the next statement we prove that the exit time from an open set is a lower semicontinuous functional $\mathscr{C} \to \mathbb{R}^+$, whereas the exit time from a closed set is upper semicontinuous. In Theorem 11.3 it will be proved that, under suitable assumptions, the exit time $\tau$ from an open set $D$ is $\mathrm{P}^x$-a.s. equal to the exit time from its closure for every $x$. Hence $\tau$ will turn out to be $\mathrm{P}^x$-a.s. continuous, which is sufficient in order to guarantee the convergence of $\overline{u}(x)$ to $u(x)$.

**Proposition 11.1** Let $D$ be an open (resp. closed) set. Then the exit time $\tau$ from $D$ as a functional $\mathscr{C} \to \mathbb{R}^+$ is lower (resp. upper) semicontinuous.

*Proof* Assume that $D$ is an open set and let $\gamma \in \mathscr{C}$. For every $\varepsilon > 0$ the set $\Gamma = \{x \in D; x = \gamma_t \text{ for some } t \le \tau(\gamma) - \varepsilon\}$ is a compact set contained in $D$. Hence $\delta := d(\Gamma, \partial D) > 0$ (the distance between a compact set and a disjoint closed set is strictly positive). Now $U = \{w \in \mathscr{C}; \sup_{0 \le t \le \tau(\gamma)-\varepsilon} |w_t - \gamma_t| \le \frac{\delta}{2}\}$ is a neighborhood of $\gamma$ such that for every path $w \in U$ $\tau(w) \ge \tau(\gamma) - \varepsilon$. Hence, by the arbitrariness of $\varepsilon$,

$$\varliminf_{w \to \gamma} \tau(w) \ge \tau(\gamma) .$$

Assume now that $D$ is closed and let $\gamma$ be such that $\tau(\gamma) < \infty$. Then there exist arbitrarily small values of $\varepsilon > 0$ such that $\gamma_{\tau(\gamma)+\varepsilon} \in D^c$. Let $\delta := d(\gamma_{\tau(\gamma)+\varepsilon}, D) > 0$. Again if $U = \{w \in \mathscr{C}; \sup_{0 \le t \le \tau(\gamma)+\varepsilon} |w_t - \gamma_t| \le \frac{\delta}{2}\}$, for every $w \in U$ we have $\tau(w) \le \tau(\gamma) + \varepsilon$. □

We must now find conditions ensuring that, for a given open set $D$, the exit time from $D$ and from its closure coincide $\mathrm{P}^x$-a.s.

The intuitive explanation of this fact is that the paths of a diffusion are subject to intense oscillations. Therefore as soon as a path has reached $\partial D$, hence has gone out of $D$, it immediately also exits from $\overline{D}$ a.s. This is an argument similar to the one developed in Example 6.4 (and in particular in Fig. 6.1). It will be necessary to assume the boundary to have enough regularity and the generator to be elliptic, which will ensure that oscillations take place in all directions.

The formal proof that we now develop, however, shall take a completely different approach.

Let $D \subset \mathbb{R}^m$ be a regular open set and let $D_n$ be a larger regular open set with $\overline{D} \subset D_n$ and such that $\operatorname{dist}(\partial D, \partial D_n) \to 0$ as $n \to \infty$. Let us consider the functions $u$, $u_n$ that are the solutions respectively of the problems

$$\begin{cases} Lu = -1 & \text{on } D \\ u_{|\partial D} = 0 \end{cases} \tag{11.27}$$

and

$$\begin{cases} Lu_n = -1 & \text{on } D_n \\ u_{n|\partial D_n} = 0 . \end{cases} \tag{11.28}$$

As a consequence of Theorem 10.2 (see also Exercise 10.1) we have, denoting by $\tau$, $\overline{\tau}$ and $\tau_n$ the exit times from $D$, $\overline{D}$, $D_n$ respectively,

$$u(x) = \mathrm{E}^x[\tau], \qquad u_n(x) = \mathrm{E}^x[\tau_n]$$

hence, for every $x \in D$, as $\tau \le \overline{\tau} \le \tau_n$

$$u(x) \le \mathrm{E}^x[\overline{\tau}] \le u_n(x) . \tag{11.29}$$

But, of course, $u_n$ is also the solution of

$$\begin{cases} Lu_n = -1 & \text{on } D \\ u_n(x) = u_n(x) & x \in \partial D \end{cases} \tag{11.30}$$

and thanks to the representation formula (10.6), for $x \in D$,

$$\mathrm{E}^x[\tau_n] = u_n(x) = \mathrm{E}^x[u_n(X_\tau)] + \mathrm{E}^x[\tau]\ . \tag{11.31}$$

Later on in this section we prove that, for every $\varepsilon > 0$, there exists an $n_0$ such that $u_n(x) \le \varepsilon$ on $\partial D$ for $n \ge n_0$. This is rather intuitive, as $u_n = 0$ on $\partial D_n$ and the boundaries of $D$ and $D_n$ are close. Hence (11.31) will give

$$u_n(x) \le \mathrm{E}^x[\tau] + \varepsilon\ ,$$

which together with (11.29) gives

$$\mathrm{E}^x[\tau] \le \mathrm{E}^x[\overline{\tau}] \le \mathrm{E}^x[\tau] + \varepsilon\ .$$

Hence, by the arbitrariness of $\varepsilon$, $\mathrm{E}^x[\tau] = \mathrm{E}^x[\overline{\tau}]$ and as $\tau \le \overline{\tau}$ this entails $\tau = \overline{\tau}$ a.s.

The result is the following. It requires the boundary of $D$ to be $C^{2,\alpha}$, i.e. locally the graph of a function that is twice differentiable with second derivatives that are Hölder continuous of order $\alpha$. It also requires that $D$ is convex, an assumption that can certainly be weakened.

**Theorem 11.3** Let $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ be the canonical diffusion associated to the differential operator

$$L = \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m b_i(x)\frac{\partial}{\partial x_i} \tag{11.32}$$

and let $D$ be a bounded open set with a $C^{2,\alpha}$ boundary with $0 < \alpha < 1$. We assume that there exists an open set $\widetilde{D}$ such that $\overline{D} \subset \widetilde{D}$ and

- the coefficients $a$ and $b$ are locally Lipschitz continuous on $\widetilde{D}$;
- $L$ is uniformly elliptic on $\widetilde{D}$, i.e. $\langle a(x)\, z, z\rangle \ge \lambda |z|^2$ for some $\lambda > 0$ and for every $x \in \widetilde{D}$, $z \in \mathbb{R}^m$.

Then the exit time $\tau$ from $D$ is a.s. continuous with respect to $\mathrm{P}^x$ for every $x$.

In order to complete the proof we must prove that for every $\varepsilon > 0$ there exists an $n_0$ such that $u_n(x) \le \varepsilon$ on $\partial D$ for $n \ge n_0$. This will follow as soon as we prove that the gradient of $u_n$ is bounded uniformly in $n$, and this will be a consequence of classical estimates for elliptic PDEs.

Let us introduce some notation. For a function $f$ on a domain $D$ and $0 < \alpha \le 1$, let us introduce the Hölder's norms

$$|f|_\alpha = \|f\|_{\infty,D} + \sup_{x,y\in D, x\neq y} \frac{|f(x)-f(y)|}{|x-y|^\alpha}$$

and, for a positive integer $k$ and denoting by $\beta = (\beta_1, \dots, \beta_m)$ a multi-index,

$$|f|_{k,\alpha} = \sum_{h=0}^{k} \sum_{|\beta|=h} \Big| \frac{\partial^{|\beta|} f}{\partial x^\beta} \Big|_\alpha \cdot$$

The following result (see Gilbarg and Trudinger 2001, Theorem 6.6, p. 98 for example) is a particular case of the classical Schauder estimates.

**Theorem 11.4** Let $D$ be an open set with a $C^{2,\alpha}$ boundary and let $u$ be a solution of (11.27), where $L$ is the differential operator (11.32). Assume that the coefficients satisfy the conditions

$$\langle a(x)\xi, \xi\rangle \ge \lambda|\xi|^2 \tag{11.33}$$

for every $x \in D$, $\xi \in \mathbb{R}^m$ and, for some $\alpha$, $0 < \alpha < 1$,

$$|a_{ij}|_\alpha, |b_i|_\alpha, \le \Lambda \tag{11.34}$$

for every $i, j = 1, \dots, m$, for some constants $\lambda, \Lambda > 0$. Then for the solution $u$ of (11.27) we have the bound

$$|u|_{2,\alpha} \le C(\|u\|_\infty + 1) \tag{11.35}$$

for some constant $C = C(m, \alpha, \lambda, \Lambda, D)$.

*End of the proof of Theorem 11.3* We must prove that, for a suitable family $(D_n)_n$ as above, the gradients of the solutions $u_n$ of (11.30) are bounded uniformly in $n$. This will be a consequence of the Schauder inequalities of Theorem 11.4, as the Hölder norm $|\ |_{2,\alpha}$ majorizes the supremum norm of the gradient.

The first task is to prove that the constant $C$ appearing in (11.35) can be chosen to hold for $D_n$ for every $n$.

Let us assume that the convex set $D$ contains the origin so that we can choose $D_n$ as the homothetic domain $(1+\frac{1}{n})D$. Note that, as $D$ is assumed to be convex, $D \subset D_n$. Let us define $r_n = 1 + \frac{1}{n}$ and let $\overline{n}$ be such that $D_{\overline{n}} \subset \widetilde{D}$. Let

$$\begin{aligned} \widetilde{u}_n(x) &= u_n(r_n x) \\ \widetilde{L}_n &= \frac{1}{2}\sum_{i,j=1}^m \widetilde{a}_{ij}^{(n)}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m \widetilde{b}_i^{(n)}(x)\frac{\partial}{\partial x_i} \end{aligned} \tag{11.36}$$

with $\widetilde{a}_{ij}^{(n)}(x) = \frac{1}{r_n^2}\, a_{ij}(r_n x)$, $\widetilde{b}^{(n)}(x) = \frac{1}{r_n}\, b_i(r_n x)$. Then $\widetilde{L}_n \widetilde{u}_n(x) = (L_n u_n)(r_n x)$, hence $\widetilde{u}_n$ is the solution of

$$\begin{cases} \widetilde{L}_n \widetilde{u}_n = -1 & \text{on } D \\ \widetilde{u}_{n|_{\partial D}} = 0\,. \end{cases} \tag{11.37}$$

Let us consider on $D$ the differential operators $\widetilde{L}_n$, $L$ and let us verify that there exist constants $\lambda$, $\Lambda$ such that (11.33) and (11.34) are satisfied for all of them. Indeed if $\langle a(x)\xi, \xi\rangle \ge \lambda|\xi|^2$ for every $x \in D_{\overline{n}}$, then, as $r_n \le 2$,

$$\langle a^{(n)}(x)\xi, \xi\rangle \ge \frac{\lambda}{r_n^2}\,|\xi|^2 \ge \frac{\lambda}{4}\,|\xi|^2$$

and if $\Lambda$ is such that $|a_{ij}|_\alpha \le \Lambda$ (the Hölder norm being taken on $D_{\overline{n}}$), then

$$|a_{ij}^{(n)}|_\alpha \le \frac{r_n^\alpha}{r_n^2}\,\Lambda \le \Lambda$$

and a similar majorization holds for the coefficients $\widetilde{b}^{(n)}$. Hence by Theorem 11.3 there exists a constant $C$ such that, for every $n$,

$$r_n\|u_n{}'\|_\infty = \|\widetilde{u}_n{}'\|_\infty \le |\widetilde{u}|_{2,\alpha} \le C(\|\widetilde{u}_n\|_\infty + 1) = C(\|u_n\|_\infty + 1)\,. \tag{11.38}$$

Finally, simply observe that, as $D_n \subset D_{\overline{n}}$ and the generator $L$ is elliptic on $D_{\overline{n}}$, the exit time from $D_{\overline{n}}$ is integrable thanks to Proposition 10.1 and we have the bound

$$\|u_n\|_\infty = \sup_{x\in D_n} \mathrm{E}^x[\tau_n] \le \sup_{x\in D_{\overline{n}}} \mathrm{E}^x[\tau_{\overline{n}}] < +\infty \tag{11.39}$$

and, putting together (11.38) and (11.39), we obtain that the supremum norm of the gradient of $u_n$ is bounded uniformly in $n$, thus concluding the proof of Theorem 11.3. □

Note that the assumption of convexity for the domain $D$ is required only in order to have $D \subset (1 + \frac{1}{n})D$ and can be weakened ($D$ starshaped is also a sufficient assumption, for example).

*Example 11.2* Consider the Poisson problem in dimension 2

$$\begin{cases} \frac{1}{2}\triangle u = 0 & \text{on } D \\ u_{|\partial D}(x_1, x_2) = x_2 \vee 0\ , \end{cases} \tag{11.40}$$

where $D$ is the ball centered at 0 of radius 1. The representation formula (10.6) gives

$$u(x) = \mathrm{E}^x[X_2(\tau) \vee 0]\ ,$$

where $(\mathscr{C}, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, (\mathrm{P}^x)_x)$ denotes the realization of the diffusion process associated to $L = \frac{1}{2}\triangle$ so that, under $\mathrm{P}^x$, $X = (X_1, X_2)$ is a two-dimensional Brownian motion starting at $x$ and $\tau$ is the exit time from $D$. The value $u(x)$ of the solution at $x \in D$ can therefore be computed by simulating many paths of the Brownian motion starting at $x$ and taking the mean of the values of $X_2(\tau)\vee 0$, i.e. the positive part of the ordinate of the exit point. Doing this for a grid of starting points in $D$, the result would appear as in Fig. 11.1.

This can be compared with the solution obtained with the aid of some specific numerical method for PDEs (finite elements for instance), shown in Fig. 11.2.

It is apparent that finite elements are much more accurate, as a comparison with the exact solution given by the formulas of Example 10.1 would show. Simulation therefore performs poorly in comparison with these numerical methods and requires, for an accurate result, a very small value of the discretization step $h$ and a large number of simulated paths. However it has two great advantages.

a) It is very simple to implement.
b) Very often it is the only method available. This is the case, for instance, if the dimension is larger than, say, 4. In these situations other numerical methods are unavailable or become really complicated to implement.

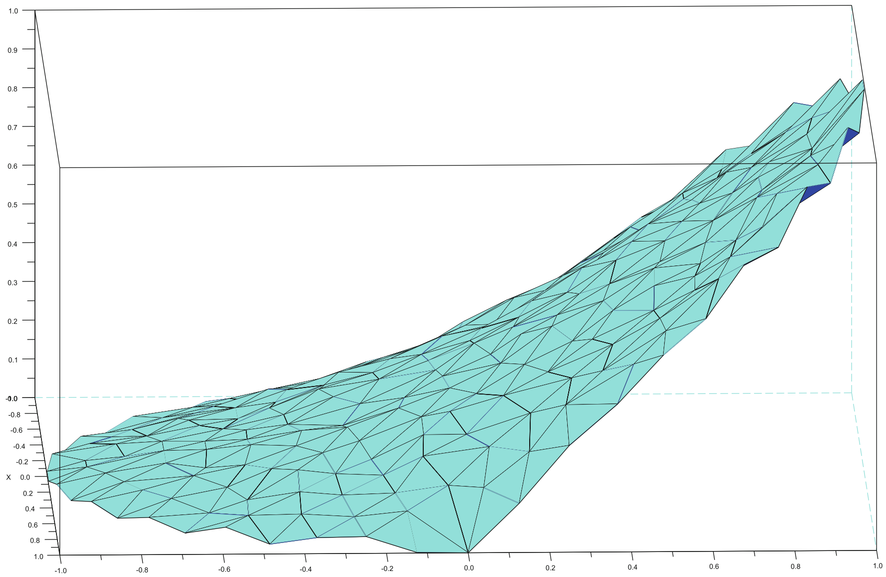


**Fig. 11.1** The solution of (11.40) computed numerically by simulation of a two-dimensional Brownian motion. At each point the value was determined as the average on $n = 1600$ paths, each obtained by a time discretization $h = \frac{1}{50}$

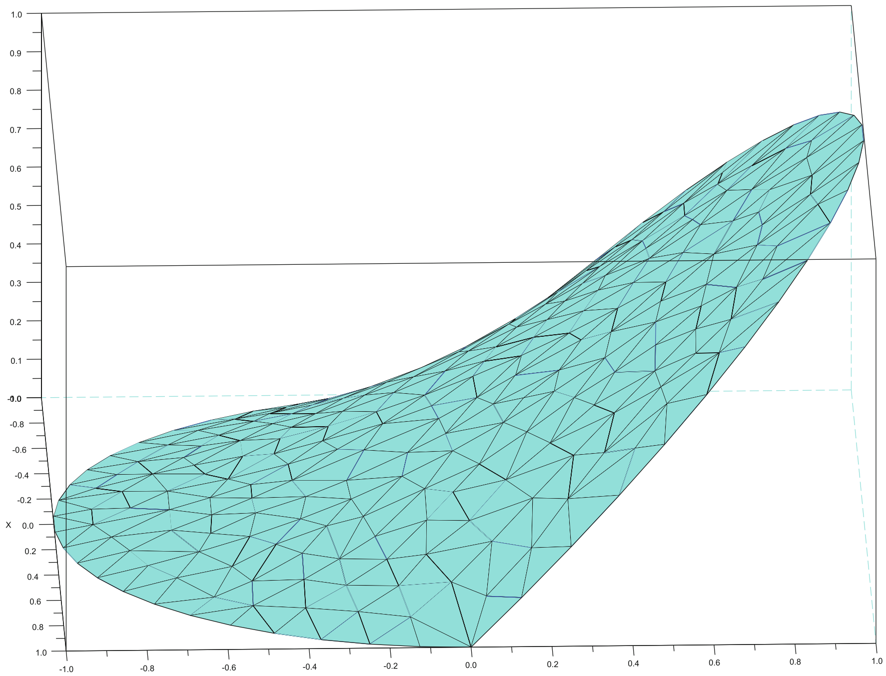


**Fig. 11.2** The solution of (11.40) computed numerically with the finite elements method

## 11.5 Other schemes

The Euler scheme is possibly the most natural approach to the approximation of the solution of an SDE but not the only one, by far. In this section we provide another approximation scheme. Many more have been developed so far, the interested reader can refer to the suggested literature.

Going back to (11.2) i.e.

$$\xi_{t_k} = \xi_{t_{k-1}} + \int_{t_{k-1}}^{t_k} b(\xi_s, s)\, ds + \int_{t_{k-1}}^{t_k} \sigma(\xi_s, s)\, dB_s \,, \tag{11.41}$$

a natural idea is to find a better approximation of the two integrals. The Euler scheme can be thought of as a zero-th order development: what if we introduce higher order developments of these integrals?

In this direction we have the Milstein scheme. Let us apply Ito's formula to the process $s \mapsto \sigma(\xi_s, s)$: assuming that $\sigma$ is twice differentiable, we have for $t_{k-1} \le s < t_k$,

$$\begin{aligned} \sigma_{ij}(\xi_s, s) =& \sigma_{ij}(\xi_{t_{k-1}}, t_{k-1}) \\ &+ \int_{t_{k-1}}^{s} \Big( \frac{\partial \sigma_{ij}}{\partial u} + \sum_{l=1}^{m} \frac{\partial \sigma_{ij}}{\partial x_l} b_l + \frac{1}{2} \sum_{r,l=1}^{m} \frac{\partial^2 \sigma_{ij}}{\partial x_r \partial x_l} a_{rl} \Big)(\xi_u, u)\, du \\ &+ \int_{t_{k-1}}^{s} \sum_{l=1}^{d} \Big( \sum_{r=1}^{m} \frac{\partial \sigma_{ij}}{\partial x_r} \sigma_{rl} \Big)(\xi_u, u)\, dB_l(u) \,. \end{aligned}$$

The integral in $du$ is of order $h$ and, after integration from $t_{k-1}$ to $t_k$ in (11.41), will give a contribution of order $o(h)$, which is negligible with respect to the other terms. Also we can approximate

$$\begin{aligned} &\int_{t_{k-1}}^{s} \sum_{l=1}^{d} \Big( \sum_{r=1}^{m} \frac{\partial \sigma_{ij}}{\partial x_r} \sigma_{rl} \Big)(\xi_u, u)\, dB_l(u) \\ \simeq& \sum_{r=1}^{m} \frac{\partial \sigma_{ij}}{\partial x_r}(\xi_{t_{k-1}}, t_{k-1}) \sum_{l=1}^{d} \sigma_{rl}(\xi_{t_{k-1}}, t_{k-1})(B_l(s) - B_l(t_{k-1})) \,, \end{aligned}$$

which gives finally the Milstein scheme, which is obtained by adding to the Euler iteration rule (11.2) the term whose $i$-th component is

$$\sum_{j=1}^{d} \sum_{r=1}^{m} \frac{\partial \sigma_{ij}}{\partial x_r}(\xi_{t_{k-1}}, t_{k-1}) \sum_{l=1}^{d} \sigma_{rl}(\xi_{t_{k-1}}, t_{k-1}) \int_{t_{k-1}}^{t_k} (B_l(s) - B_l(t_{k-1}))\, dB_j(s) \,.$$

If $d = 1$, i.e. the Brownian motion is one-dimensional and $\sigma$ is an $m \times 1$ dimensional matrix, by Ito's formula we have

$$\int_{t_{k-1}}^{t_k} (B_s - B_{t_{k-1}})\, dB_s = \frac{1}{2}\Big((B_{t_k} - B_{t_{k-1}})^2 - h\Big)$$

and the iteration rule becomes, for $i = 1, \dots, m$,

$$\begin{aligned}\xi_i(t_k) =& \xi_i(t_{k-1}) + b_i(\xi_{t_{k-1}}, t_{k-1})h + \sigma_i(\xi_{t_{k-1}}, t_{k-1})(B_{t_k} - B_{t_{k-1}}) \\ &+ \sum_{l=1}^m \frac{\partial \sigma_i}{\partial x_l}(\xi_{t_{k-1}}, t_{k-1})\sigma_l(\xi_{t_{k-1}}, t_{k-1}) \int_{t_{k-1}}^{t_k} (B_s - B_{t_{k-1}})\, dB_s \\ =& \xi_i(t_{k-1}) + b_i(\xi_{t_{k-1}}, t_{k-1})h + \sigma_i(\xi_{t_{k-1}}, t_{k-1})(B_{t_k} - B_{t_{k-1}}) \\ &+ \frac{1}{2}\sum_{l=1}^m \frac{\partial \sigma_i}{\partial x_l}(\xi_{t_{k-1}}, t_{k-1})\sigma_l(\xi_{t_{k-1}}, t_{k-1})\Big((B_{t_k} - B_{t_{k-1}})^2 - h\Big)\,.\end{aligned}$$

In practice, the subsequent positions of the approximating process will be obtained by choosing $\overline{\xi}_u \sim \eta$ and iterating

$$\begin{aligned}\overline{\xi}_i(t_k) =& \overline{\xi}_i(t_{k-1}) + b_i(\overline{\xi}_{t_{k-1}}, t_{k-1})h + \sqrt{h}\,\sigma_i(\overline{\xi}_{t_{k-1}}, t_{k-1})Z_k \\ &+ \frac{h}{2}\sum_{l=1}^m \frac{\partial \sigma_i}{\partial x_l}(\overline{\xi}_{t_{k-1}}, t_{k-1})\sigma_l(\overline{\xi}_{t_{k-1}}, t_{k-1})(Z_k^2 - 1)\end{aligned}$$

for a sequence $(Z_k)_k$ of independent $N(0, 1)$-distributed r.v.'s.

In dimension larger than 1 this scheme requires us to be able to simulate the joint distribution of the r.v.'s

$$B_i(t_j),\quad \int_{t_{k-1}}^{t_k} (B_l(t) - B_l(t_{k-1}))\, dB_j(t), \quad i, j, l = 1, \dots, d,\ \ k = 1, \dots, n\,,$$

which is not easy. In addition this scheme requires the computation of the derivatives of $\sigma$. The Milstein scheme in practice is confined to the simulation of diffusions in dimension 1 or, at least, with a one-dimensional Brownian motion.

Also for the Milstein scheme there are results concerning strong and weak convergence. Without entering into details, the Milstein scheme is of strong order of convergence equal to 1 (i.e. better that the Euler scheme) and of weak order of convergence also 1 (i.e. the same as the Euler scheme).

Let us mention, among others, the existence of higher-order schemes, usually involving a higher order development of the coefficients.

## 11.6 Practical remarks

In this section we discuss some issues a researcher may be confronted with when setting up a simulation procedure.

In practice, in order to compute the expectation of a continuous functional of the diffusion process $\xi$ it will be necessary to simulate many paths $\overline{\xi}_1, \dots, \overline{\xi}_N$ using the Euler or some other scheme with discretization step $h$ and then to take their average.

If the functional is of the form $\Phi(\xi)$ for a continuous map $\Phi : \mathscr{C}([0,T], \mathbb{R}^m) \to \mathbb{R}$, the expectation $\mathrm{E}[\Phi(\xi)]$ will be approximated by the average

$$\overline{\Phi}_N := \frac{1}{N}\sum_{k=1}^{N} \Phi(\overline{\xi}_k) .$$

In order to evaluate the error of this approximation it is natural to consider the mean square error

$$\mathrm{E}\big[(\overline{\Phi}_N - \mathrm{E}[\Phi(\xi)])^2\big] .$$

We have, adding and subtracting $\mathrm{E}[\Phi(\overline{\xi})]$,

$$\begin{aligned}
&\mathrm{E}\big[(\overline{\Phi}_N - \mathrm{E}[\Phi(\xi)])^2\big]\\
&= \mathrm{E}\big[(\overline{\Phi}_N - \mathrm{E}[\Phi(\overline{\xi})])^2\big] + 2\mathrm{E}\big(\overline{\Phi}_N - \mathrm{E}[\Phi(\overline{\xi})]\big)\big(\mathrm{E}[\Phi(\overline{\xi})] - \mathrm{E}[\Phi(\xi)]\big)\\
&\qquad + (\mathrm{E}[\Phi(\overline{\xi})] - \mathrm{E}[\Phi(\xi)])^2 .
\end{aligned}$$

The double product above vanishes as $\mathrm{E}[\overline{\Phi}_N] = \mathrm{E}[\Phi(\overline{\xi})]$ so that the mean square error is the sum of the 2 terms

$$I_1 := \mathrm{E}\big[(\overline{\Phi}_N - \mathrm{E}[\Phi(\overline{\xi})])^2\big], \quad I_2 := (\mathrm{E}[\Phi(\overline{\xi})] - \mathrm{E}[\Phi(\xi)])^2 .$$

The first term, $I_1$, is equal to $\frac{1}{N}\mathrm{Var}(\Phi(\overline{\xi}_1))$, i.e. is the Monte Carlo error. The second one is the square of the difference between the expectation of $\Phi(\xi)$ and the expectation of the approximation $\Phi(\overline{\xi})$. If we consider the Euler scheme and the functional $\Phi$ is of the form $\Phi(\xi) = f(\xi_T)$ and the assumptions of Theorem 11.2 are satisfied, we then have $I_2 \sim c\,h^2$.

These remarks are useful when designing a simulation program, in order to decide the best values for $N$ (number of simulated paths) and $h$ (amplitude of the discretization step). Large numbers of $N$ make the Monte Carlo error $I_1$ smaller, whereas small values of $h$ make the bias error $I_2$ smaller.

In the numerical approximation problems of this chapter it is possible to take advantage of the classical Romberg artifice, originally introduced in the context of the numerical solution of Ordinary Differential Equations. Assume that we know that an approximation scheme has an error of order $h^\alpha$, i.e. that

$$a_h := \mathrm{E}[f(\overline{\xi}_T^{(n)})] = \mathrm{E}[f(\xi_T)] + ch^\alpha + o(h^\alpha)\ . \tag{11.42}$$

Then if we have an estimate $a_h$ for the value $h$ and an estimate $a_{h/2}$ for $h/2$, we have

$$2^\alpha a_{h/2} - a_h = (2^\alpha - 1)\mathrm{E}[f(\xi_T)] + o(h^\alpha)$$

i.e., the linear combination

$$\frac{2^\alpha a_{h/2} - a_h}{2^\alpha - 1} = \mathrm{E}[f(\xi_T)] + o(h^\alpha)$$

of the two approximations $a_h$ and $a_{h/2}$ gives an approximation of higher order.

Note that the knowledge of the constant $c$ in (11.42) is not needed. In order to apply this artifice the value of $\alpha$ must be known, otherwise there is no improvement (but the estimate should not become worse by much).

In the case of an Euler approximation, in the assumptions of Theorem 11.2 we have $\alpha = 1$ and the Romberg estimator will give an approximation of order $h^2$.

*Example 11.3* Imagine that you want to obtain an approximation of the expectation at time $T$ of a geometric Brownian motion using the Euler scheme. Of course, the true value is known and for geometric Brownian motion there are more reasonable ways to perform a simulation, but this example allows us to make some explicit estimates.

By Theorem 11.2 the error is of order $h$. The Romberg artifice says that if, instead of a simulation with a very small value of $h$, you make two simulations with moderately small values $h$ and $h/2$ you may obtain a better estimate at a lower price. Table 11.1 shows some results. Here the drift and diffusion coefficient are $b = 1$ and $\sigma = 0.5$ and we set $T = 1$. Hence the true value is $\mathrm{e} = 2.7182$.

**Table 11.1** The quantity of interest appears to be better estimated making two sets of approximations with the values $h = 0.02$ and $h = 0.01$ than with a single simulation with $h = 0.001$, which is probably more costly in time. Here $10^6$ paths were simulated for every run

| | Value | Error |
|---|---|---|
| $h = 0.02$ | 2.6920 | 0.0262 |
| $h = 0.01$ | 2.7058 | 0.0124 |
| $2a_{0.01} - a_{0.02}$ | 2.7196 | 0.0014 |
| $h = 0.001$ | 2.7166 | 0.0016 |

*Example 11.4* Let us go back to Example 3.3 where we were dealing with the estimation by simulation of the probability

$$P\Big(\sup_{0\le t\le 1} B_t > 1\Big)\, .$$

We remarked there that the discrepancy between the estimate provided by the Euler scheme (which in this case coincides with a discretization of the Brownian motion) and the true value was decreasing very slowly. Actually it is believed that it is of the order $h^{1/2}$. Granting this fact, the Romberg artifice gives, combining the estimates for $h = \frac{1}{200}$ and $h = \frac{1}{400}$, the value

$$\frac{\sqrt{2} \times 0.3031 - 0.2975}{\sqrt{2} - 1} = 0.3166$$

with a 0.15% of relative error with respect to the true value 0.3173. Note that this value is much better than the result obtained by simulation with $h = \frac{1}{1600}$.

## Exercises

**11.1** (p. 602) Let us consider the geometric Brownian motion that is the solution of the SDE

$$\begin{aligned} d\xi_t &= b\xi_t\, dt + \sigma\xi_t\, dB_t \\ \xi_0 &= x\, . \end{aligned} \tag{11.43}$$

In this exercise we compute explicitly some weak type estimates for the Euler scheme for this process. Let us denote by $\overline{\xi}$ the Euler scheme with the time interval $[0, T]$ divided into $n$ subintervals. Of course the discretization step is $h = \frac{T}{n}$. Let $(Z_k)_k$ be a sequence of independent $N(0, 1)$-distributed r.v.'s.

a1) Prove that

$$\overline{\xi}_T \sim x \prod_{i=1}^{n} (1 + bh + \sigma\sqrt{h}\, Z_k) \ . \tag{11.44}$$

a2) Compute the mean and variance of $\overline{\xi}_T$ and verify that they converge to the mean and variance of $\xi_T$, respectively.

a3) Prove that

$$|\mathrm{E}[\xi_T] - \mathrm{E}[\overline{\xi}_T]| = c_1\, h + o(h) \tag{11.45}$$

$$|\mathrm{E}[\xi_T^2] - \mathrm{E}[\overline{\xi}_T^2]| = c_2\, h + o(h) \ . \tag{11.46}$$

b1) Find a formula similar to (11.44) for the Milstein approximation of $\xi_T$.

b2) Let us denote this approximation by $\widetilde{\xi}_T$. Prove that

$$|\mathrm{E}[\xi_T] - \mathrm{E}[\widetilde{\xi}_T]| = \widetilde{c}_1\, h + o(h) \tag{11.47}$$

and compare the values $\widetilde{c}_1$ and $c_1$ of (11.45).

# Chapter 12
# Back to Stochastic Calculus

## 12.1 Girsanov's theorem

*Example 12.1* Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion. We know that

$$M_t = \mathrm{e}^{\lambda B_t - \frac{1}{2}\lambda^2 t}$$

is a positive martingale with expectation equal to 1. We can therefore define on $(\Omega, \mathscr{F})$ a new probability Q letting, for every $A \in \mathscr{F}$,

$$\mathrm{Q}(A) = \mathrm{E}[M_T 1_A] = \int_A M_T \, d\mathrm{P} \, .$$

What has become of $(B_t)_t$ with respect to this new probability? Is it still a Brownian motion?

We can first investigate the law of $B_t$ with respect to Q. Its characteristic function is

$$\mathrm{E}^{\mathrm{Q}}[\mathrm{e}^{i\theta B_t}] = \mathrm{E}[\mathrm{e}^{i\theta B_t} M_T] \, .$$

Let us assume $t \le T$. Then, as $B_t$ is $\mathscr{F}_t$-measurable,

$$\begin{aligned}\mathrm{E}[\mathrm{e}^{i\theta B_t} M_T] &= \mathrm{E}\big[\mathrm{E}[\mathrm{e}^{i\theta B_t} M_T \,|\, \mathscr{F}_t]\big] = \mathrm{E}[\mathrm{e}^{i\theta B_t} M_t] = \mathrm{E}[\mathrm{e}^{(\lambda+i\theta)B_t - \frac{1}{2}\lambda^2 t}] \\ &= \underbrace{\mathrm{E}[\mathrm{e}^{(\lambda+i\theta)B_t - \frac{1}{2}(\lambda+i\theta)^2 t}]}_{=1} \mathrm{e}^{-\frac{1}{2}\theta^2 t + i\lambda\theta t} = \mathrm{e}^{-\frac{1}{2}\theta^2 t + i\lambda\theta t} \, ,\end{aligned}$$

(continued)

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_12

*Example 12.1* (continued)
where the expectation of the quantity over the brace is equal to 1 because we recognize an exponential complex martingale. The quantity on the right-hand side is the characteristic function of an $N(\lambda t, t)$ distribution. Hence, with respect to Q, $B_t \sim N(\lambda t, t)$; it is not a centered r.v. and $(B_t)_t$ cannot be a Brownian motion under Q. What is it then?

The previous computation suggests that $B_t$ might be equal to a Brownian motion plus a motion $t \mapsto \lambda t$. To check this we must prove that $W_t = B_t - \lambda t$ is a Brownian motion with respect to Q. How can this be done? Among the possible ways of checking whether a given process $W$ is a Brownian motion recall Theorem 5.17, reducing the task to the verification that

$$Y_t = \mathrm{e}^{i\theta W_t + \frac{1}{2}\theta^2 t} = \mathrm{e}^{i\theta(B_t - \lambda t) + \frac{1}{2}\theta^2 t}$$

is a Q-martingale for every $\theta \in \mathbb{R}$. This is not difficult: if $s \le t \le T$ and $A \in \mathscr{F}_s$, we have, again first conditioning with respect to $\mathscr{F}_t$,

$$\begin{aligned}
\mathrm{E}^{\mathrm{Q}}[Y_t 1_A] &= \mathrm{E}[Y_t 1_A M_T] = \mathrm{E}\big[Y_t 1_A \mathrm{E}[M_T | \mathscr{F}_t]\big] = \mathrm{E}[Y_t 1_A M_t] \\
&= \mathrm{E}[\mathrm{e}^{i\theta(B_t - \lambda t) + \frac{1}{2}\theta^2 t} \mathrm{e}^{\lambda B_t - \frac{1}{2}\lambda^2 t} 1_A] = \mathrm{E}[\underbrace{\mathrm{e}^{(\lambda + i\theta)B_t - \frac{1}{2}t(\lambda + i\theta)^2}}_{:=Z_t} 1_A]
\end{aligned}$$

$$= \mathrm{E}[Z_s 1_A] = \mathrm{E}[\mathrm{e}^{(\lambda + i\theta)B_s - \frac{1}{2}s(\lambda + i\theta)^2} 1_A] = \mathrm{E}[Y_s M_s 1_A] = \mathrm{E}[Y_s M_T 1_A] = \mathrm{E}^{\mathrm{Q}}[Y_s 1_A] ,$$

as we recognized that $Z$ is a complex martingale with respect to P.

Hence, for $0 \le t \le T$, $B$ is a Brownian motion plus a drift $t \mapsto \lambda t$.

The example above introduces a subtle way of obtaining new processes from old: just change the underlying probability. In this section we develop this idea in full generality, but we shall see that the main ideas are more or less the same as in this example.

From now on let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ denote an $m$-dimensional Brownian motion. Let $\Phi$ be a $\mathbb{C}^m$-valued process in $M^2_{loc}([0, T])$ and

$$Z_t = \exp\Big[\int_0^t \Phi_s \, dB_s - \frac{1}{2}\int_0^t \Phi_s^2 \, ds\Big] , \tag{12.1}$$

where, if $z \in \mathbb{C}^m$, we mean $z^2 = z_1^2 + \dots + z_m^2$. By the same computation leading to (8.40) we have

$$dZ_t = Z_t \Phi_t \, dB_t \tag{12.2}$$

and therefore $Z$ is a local martingale. If $\Phi$ is $\mathbb{R}^m$-valued, $Z$ is a positive supermartingale and $\mathrm{E}(Z_t) \le \mathrm{E}(Z_0) = 1$. If $\mathrm{E}(Z_T) = 1$ for $T > 0$, we can consider on $(\Omega, \mathscr{F})$ the new probability $d\mathrm{Q} = Z_T\, d\mathrm{P}$. Girsanov's Theorem below investigates the nature of the process $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{Q})$ in general.

First let us look for conditions guaranteeing that $Z$ is a martingale. Let $|z|^2 = \sum_{k=1}^m |z_k|^2$ for $z \in \mathbb{C}^m$. If $\mathrm{E}\big[\int_0^T |Z_t\Phi_t|^2\, dt\big] < +\infty$, then $Z\Phi$ is an $m$-dimensional process whose real and imaginary parts are in $M^2([0,T])$ and therefore, recalling (12.2), $Z$ is a complex martingale.

**Proposition 12.1** If $\int_0^T |\Phi_s|^2\, ds \le K$ for some $K \in \mathbb{R}$ then $(Z_t)_{0\le t\le T}$ is a complex martingale bounded in $L^p$ for every $p$.

*Proof* Let $Z^* = \sup_{0\le s\le T} |Z_s|$ and let us show first that $\mathrm{E}(Z^{*p}) < +\infty$. We shall use the elementary relation (see Exercise 1.3)

$$\mathrm{E}(Z^{*p}) = \int_0^{+\infty} pr^{p-1}\mathrm{P}(Z^* > r)\, dr\,. \tag{12.3}$$

Therefore we must just show that $r \mapsto \mathrm{P}(Z^* \ge r)$ goes to 0 fast enough. Observe that, as the modulus of the exponential of a complex number is equal to the exponential of its real part,

$$\begin{aligned}|Z_t| &= \exp\Big[\int_0^t \mathrm{Re}\, \Phi_s\, dB_s - \frac{1}{2}\int_0^t |\mathrm{Re}\, \Phi_s|^2\, ds + \frac{1}{2}\int_0^t |\mathrm{Im}\, \Phi_s|^2\, ds\Big]\\ &\le \exp\Big[\int_0^t \mathrm{Re}\, \Phi_s\, dB_s + \frac{1}{2}\int_0^t |\mathrm{Im}\, \Phi_s|^2\, ds\Big]\,.\end{aligned}$$

Therefore, if $r > 0$, by the exponential inequality (8.41),

$$\begin{aligned}\mathrm{P}(Z^* \ge r) &\le \mathrm{P}\Big(\sup_{0\le t\le T} \exp\Big[\int_0^t \mathrm{Re}\, \Phi_s\, dB_s + \frac{1}{2}\int_0^t |\mathrm{Im}\, \Phi_s|^2\, ds\Big] \ge r\Big)\\ &= \mathrm{P}\Big(\sup_{0\le t\le T} \int_0^t \mathrm{Re}\, \Phi_s\, dB_s \ge \log r - \frac{1}{2}\int_0^T |\mathrm{Im}\, \Phi_s|^2\, ds\Big)\\ &\le \mathrm{P}\Big(\sup_{0\le t\le T} \int_0^t \mathrm{Re}\, \Phi_s\, dB_s \ge \log r - \frac{K}{2}\Big) \le 2\exp\Big[-\frac{(\log r - \frac{K}{2})^2}{2K}\Big]\,.\end{aligned}$$

The right-hand side at infinity is of order $\mathrm{e}^{-c(\log r)^2} = \frac{1}{r^{c\log r}}$ and therefore $r \mapsto \mathrm{P}(Z^* \ge r)$ tends to 0 as $r \to \infty$ faster than $r^{-\alpha}$ for every $\alpha > 0$ and the integral in (12.3) is convergent for every $p > 1$. We deduce that $\Phi Z \in M^2$, as

$$\mathrm{E}\Big(\int_0^T |\Phi_s|^2|Z_s|^2\, ds\Big) \le \mathrm{E}\Big(Z^{*2}\int_0^T |\Phi_s|^2\, ds\Big) \le K\mathrm{E}(Z^{*2}) < +\infty$$

and, thanks to (12.2), $Z$ is a martingale. It is also bounded in $L^p$ for every $p$ as we have seen that $Z^* \in L^p$ for every $p$.

□

If $\Phi$ is $\mathbb{R}^m$-valued, then $Z_t \geq 0$ and if $Z$ is a martingale for $t \in [0, T]$, then $\mathrm{E}(Z_T) = \mathrm{E}(Z_0) = 1$. In this case we can consider on $(\Omega, \mathscr{F}_T)$ the probability Q having density $Z_T$ with respect to P. From now on $\mathrm{E}^{\mathrm{Q}}$ will denote the expectation with respect to Q.

**Lemma 12.1** Let $(Y_t)_t$ be an adapted process. A sufficient condition for it to be an $(\mathscr{F}_t)_t$-martingale with respect to Q is that $(Z_t Y_t)_t$ is an $(\mathscr{F}_t)_t$-martingale with respect to P.

*Proof* Let $s \leq t \leq T$. From Exercise 4.4 we know that

$$\mathrm{E}^{\mathrm{Q}}(Y_t | \mathscr{F}_s) = \frac{\mathrm{E}(Z_T Y_t | \mathscr{F}_s)}{\mathrm{E}(Z_T | \mathscr{F}_s)} \cdot \tag{12.4}$$

But if $(Z_t Y_t)_t$ is an $(\mathscr{F}_t)_t$-martingale with respect to P

$$\mathrm{E}(Z_T Y_t | \mathscr{F}_s) = \mathrm{E}\big[\mathrm{E}(Z_T Y_t | \mathscr{F}_t) | \mathscr{F}_s\big] = \mathrm{E}\big[Y_t \mathrm{E}(Z_T | \mathscr{F}_t) | \mathscr{F}_s\big] = \mathrm{E}(Z_t Y_t | \mathscr{F}_s) = Z_s Y_s$$

so that from (12.4)

$$\mathrm{E}^{\mathrm{Q}}(Y_t | \mathscr{F}_s) = \frac{Z_s Y_s}{Z_s} = Y_s \, .$$

□

**Theorem 12.1 (Girsanov)** Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian and $\Phi \in M^2_{loc}([0, T])$. Let us assume that the process $Z$ defined in (12.1) is a martingale on $[0, T]$ and let Q be the probability on $(\Omega, \mathscr{F})$ having density $Z_T$ with respect to P. Then the process

$$\widetilde{B}_t = B_t - \int_0^t \Phi_s \, ds$$

is an $(\mathscr{F}_t)_t$-Brownian motion on $[0, T]$ with respect to Q.

*Proof* Thanks to Theorem 5.17 we just need to prove that, for every $\lambda \in \mathbb{R}^m$,

$$Y_t^\lambda = \mathrm{e}^{i\langle \lambda, \widetilde{B}_t \rangle + \frac{1}{2}|\lambda|^2 t}$$

is an $(\mathscr{F}_t)_t$-martingale with respect to Q. This follows from Lemma 12.1 if we verify that $X_t^\lambda := Z_t Y_t^\lambda$ is a $(\mathscr{F}_t)_t$-martingale with respect to P. We have

$$\begin{aligned} X_t^\lambda = Z_t Y_t^\lambda &= \exp\Big[\int_0^t \Phi_s\, dB_s - \frac{1}{2}\int_0^t \Phi_s^2\, ds + i\langle \lambda, \widetilde{B}_t\rangle + \frac{1}{2}\,|\lambda|^2 t\Big] \\ &= \exp\Big[\int_0^t (\Phi_s + i\lambda)\, dB_s - \frac{1}{2}\int_0^t \Phi_s^2\, ds - i\int_0^t \langle \lambda, \Phi_s\rangle\, ds + \frac{1}{2}\,|\lambda|^2 t\Big]] \\ &= \exp\Big[\int_0^t (\Phi_s + i\lambda)\, dB_s - \frac{1}{2}\int_0^t (\Phi_s + i\lambda)^2\, ds\Big] \end{aligned}$$

(recall that $z^2$ is the function $\mathbb{C}^m \to \mathbb{C}$ defined as $z^2 = \sum_{k=1}^m z_k^2$). If the r.v. $\int_0^T |\Phi_s|^2\, ds$ is bounded then $X^\lambda$ is a martingale by Proposition 12.1. In general, let

$$\tau_n = \inf\Big\{t \le T; \int_0^t |\Phi_s|^2\, ds > n\Big\} \tag{12.5}$$

and $\Phi_n(s) = \Phi_s 1_{\{s<\tau_n\}}$. We have

$$\begin{aligned} &\exp\Big[\int_0^t \Phi_n(s)\, dB_s - \frac{1}{2}\int_0^t |\Phi_n(s)|^2\, ds\Big] \\ &= \exp\Big[\int_0^{t\wedge\tau_n} \Phi(s)\, dB_s - \frac{1}{2}\int_0^{t\wedge\tau_n} |\Phi(s)|^2\, ds\Big] = Z_{t\wedge\tau_n}\,. \end{aligned}$$

If

$$Y_n^\lambda(t) = \exp\Big[i\Big\langle\lambda, B_t - \int_0^t \Phi_n(s)\, ds\Big\rangle + \frac{1}{2}|\,\lambda|^2 t\Big]\,,$$

then, as $\int_0^T |\Phi_n(s)|^2\, ds \le n$, $X_n^\lambda(t) = Z_{t\wedge\tau_n} Y_n^\lambda(t)$ is a P-martingale by the first part of the proof. In order to show that $X_t^\lambda = Z_t Y_t^\lambda$ is a martingale, we need only to prove that $X_n^\lambda(t) \to X_t^\lambda$ as $n \to \infty$ a.s. and in $L^1$. This will allow us to pass to the limit as $n \to \infty$ in the martingale relation

$$\mathrm{E}[X_n^\lambda(t)\,|\,\mathscr{F}_s] = X_n^\lambda(s) \qquad \text{a.s.}$$

First of all $Z_{t\wedge\tau_n} \to_{n\to\infty} Z_t$ a.s., as $\tau_n \to_{n\to\infty} +\infty$. If $H_n = \min(Z_{t\wedge\tau_n}, Z_t)$, then $H_n \to_{n\to\infty} Z_t$ and, as $0 \le H_n \le Z_t$, by Lebesgue's theorem $\mathrm{E}[H_n] \to \mathrm{E}[Z_t]$. Finally, as $\mathrm{E}[Z_t] = \mathrm{E}[Z_{t\wedge\tau_n}] = 1$ (both $(Z_t)_t$ and $(Z_{t\wedge\tau_n})_t$ are martingales equal to 1 for $t = 0$),

$$\mathrm{E}[|Z_t - Z_{t\wedge\tau_n}|] = \mathrm{E}[Z_t - H_n] + \mathrm{E}[Z_{t\wedge\tau_n} - H_n] = 2\mathrm{E}[Z_t - H_n] \underset{n\to\infty}{\to} 0$$

and therefore $Z_{t\wedge\tau_n} \to_{n\to\infty} Z_t$ in $L^1$. Moreover, note that $Y_n^\lambda(t) \to_{n\to\infty} Y_t^\lambda$ a.s. and $|Y_n^\lambda(t)| \le e^{\frac{1}{2}|\lambda|^2 t}$. Then

$$\begin{aligned} \mathrm{E}\big[|X_n^\lambda(t) - X_t^\lambda|\big] &= \mathrm{E}\big[|Z_n^\lambda(t)Y_n^\lambda(t) - Z_t^\lambda Y_t^\lambda|\big] \\ &\le \mathrm{E}\big[|Z_n^\lambda(t) - Z_t^\lambda||Y_n^\lambda(t)|\big] + \mathrm{E}\big[Z_t^\lambda|Y_n^\lambda(t) - Y_t^\lambda|\big] \end{aligned}$$

and by Lebesgue's theorem both terms on the right-hand side tend to 0 as $n \to \infty$. It follows that, for every $t$, $X_n^\lambda(t) \to_{n\to\infty} X_t^\lambda$ in $L^1$ which concludes the proof.

□

Girsanov's theorem has many applications, as the next examples will show. Note, however, a first important consequence: if $X$ is an Ito process with stochastic differential

$$dX_t = A_t\,dt + G_t\,dB_t$$

then, with the notations of Theorem 12.1,

$$dX_t = (A_t + G_t\Phi_t)\,dt + G_t\,d\widetilde{B}_t\ .$$

As both $(G_t)_t$ and $(\Phi_t)_t$ belong to $M^2_{loc}$, the process $t \mapsto G_t\Phi_t$ belongs to $M^1_{loc}$, so that $X$ is also an Ito process with respect to the new probability Q and with the same stochastic component. Because of the importance of Girsanov's theorem it is useful to have weaker conditions than that of Proposition 12.1 ensuring that $Z$ is a martingale. The following statements provide some sufficient conditions. Let us recall that anyway $Z$ is a positive supermartingale: in order to prove that it is a martingale, it is sufficient to show that $\mathrm{E}(Z_T) = 1$.

**Theorem 12.2** Let $\Phi \in M^2_{loc}([0,T])$, $M_t = \int_0^t \Phi_s\,dB_s$, $0 \le t \le T$, and

$$Z_t = e^{M_t - \frac{1}{2}\langle M\rangle_t}, \quad t \le T\ .$$

Consider the following properties:

a) $\mathrm{E}[e^{\frac{1}{2}\langle M\rangle_T}] = \mathrm{E}[e^{\frac{1}{2}\int_0^T |\Phi_s|^2\,ds}] < +\infty$ (the Novikov's criterion).
b) $(M_t)_{0\le t\le T}$ is a martingale that is bounded in $L^2$ and $\mathrm{E}[e^{\frac{1}{2}M_T}] < +\infty$ (the Kazamaki's criterion).
c) $(Z_t)_{0\le t\le T}$ is a uniformly integrable martingale.

Then a)⇒b)⇒c).

*Proof* If a) is true then the positive r.v. $\langle M\rangle_T$ has a finite Laplace transform at $\theta = \frac{1}{2}$. It is therefore integrable:

$$\mathrm{E}[\langle M\rangle_T] = \mathrm{E}\Big[\int_0^T |\Phi_s|^2\,ds\Big] < +\infty$$

hence $\Phi \in M^2([0,T])$ and $(M_t)_{0\le t\le T}$ is a martingale which is bounded in $L^2$. Moreover, note that

$$\mathrm{e}^{\frac{1}{2}M_T} = Z_T^{1/2}(\mathrm{e}^{\frac{1}{2}\langle M\rangle_T})^{1/2}$$

so that, thanks to the Cauchy–Schwarz inequality,

$$\mathrm{E}[\mathrm{e}^{\frac{1}{2}M_T}] \le \mathrm{E}[Z_T]^{1/2}\mathrm{E}[\mathrm{e}^{\frac{1}{2}\langle M\rangle_T}]^{1/2} \le \mathrm{E}[\mathrm{e}^{\frac{1}{2}\langle M\rangle_T}]^{1/2} < +\infty$$

(recall that $\mathrm{E}[Z_T] \le 1$) so that b) is true.

Let us prove that b)⇒c). By the stopping theorem, for every stopping time $\tau \le T$ we have $M_\tau = \mathrm{E}[M_T\,|\,\mathscr{F}_\tau]$ and by Jensen's inequality

$$\mathrm{e}^{\frac{1}{2}M_\tau} \le \mathrm{E}[\mathrm{e}^{\frac{1}{2}M_T}\,|\,\mathscr{F}_\tau]\,.$$

As $\mathrm{e}^{\frac{1}{2}M_T}$ is assumed to be integrable, the r.v.'s $\mathrm{E}[\mathrm{e}^{\frac{1}{2}M_T}\,|\,\mathscr{F}_\tau]$ with $\tau$ ranging among the stopping times that are smaller than $T$ form a uniformly integrable family (Proposition 5.4). Hence the family of r.v.'s $\mathrm{e}^{\frac{1}{2}M_\tau}$ (which are positive) is also uniformly integrable.

Let $0 < a < 1$ and $Y_t^{(a)} = \mathrm{e}^{\frac{a}{1+a}M_t}$. We have

$$\mathrm{e}^{aM_t-\frac{1}{2}a^2\langle M\rangle_t} = \big(\mathrm{e}^{M_t-\frac{1}{2}\langle M\rangle_t}\big)^{a^2}\big(Y_t^{(a)}\big)^{1-a^2}\,.$$

If $A \in \mathscr{F}_T$ and $\tau \le T$ is a stopping time we have by Hölder's inequality, as $\frac{1+a}{2a} > 1$,

$$\begin{aligned}\mathrm{E}[1_A\mathrm{e}^{aM_\tau-\frac{1}{2}a^2\langle M\rangle_\tau}] &\le \mathrm{E}[\mathrm{e}^{M_\tau-\frac{1}{2}\langle M\rangle_\tau}]^{a^2}\mathrm{E}[1_A Y_\tau^{(a)}]^{1-a^2}\\ \le \mathrm{E}[1_A Y_\tau^{(a)}]^{1-a^2} &\le \mathrm{E}[1_A (Y^{(a)})^{\frac{1+a}{2a}}]^{\frac{2a}{1+a}\cdot(1-a^2)} = \mathrm{E}[1_A\mathrm{e}^{\frac{1}{2}M_\tau}]^{a(1-a)}\,.\end{aligned} \tag{12.6}$$

We used here the relation $\mathrm{E}[\mathrm{e}^{M_\tau-\frac{1}{2}\langle M\rangle_\tau}] = \mathrm{E}[Z_\tau] \le \mathrm{E}[Z_T] \le 1$, which follows from the stopping theorem applied to the supermartingale $Z$.

As we know that the family $(\mathrm{e}^{\frac{1}{2}M_\tau})_\tau$ for $\tau$ ranging among the stopping times that are smaller than $T$ forms a uniformly integrable family, by (12.6) and using the criterion of Proposition 5.2, the same is true for the r.v.'s $\mathrm{e}^{aM_\tau-\frac{1}{2}a^2\langle M\rangle_\tau}$.

Let us prove that $(\mathrm{e}^{aM_t-\frac{1}{2}a^2\langle M\rangle_t})_t$ is a uniformly integrable (true) martingale: as we know that it is a local martingale, let $(\tau_n)_n$ be a sequence of reducing stopping times. Then for every $s \le t$, $A \in \mathscr{F}_s$,

$$\mathrm{E}[\mathrm{e}^{aM_{t\wedge\tau_n}-\frac{1}{2}a^2\langle M\rangle_{t\wedge\tau_n}}1_A] = \mathrm{E}[\mathrm{e}^{aM_{s\wedge\tau_n}-\frac{1}{2}a^2\langle M\rangle_{s\wedge\tau_n}}1_A]\,.$$

As $e^{aM_{t\wedge\tau_n}-\frac{1}{2}a^2\langle M\rangle_{t\wedge\tau_n}}1_A \to_{n\to+\infty} e^{aM_t-\frac{1}{2}a^2\langle M\rangle_t}1_A$ and these r.v.'s are uniformly integrable, we can take the limit in the expectations and obtain the martingale relation

$$\mathrm{E}[e^{aM_t-\frac{1}{2}a^2\langle M\rangle_t}1_A] = \mathrm{E}[e^{aM_s-\frac{1}{2}a^2\langle M\rangle_s}1_A] .$$

Again by Jensen's inequality, repeating the argument above for $\tau = T$, $A = \Omega$, we have

$$1 = \mathrm{E}[e^{aM_T-\frac{1}{2}a^2\langle M\rangle_T}] \le \mathrm{E}[e^{M_T-\frac{1}{2}\langle M\rangle_T}]^{a^2}\mathrm{E}[Y_T^{(a)}]^{1-a^2} \le \mathrm{E}[e^{M_T-\frac{1}{2}\langle M\rangle_T}]^{a^2}\mathrm{E}[e^{\frac{1}{2}M_T}]^{a(1-a)}.$$

Letting $a \to 1$ this inequality becomes $\mathrm{E}[e^{M_T-\frac{1}{2}\langle M\rangle_T}] \ge 1$, which allows us to conclude the proof.

□

**Corollary 12.1** With the notations of Theorem 12.2, if for some $\mu > 0$ $\mathrm{E}[e^{\mu|\Phi_t|^2}] < C < +\infty$ for every $0 \le t \le T$, then $(Z_t)_{0\le t\le T}$ is a martingale.

*Proof* By Jensen's inequality (the exponential is a convex function)

$$\begin{aligned}\exp\Big[\frac{1}{2}\int_{t_1}^{t_2}|\Phi_s|^2\,ds\Big] &= \exp\Big[\frac{1}{t_2-t_1}\int_{t_1}^{t_2}\frac{1}{2}(t_2-t_1)|\Phi_s|^2\,ds\Big]\\ &\le \frac{1}{t_2-t_1}\int_{t_1}^{t_2}e^{\frac{1}{2}(t_2-t_1)|\Phi_s|^2}\,ds .\end{aligned}$$

Let $t_2 - t_1 < 2\mu$. With this constraint we have

$$\mathrm{E}\Big[\exp\Big(\frac{1}{2}\int_{t_1}^{t_2}|\Phi_s|^2\,ds\Big)\Big] \le \frac{1}{t_2-t_1}\int_{t_1}^{t_2}\mathrm{E}[e^{\mu|\Phi_s|^2}]\,ds < +\infty .$$

Let $0 = t_1 < t_2 < \dots < t_{n+1} = T$ be chosen so that $t_{i+1} - t_i < 2\mu$ and let $\Phi_i(s) = \Phi_s 1_{[t_i,t_{i+1}[}(s)$. Each of the processes $\Phi_i$, $i = 1, \dots, n$, satisfies condition a) of Theorem 12.2 as

$$\mathrm{E}\Big[e^{\frac{1}{2}\int_0^T|\Phi_i(s)|^2\,ds}\Big] = \mathrm{E}\Big[\exp\Big(\frac{1}{2}\int_{t_i}^{t_{i+1}}|\Phi_s|^2\,ds\Big)\Big] < +\infty$$

and if

$$Z_t^i = \exp\Big[\int_0^t\Phi_i(s)\,dB_s - \frac{1}{2}\int_0^t|\Phi_i(s)|^2\,ds\Big]$$

then $Z^i$ is a martingale. We have $Z_T = Z_T^1 \dots Z_T^n$; moreover, $Z_{t_i}^i = 1$, as $\Phi_i = 0$ on $[0, t_i[$ and therefore $\mathrm{E}(Z_T^i \,|\, \mathscr{F}_{t_i}) = Z_{t_i}^i = 1$. Hence it is easy now to deduce that $\mathrm{E}[Z_T] = 1$ as

$$\mathrm{E}(Z_T) = \mathrm{E}(Z_T^1 \dots Z_T^n) = \mathrm{E}[Z_T^1 \dots Z_T^{n-1} \mathrm{E}(Z_T^n \,|\, \mathscr{F}_{t_n})] = \mathrm{E}(Z_T^1 \dots Z_T^{n-1}) = \dots = 1$$

which allows us to conclude the proof.

□

*Example 12.2* Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion. Let us prove that

$$Z_t = \exp\Big(\theta \int_0^t B_s\, dB_s - \frac{\theta^2}{2} \int_0^t B_s^2\, ds\Big)$$

is a martingale for $t \in [0, T]$.

Novikov's criterion requires us to prove that

$$\mathrm{E}\big[\mathrm{e}^{\frac{\theta^2}{2} \int_0^T B_s^2\, ds}\big] < +\infty\,.$$

Here, however, it is much easier to apply the criterion of Corollary 12.1, which requires us to check that for some $\mu > 0$

$$\mathrm{E}\big[\mathrm{e}^{\mu \theta^2 B_t^2}\big] < +\infty \tag{12.7}$$

for every $t \in [0, T]$. Recalling Remark 3.3 we have that if $2\mu\theta^2 T < 1$ then $\mathrm{E}[\mathrm{e}^{\mu\theta^2 B_t^2}] = (1 - 2\mu\theta^2 T)^{-1/2}$, so that (12.7) is satisfied as soon as

$$\mu < \frac{1}{2\theta^2 T}\,\cdot$$

Therefore the criterion of Corollary 12.1 is satisfied and $(Z_t)_{0 \le t \le T}$ is a martingale.

We can therefore consider on $\mathscr{F}$ the probability Q having density $Z_T$ with respect to P. What can be said of $B$ under this new probability Q?

Girsanov's theorem states that, for $0 \le t \le T$,

$$W_t = B_t - \theta \int_0^t B_s\, dB_s$$

(continued)

*Example 12.2* (continued)
is a Q-Brownian motion. Therefore, under Q, $(B_t)_{0\le t\le T}$ has the differential

$$dB_t = \theta B_t\,dt + dW_t$$

and, with respect to Q, $(B_t)_{0\le t\le T}$ is an Ornstein–Uhlenbeck process.

*Example 12.3* As an example of application of Girsanov's theorem, let us investigate the distribution of the passage time $\tau_a$ at $a > 0$ of a Brownian motion with a drift $\mu$. Recall that the distribution of $\tau_a$ for the Brownian motion has been computed in Exercise 3.20. Let $B$ be a Brownian motion and $Z_t = \mathrm{e}^{\mu B_t - \frac{\mu^2}{2}t}$. Then $Z$ is a martingale and, by Girsanov's formula, if $d\mathrm{Q} = Z_t\,d\mathrm{P}$, then, with respect to Q, $W_s = B_s - \mu s$ is, for $s \le t$, a Brownian motion. Hence, with respect to Q, $B_s = W_s + \mu s$.
If $\tau_a = \inf\{t; B_t \ge a\}$, let us compute $\mathrm{Q}(\tau_a \le t)$, which is then the partition function of the passage time of a Brownian motion with drift $\mu$. We have

$$\mathrm{Q}(\tau_a \le t) = \mathrm{E}^{\mathrm{Q}}[1_{\{\tau_a\le t\}}] = \mathrm{E}^{\mathrm{P}}\big[1_{\{\tau_a\le t\}} Z_t\big]\,.$$

Recall now that $\{\tau_a \le t\} \in \mathscr{F}_{\tau_a} \cap \mathscr{F}_t = \mathscr{F}_{\tau_a\wedge t}$ and that $Z$ is a martingale. Therefore, by the stopping theorem for martingales,

$$\mathrm{E}^{\mathrm{P}}\big[1_{\{\tau_a\le t\}} Z_t\big] = \mathrm{E}^{\mathrm{P}}\big[\mathrm{E}[1_{\{\tau_a\le t\}} Z_t\,|\,\mathscr{F}_{\tau_a\wedge t}]\big] = \mathrm{E}^{\mathrm{P}}\big[1_{\{\tau_a\le t\}} Z_{\tau_a\wedge t}\big] = \mathrm{E}^{\mathrm{P}}\big[1_{\{\tau_a\le t\}} Z_{\tau_a}\big]\,,$$

the last equality coming from the fact that, of course, on $\{\tau_a \le t\}$ we have $\tau_a \wedge t = \tau_a$. But

$$Z_{\tau_a} = \mathrm{e}^{\mu B_{\tau_a} - \frac{\mu^2}{2}\tau_a} = \mathrm{e}^{\mu a - \frac{\mu^2}{2}\tau_a}$$

and, under P, $B$ is a Brownian motion so that the law of $\tau_a$ under P is given by (3.23). In conclusion,

$$\begin{aligned}\mathrm{Q}(\tau_a \le t) &= \mathrm{e}^{\mu a}\mathrm{E}^{\mathrm{P}}[1_{\{\tau_a\le t\}}\mathrm{e}^{-\frac{\mu^2}{2}\tau_a}]\\ &= \mathrm{e}^{\mu a}\int_0^t \frac{a}{(2\pi)^{1/2}\,s^{3/2}}\,\mathrm{e}^{-a^2/2s}\,\mathrm{e}^{-\frac{\mu^2}{2}s}\,ds \qquad (12.8)\\ &= \int_0^t \frac{a}{(2\pi)^{1/2}\,s^{3/2}}\,\mathrm{e}^{-\frac{1}{2s}(\mu s-a)^2}\,ds\,.\end{aligned}$$

(continued)

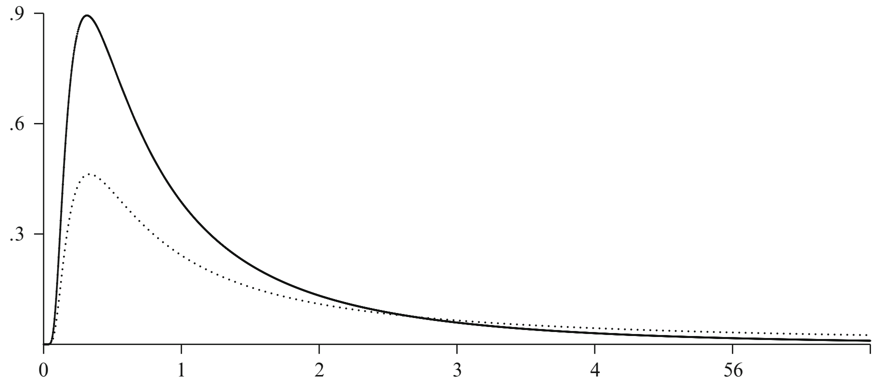


**Fig. 12.1** The graph of the densities of the passage time $\tau_a$ for $a = 1$ and $\mu = \frac{3}{4}$ (solid) and $\mu = 0$ (dots). The first one decreases much faster as $t \to +\infty$

*Example 12.3* (continued)
If $\mu > 0$, then we know that $\tau_a$ is Q-a.s. finite and taking the derivative we find that the passage time $\tau_a$ of the Brownian motion with a positive drift $\mu$ has density (Fig. 12.1)

$$f(t) = \frac{a}{(2\pi)^{1/2}\, t^{3/2}}\, \mathrm{e}^{-\frac{1}{2t}\,(\mu t - a)^2}\; . \tag{12.9}$$

In particular,

$$\int_0^{+\infty} \frac{a}{(2\pi)^{1/2}\, s^{3/2}}\, \mathrm{e}^{-\frac{1}{2s}\,(\mu s - a)^2}\, ds = 1\; . \tag{12.10}$$

Note the rather strange fact that the integral in (12.10) does not depend on $\mu$, as far as $\mu \geq 0$.

For $\mu > 0$ the passage time has a finite expectation, as its density goes to 0 at $+\infty$ exponentially fast. This is already known: in Exercise 5.33 we obtained, using a Laplace transform argument, that $\mathrm{E}[\tau_a] = \frac{a}{\mu}$.

If $\mu < 0$, conversely, we know that $\tau_a = +\infty$ with probability $1 - \mathrm{e}^{2\mu a}$ (Exercise 5.17).

*Example 12.4 (Same as the previous one, but from a different point of view)*
Imagine we want to compute the quantity

$$\mathrm{E}[f(\tau_a)]$$

(continued)

*Example 12.4* (continued)
$\tau_a$ being the passage time at $a$ of a Brownian motion and $f$ some bounded measurable function. Imagine, moreover, that we do not know the distribution of $\tau_a$ (or that the computation of the integral with respect to its density is too complicated) so that we are led to make this computation by simulation. This means that we are going to simulate $N$ paths of the Brownian motion and for each one of these to note the passage times, $T_1, \dots, T_N$ say. Then by the law of large numbers

$$\frac{1}{n}\sum_{j=1}^{N} f(T_i)$$

is, for $N$ large, an approximation of $\mathrm{E}[f(\tau_a)]$.

However, in practice this procedure is unfeasible as $\tau_a$ is finite but can take very large values (recall that $\tau_a$ does not have finite expectation). Therefore some of the values $T_i$ may turn out to be very large, so large that the computer can stay on a single path for days before it reaches the level $a$ (see Exercise 3.20 b)).

The computation of the previous example suggests that we can do the following: instead of a Brownian motion, let us simulate a Brownian motion with a positive drift $\mu$, collect the passage times, $T_1^\mu, \dots, T_N^\mu$ say, and consider the quantity

$$\frac{\mathrm{e}^{-\mu a}}{N}\sum_{j=1}^{N} f(T_i^\mu)\mathrm{e}^{\frac{\mu^2}{2}T_i^\mu}\;.$$

Let us check that $\mathrm{e}^{-\mu a}\mathrm{E}[f(T_i^\mu)\mathrm{e}^{\frac{\mu^2}{2}T_i^\mu}] = \mathrm{E}[f(\tau_a)]$ so that, again by the law of large numbers, this is also an estimate of $\mathrm{E}[f(\tau_a)]$. Actually, denoting $g_\mu, g_0$ the densities of the r.v.'s $T_i^\mu$ and $T_i$ ,respectively (given by (12.9) and (3.23) respectively), we have

$$\begin{aligned}
\mathrm{e}^{-\mu a}\mathrm{E}\big[f(T_i^\mu)\mathrm{e}^{\frac{\mu^2}{2}T_i^\mu}\big] &= \mathrm{e}^{-\mu a}\int_0^{+\infty} f(t)\,\mathrm{e}^{\frac{\mu^2}{2}t} g_\mu(t)\,dt\\
&= \mathrm{e}^{-\mu a}\int_0^{+\infty} f(t)\,\mathrm{e}^{\frac{\mu^2}{2}t}\frac{a}{(2\pi)^{1/2}\,t^{3/2}}\,\mathrm{e}^{-\frac{1}{2t}(\mu t-a)^2}\,dt\\
&= \int_0^{+\infty} f(t)\frac{a}{(2\pi)^{1/2}\,t^{3/2}}\,\mathrm{e}^{-\frac{a^2}{2t}}\,dt\\
&= \int_0^{+\infty} f(t)g_0(t)\,dt = \mathrm{E}\big[f(T_i)\big]\,.
\end{aligned}$$

(continued)

*Example 12.4* (continued)
Now, as the passage time of the Brownian motion with a positive drift $\mu$ has finite expectation, the simulation of these paths is much quicker and the problem mentioned above disappears.

In other words, Girsanov's theorem can by applied to simulation in the following way: instead of simulating a process $X$ ($B$ in the example above), which can lead to problems, simulate instead another process $Y$ ($B_t + \mu t$ above) that is less problematic and then compensate with the density of the law of $X$ with respect to the law of $Y$.

*Example 12.5* Let $X$ be a Brownian motion and $\mu$ a real number. Let $a > 0$ and

$$\tau = \inf\{t; X_t \geq a + \mu t\} \ .$$

We want to compute $\mathrm{P}(\tau \leq T)$, i.e. the probability for $X$ to cross the linear barrier $t \mapsto a + \mu t$ before time $T$. For $\mu = 0$ we already know the answer thanks to the reflection principle, Corollary 3.4. Thanks to Example 12.3 we know, for $\mu < 0$, the density of $\tau$ so we have

$$\mathrm{P}(\tau \leq T) = \int_0^T \frac{a}{(2\pi)^{1/2}\, t^{3/2}}\, \mathrm{e}^{-\frac{1}{2t}\,(-\mu t - a)^2}\, dt \ ,$$

but let us avoid the computation of the primitive by taking another route and applying Girsanov's theorem directly. The idea is to write

$$\mathrm{P}(\tau \leq T) = \mathrm{P}\Big(\sup_{0\leq t\leq T} (X_t - \mu t) \geq a\Big)$$

and then to make a change of probability so that, with respect to the new probability, $t \mapsto X_t - \mu t$ is a Brownian motion for which known formulas are available. Actually, if

$$Z_T = \mathrm{e}^{\mu X_T - \frac{1}{2}\mu^2 T}$$

and $d\mathrm{Q} = Z_T\, d\mathrm{P}$, then with respect to Q the process

$$W_t = X_t - \mu\, t$$

(continued)

*Example 12.5* (continued)
is a Brownian motion up to time $T$. Hence

$$\begin{aligned}
\mathrm{P}\Big(\sup_{0\le s\le T}(X_s-\mu s)\ge a\Big) &= \mathrm{E}^{\mathrm{Q}}\Big[Z_T^{-1}1_{\{\sup_{0\le s\le T}(X_s-\mu s)\ge a\}}\Big]\\
&= \mathrm{E}^{\mathrm{Q}}\Big[\mathrm{e}^{-\mu X_T+\frac12\mu^2T}1_{\{\sup_{0\le s\le T}(X_s-\mu s)\ge a\}}\Big]\\
&= \mathrm{e}^{-\frac12\mu^2T}\mathrm{E}^{\mathrm{Q}}\Big[\mathrm{e}^{-\mu W_T}1_{\{\sup_{0\le s\le T}W_s\ge a\}}\Big]\,.
\end{aligned}$$

The expectation on the right-hand side can be computed analytically as, from Corollary 3.3, we know the joint density of $W_T$ and $\sup_{0\le s\le T}W_s$, i.e.

$$\begin{aligned}
&\mathrm{P}\Big(\sup_{0\le s\le T}(X_s-\mu s)\ge a\Big)\\
&=\frac{2}{\sqrt{2\pi T^3}}\,\mathrm{e}^{-\frac12\mu^2T}\int_a^{+\infty}ds\int_{-\infty}^{s}(2s-x)\,\mathrm{e}^{-\frac{1}{2T}(2s-x)^2}\mathrm{e}^{-\mu x}\,dx\,.
\end{aligned}$$

With the change of variable $y=2s-x$ and using Fubini's theorem

$$\begin{aligned}
\mathrm{P}\Big(\sup_{0\le s\le T}(X_s-\mu s)\ge a\Big) &= \frac{2}{\sqrt{2\pi T^3}}\,\mathrm{e}^{-\frac12\mu^2T}\int_a^{+\infty}ds\int_s^{+\infty}y\mathrm{e}^{-\frac{1}{2T}y^2}\mathrm{e}^{-\mu(2s-y)}\,dy\\
&=\frac{2}{\sqrt{2\pi T^3}}\,\mathrm{e}^{-\frac12\mu^2T}\int_a^{+\infty}y\mathrm{e}^{-\frac{1}{2T}y^2}\mathrm{e}^{\mu y}\,dy\int_a^{y}\mathrm{e}^{-2\mu s}\,ds\\
&=\frac{1}{\mu\sqrt{2\pi T^3}}\,\mathrm{e}^{-\frac12\mu^2T}\int_a^{+\infty}y\mathrm{e}^{-\frac{1}{2T}y^2}\mathrm{e}^{\mu y}(\mathrm{e}^{-2\mu a}-\mathrm{e}^{-2\mu y})\,dy\,.
\end{aligned}$$

Now, integrating by parts,

$$\begin{aligned}
\mathrm{P}\Big(\sup_{0\le s\le T}(X_s-\mu s)\ge a\Big) &= -\frac{1}{\mu\sqrt{2\pi T}}\,\mathrm{e}^{-\frac12\mu^2T}\underbrace{\mathrm{e}^{-\frac{1}{2T}y^2}\mathrm{e}^{\mu y}(\mathrm{e}^{-2\mu a}-\mathrm{e}^{-2\mu y})\Big|_a^{+\infty}}_{=0}\\
&\quad+\frac{1}{\sqrt{2\pi T}}\,\mathrm{e}^{-\frac12\mu^2T}\int_a^{+\infty}\mathrm{e}^{-\frac{1}{2T}y^2}(\mathrm{e}^{\mu(y-2a)}+\mathrm{e}^{-\mu y})\,dy
\end{aligned}$$

and with the change of variable $z=\frac{1}{\sqrt{T}}(y-\mu T)$

$$\frac{1}{\sqrt{2\pi T}}\,\mathrm{e}^{-2\mu a}\mathrm{e}^{-\frac12\mu^2T}\int_a^{+\infty}\mathrm{e}^{-\frac{1}{2T}y^2}\mathrm{e}^{\mu y}\,dy$$

(continued)

*Example 12.5* (continued)

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi T}}\,\mathrm{e}^{-2\mu a}\int_a^{+\infty}\mathrm{e}^{-\frac{1}{2T}(y-\mu T)^2}\,dy\\
&= \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-2\mu a}\int_{(a-\mu T)/\sqrt{T}}^{+\infty}\mathrm{e}^{-z^2}\,dz = \mathrm{e}^{-2\mu a}\Big(1-\Phi\Big(\frac{a-\mu T}{\sqrt{T}}\Big)\Big)\,.
\end{aligned}
$$

Similarly

$$
\frac{1}{\sqrt{2\pi T}}\,\mathrm{e}^{-\frac{1}{2}\mu^2 T}\int_a^{+\infty}\mathrm{e}^{-\frac{1}{2T}y^2}\mathrm{e}^{-\mu y}\,dy\frac{1}{\sqrt{2\pi T}}\int_a^{+\infty}\mathrm{e}^{-\frac{1}{2T}(y+\mu T)^2}\,dy
$$

$$
\frac{1}{\sqrt{2\pi}}\int_{(a+\mu T)/\sqrt{T}}^{+\infty}\mathrm{e}^{-z^2/2}\,dy = \Big(1-\Phi\Big(\frac{a+\mu T}{\sqrt{T}}\Big)\Big)
$$

and finally

$$
\begin{aligned}
&\mathrm{P}\Big(\sup_{0\le s\le T}(X_s-\mu s)\ge a\Big)\\
&= \mathrm{e}^{-2\mu a}\Big(1-\Phi\Big(\frac{a-\mu T}{\sqrt{T}}\Big)\Big)+\Big(1-\Phi\Big(\frac{a+\mu T}{\sqrt{T}}\Big)\Big)\,.
\end{aligned}
\tag{12.11}
$$

## 12.2 The Cameron–Martin formula

In this section we see how we can use Girsanov's Theorem 12.1 in order to construct weak solutions, in the sense of Definition 9.1, for SDEs that may not satisfy Assumption (A').

Let us assume that $\sigma$ satisfies Assumption (A') and, moreover, that it is a symmetric $d\times d$ matrix field and that, for every $(x,t)$, the smallest eigenvalue of $\sigma(x,t)$ is bounded below by a constant $\lambda>0$. The last hypothesis implies that the matrix field $(x,t)\mapsto\sigma(x,t)^{-1}$ is well defined and bounded. Let $b:\mathbb{R}^m\times\mathbb{R}^+\to\mathbb{R}^m$ be a bounded measurable vector field. We know, by Theorem 9.4, that there exists a solution of the SDE

$$
\begin{aligned}
d\xi_t &= \sigma(\xi_t,t)\,dB_t\\
\xi_0 &= x
\end{aligned}
\tag{12.12}
$$

which, moreover, is unique in law. Let now, for a fixed $T>0$,

$$
Z_T = \exp\Big[\int_0^T\sigma^{-1}(\xi_s,s)b(\xi_s,s)\,dB_s-\frac{1}{2}\int_0^T|\sigma^{-1}(\xi_s,s)b(\xi_s,s)|^2\,ds\Big]\,. \tag{12.13}
$$

As $\sigma^{-1}$ and $b$ are bounded, by Proposition 12.1 and Theorem 12.1 $\mathrm{E}[Z_T] = 1$ and if $\widetilde{\mathrm{P}}$ is the probability on $(\Omega, \mathscr{F})$ with density $Z_T$ with respect to P and

$$\widetilde{B}_t = B_t - \int_0^t \sigma^{-1}(\xi_s, s) b(\xi_s, s)\, ds\,,$$

then $\widetilde{B} = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\le t\le T}, (\widetilde{B}_t)_{0\le t\le T}, \widetilde{\mathrm{P}})$ is a Brownian motion; hence (12.12) can be written as

$$\begin{aligned} d\xi_t &= b(\xi_t, t)\, dt + \sigma(\xi_t, t)\, d\widetilde{B}_t \\ \xi_0 &= x\,. \end{aligned} \tag{12.14}$$

Therefore, with respect to $\widetilde{\mathrm{P}}$, $\xi$ is a solution of (12.14).

**Theorem 12.3** Under the previous assumptions equation (12.14) admits a weak solution on $[0, T]$ for every $T$ and there is uniqueness in law.

*Proof* The existence has already been proved. As for the uniqueness the idea of the proof is that for Eq. (12.12) there is uniqueness in law (as Assumption (A') is satisfied) and that the law of the solution of (12.14) has a density with respect to the law of the solution of (12.12).

Let $(\Omega', \mathscr{F}', (\mathscr{F}'_t)_{0\le t\le T}, (\xi'_t)_{0\le t\le T}, (\widetilde{B}'_t)_{0\le t\le T}, \widetilde{\mathrm{P}}')$ be another solution of (12.14), i.e.

$$d\xi'_t = b(\xi'_t, t)\, dt + \sigma(\xi'_t, t)\, d\widetilde{B}'_t\,.$$

We must prove that $\xi'$ has the same law as the solution $\xi$ constructed in (12.14). Let

$$\widetilde{Z}'_T = \exp\Big[ -\int_0^T \sigma^{-1}(\xi'_s, s) b(\xi'_s, s)\, d\widetilde{B}_s - \frac{1}{2} \int_0^T |\sigma^{-1}(\xi'_s, s) b(\xi'_s, s)|^2\, ds \Big]$$

and let $\mathrm{P}'$ be the probability on $(\Omega', \mathscr{F}')$ having density $\widetilde{Z}'_T$ with respect to $\widetilde{\mathrm{P}}'$. Then, by Proposition 12.1 and Theorem 12.1, the process

$$B'_t = \widetilde{B}'_t + \int_0^t \sigma^{-1}(\xi'_s, s) b(\xi'_s, s)\, ds$$

is a $\mathrm{P}'$-Brownian motion. Expressing $B'$ in terms of $\widetilde{B}'$ in (12.14), it turns out that $(\Omega', \mathscr{F}', (\mathscr{F}'_t)_{0\le t\le T}, (\xi'_t)_{0\le t\le T}, (B'_t)_{0\le t\le T}, \mathrm{P}')$ is a solution of (12.12). Moreover, $\widetilde{\mathrm{P}}'$ has density with respect to $\mathrm{P}'$ given by

$$\begin{aligned} Z'_T = \widetilde{Z}'^{-1}_T &= \exp\Big[ \int_0^T \sigma^{-1}(\xi'_s, s) b(\xi'_s, s)\, dB'_s + \frac{1}{2} \int_0^T |\sigma^{-1}(\xi'_s, s) b(\xi'_s, s)|^2\, ds \Big] \\ &= \exp\Big[ \int_0^T \sigma^{-1}(\xi'_s, s) b(\xi'_s, s)\, d\widetilde{B}'_s - \frac{1}{2} \int_0^T |\sigma^{-1}(\xi'_s, s) b(\xi'_s, s)|^2\, ds \Big]\,. \end{aligned}$$

As Eq. (12.12) satisfies Assumption (A'), there is uniqueness in law and the laws $\xi(\mathrm{P})$ and $\xi'(\mathrm{P}')$ coincide. The law of the solution of (12.14) is $\xi(\mathrm{P})$ (image of P through $\xi$), but

$$\xi(\mathrm{P}) = \xi(Z_T \cdot \widetilde{\mathrm{P}}) \ .$$

Also the joint laws of $(\xi, Z_T)$ and $(\xi', Z'_T)$ (Lemma 9.2) coincide and from this it follows (Exercise 4.3) that the laws $\xi(Z_T\mathrm{P}) = \xi(\widetilde{\mathrm{P}})$ and $\xi'(Z'_T\mathrm{P}) = \xi'(\widetilde{\mathrm{P}'})$ coincide, which is what had to be proved.

□

The arguments of this section stress that, under the hypotheses under consideration, if $\xi$ is a solution of (12.14), then its law is absolutely continuous with respect to the law of the solution of (12.12). More precisely, with the notations of the proof of Theorem 12.3, if $A$ is a Borel set of the paths space $\mathscr{C}$,

$$\mathrm{P}(\xi \in A) = \mathrm{E}^{\widetilde{\mathrm{P}}}(Z_T 1_{\{\xi \in A\}}) \ , \tag{12.15}$$

where $Z_T$ is defined in (12.13). (12.15) is the *Cameron–Martin formula*. In fact, it holds under hypotheses weaker than those considered here. These are actually rather restrictive, since often one is led to the consideration of diffusions having an unbounded drift or a diffusion coefficient that it is not uniformly invertible. In this case the vector field $(x,t) \mapsto \sigma(x,t)^{-1}b(x,t)$ is not bounded and therefore Proposition 12.1 cannot be used to obtain that $\mathrm{E}(Z_T) = 1$; it will be necessary to turn instead to Proposition 12.2. In Exercise 12.5 the absolute continuity of the laws of the Ornstein–Uhlenbeck process with respect to Wiener measure is considered: this is a simple example where the hypotheses of this section do not hold, the drift being linear.

## 12.3 The martingales of the Brownian filtration

Let $B$ be an $m$-dimensional Brownian motion with respect to a filtration $(\mathscr{F}_t)_t$. We have seen in Sect. 8.4 that the stochastic integral of $B$ with respect to an integrand of $M^2$ is a square integrable $(\mathscr{F}_t)_t$-martingale. Is the converse also true? Is it true that every square integrable $(\mathscr{F}_t)_t$-martingale is a stochastic integral with respect to some integrand of $M^2$?

We see in this section that the answer is yes if $(\mathscr{F}_t)_t = (\overline{\mathscr{G}}_t)_t$, the natural filtration augmented with the events of probability zero of $\mathscr{F}$ which was introduced in Remark 3.1. This is a particularly important representation result with deep applications, some of which will appear in Chap. 13. Now just observe that this result implies that every $(\overline{\mathscr{G}}_t)_t$-martingale is continuous (has a continuous modification, to be precise). We know from Proposition 4.3 that $(\overline{\mathscr{G}}_t)_t$ is right-continuous, so that $(\Omega, \mathscr{F}, (\overline{\mathscr{G}}_t)_t, (B_t)_t, \mathrm{P})$ is a standard Brownian motion.

The key result is the following, which we have already mentioned in Remark 7.2.

**Theorem 12.4** Let $B = (\Omega, \mathscr{F}, (\overline{\mathscr{G}}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion with respect to its augmented natural filtration and let $T > 0$. Then every r.v. $Z \in L^2 = L^2(\Omega, \overline{\mathscr{G}}_T, \mathrm{P})$ is of the form

$$Z = c + \int_0^T H_s \, dB_s \,, \tag{12.16}$$

where $c \in \mathbb{R}$ and $H \in M^2([0, T])$ (in particular, $H$ is $(\overline{\mathscr{G}}_t)_t$-adapted). Moreover, the representation (12.16) is unique.

*Proof* The uniqueness is obvious, as $c$ is determined by $c = \mathrm{E}[Z]$ whereas, if $H_1$ and $H_2$ were two processes in $M^2([0, T])$ satisfying (12.16), then from the relation

$$\int_0^T \big(H_1(s) - H_2(s)\big) \, dB_s = 0$$

we have immediately, by the isometry property of the stochastic integral, that

$$\mathrm{E}\Big(\int_0^T \big|H_1(s) - H_2(s)\big|^2 \, ds\Big) = 0$$

and therefore $H_1(s) = H_2(s)$ for almost every $s \in [0, T]$ a.s.

As for the existence, let us denote by $\mathscr{H}$ the space of the r.v.'s of the form (12.16) and let us prove first that $\mathscr{H}$ is a closed subset of $L^2$. If $(Z_n)_n \subset \mathscr{H}$ is a Cauchy sequence in $L^2$ and

$$Z_n = c_n + \int_0^T H_n(s) \, dB_s = c_n + I_n$$

then $(c_n)_n$ is also a Cauchy sequence, as $c_n = \mathrm{E}[Z_n]$ and

$$|c_n - c_m| \le \mathrm{E}(|Z_n - Z_m|) \le \mathrm{E}(|Z_n - Z_m|^2)^{1/2} \,.$$

Therefore both $(c_n)_n$ and $(I_n)_n$ are Cauchy sequences (in $\mathbb{R}$ and in $L^2$ respectively). As the stochastic integral is an isometry between $L^2$ and $M^2([0, T])$ (Theorem 7.1), it follows that $(H_n)_n$ is a Cauchy sequence in $M^2([0, T])$. Therefore there exist $c \in \mathbb{R}$ and $H \in M^2([0, T])$ such that

$$\begin{aligned} &\lim_{n\to\infty} c_n = c \\ &\lim_{n\to\infty} H_n = H \qquad \text{in } M^2([0, T]) \end{aligned}$$

whence we get

$$\lim_{n\to\infty} Z_n = c + \int_0^T H_s\,dB_s \qquad \text{in } L^2\,.$$

Let us prove now that $\mathscr{H}$ is also dense in $L^2$. If $f \in L^2([0,T],\mathbb{R}^m)$ ($f$ is therefore a deterministic function), for $t \le T$, and we set

$$e_f(t) := \exp\Big(\int_0^t f(s)\,dB_s - \frac{1}{2}\int_0^t |f(s)|^2\,ds\Big)$$

then we know that

$$e_f(T) = 1 + \int_0^T e_f(s)f(s)\,dB_s\,. \tag{12.17}$$

Moreover, $e_f(T)$ is square integrable, as $\int_0^T f(s)\,dB_s$ is a Gaussian r.v. More precisely (see also Remark 7.3).

$$\begin{aligned} \mathrm{E}[e_f(T)^2] &= \mathrm{E}\Big[\exp\Big(2\int_0^T f(s)\,dB_s\Big)\Big]\exp\Big(-\int_0^T |f(s)|^2\,ds\Big)\\ &= \exp\Big(2\int_0^T |f(s)|^2\,ds\Big)\exp\Big(-\int_0^T |f(s)|^2\,ds\Big) = \exp\Big(\int_0^T |f(s)|^2\,ds\Big)\,. \end{aligned}$$

Therefore

$$\begin{aligned} \int_0^T \mathrm{E}[e_f(s)^2]\,|f(s)|^2\,ds &\le \mathrm{E}[e_f(T)^2]\int_0^T |f(s)|^2\,ds\\ &\le \exp\Big(\int_0^T |f(s)|^2\,ds\Big)\int_0^T |f(s)|^2\,ds\,, \end{aligned}$$

which proves that the process $H_f(s) = f(s)e_f(s)$ belongs to $M^2([0,T])$ hence, thanks to (12.17), the r.v.'s of the form $e_f(T)$ belong to $\mathscr{H}$. Let us prove that they form a *total* set, i.e. that their linear combinations are dense in $\mathscr{H}$. To this end let $Y \in L^2(\overline{\mathscr{G}}_T)$ be a r.v. that is orthogonal to each of the r.v.'s $e_f(T)$ and let us prove that $Y = 0$ a.s.

First, choosing $f = 0$ so that $e_f(T) = 1$ a.s., $Y = Y^+ - Y^-$ must have mean zero and therefore $\mathrm{E}[Y^+] = \mathrm{E}[Y^-]$. If both these mathematical expectations are equal to 0, $Y$ vanishes and there is nothing to prove. Otherwise we can multiply $Y$ by a constant so that $\mathrm{E}(Y^+) = \mathrm{E}(Y^-) = 1$. The remainder of the proof consists in checking that the two probabilities $Y^+d\mathrm{P}$ and $Y^-d\mathrm{P}$ coincide on $(\Omega,\overline{\mathscr{G}}_T)$, which will imply $Y^+ = Y^-$ and therefore $Y = 0$ a.s.

If $f(s) = \sum_{j=1}^n \lambda_j 1_{[t_{j-1},t_j[}(s)$, where $0 = t_0 < t_1 < \dots < t_n = T$ and $\lambda_1, \dots, \lambda_n \in \mathbb{R}^m$, then $\int_0^T f(s)\, dB_s = \sum_{j=1}^n \langle \lambda_j, B_{t_j} - B_{t_{j-1}} \rangle$. We have

$$e_f(T) = \exp\Big( \sum_{j=1}^n \langle \lambda_j, B_{t_j} - B_{t_{j-1}} \rangle - \frac{1}{2} \sum_{j=1}^n |\lambda_j|^2 (t_j - t_{j-1}) \Big)$$

and orthogonality of $Y$ with respect to $e_f(T)$ implies

$$0 = \mathrm{E}\Big[ Y \exp\Big( \sum_{j=1}^n \langle \lambda_j, B_{t_j} - B_{t_{j-1}} \rangle \Big) \Big] , \tag{12.18}$$

i.e.

$$\mathrm{E}\Big[ Y^+ \exp\Big( \sum_{j=1}^n \langle \lambda_j, B_{t_j} - B_{t_{j-1}} \rangle \Big) \Big] = \mathrm{E}\Big[ Y^- \exp\Big( \sum_{j=1}^n \langle \lambda_j, B_{t_j} - B_{t_{j-1}} \rangle \Big) \Big] .$$

Allowing the vectors $\lambda_1, \dots, \lambda_n$ to take all possible values, this implies that the laws of the random vector $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$ with respect to the two probabilities $Y^+ d\mathrm{P}$ and $Y^- d\mathrm{P}$ have the same Laplace transforms and therefore coincide (see Sect. 5.7 for more details). Recalling the definition of the law of a r.v., this also implies that

$$\mathrm{E}\big[ Y^+ \Phi(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}) \big] = \mathrm{E}\big[ Y^- \Phi(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}) \big]$$

for every bounded measurable function $\Phi$, from which we deduce that $Y^+ d\mathrm{P}$ and $Y^- d\mathrm{P}$ coincide on the $\sigma$-algebra $\sigma(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$, which is equal to $\sigma(B_{t_1}, \dots, B_{t_n})$.

As the union of these $\sigma$-algebras for all possible choices of $n$, and of $0 = t_0 < t_1 < \dots < t_n = T$, forms a family that is stable with respect to finite intersections and generates $\mathscr{G}_T$, the two probabilities $Y^+ d\mathrm{P}$ and $Y^- d\mathrm{P}$ coincide on $\mathscr{G}_T$ by Carathéodory's criterion, Theorem 1.1. They also coincide on $\overline{\mathscr{G}}_T$ (just repeat the argument of Remark 4.4): let $\mathscr{C}$ be the class of the events of the form $A \cap N$ with $A \in \mathscr{G}_T$ and $N \in \mathscr{F}$ is either a negligible set or $N = \Omega$. Then $\mathscr{C}$ is stable with respect to finite intersections, contains both $\mathscr{G}_T$ and the negligible sets of $\mathscr{F}$ and therefore generates $\overline{\mathscr{G}}_T$. Moreover, the two probabilities $Y^+ d\mathrm{P}$ and $Y^- d\mathrm{P}$ coincide on $\mathscr{C}$ and therefore also on $\overline{\mathscr{G}}_T$, again by Carathéodory's criterion.

□

An immediate consequence is the following

**Theorem 12.5** Let $(M_t)_{0 \le t \le T}$ be a square integrable martingale of the filtration $(\overline{\mathscr{G}}_t)_t$. Then there exist a unique process $H \in M^2([0, T])$ and a constant

(continued)

**Theorem 12.5** (continued)
$c \in \mathbb{R}$ such that

$$M_t = c + \int_0^t H_s\, dB_s \qquad \text{a.s.}$$

for $t \in [0, T]$. In particular, $(M_t)_{0 \le t \le T}$ has a continuous modification.

*Proof* As $M_T \in L^2$, by Theorem 12.4 there exists a unique process $H \in M^2([0, T])$ such that

$$M_T = c + \int_0^T H_s\, dB_s$$

and therefore

$$M_t = \mathrm{E}(M_T | \overline{\mathscr{G}}_t) = c + \int_0^t H_s\, dB_s \qquad \text{a.s.}$$

□

Note that in the statement of Theorem 12.5 we make no assumption of continuity. Therefore every square integrable martingale of the filtration $(\overline{\mathscr{G}}_t)_t$ always admits a continuous version.

The representation Theorem 12.5 can be extended to local martingales.

**Theorem 12.6** Let $(M_t)_{0 \le t \le T}$ be a local martingale of the filtration $(\overline{\mathscr{G}}_t)_t$. Then there exists a unique process $H \in M^2_{loc}([0, T])$ and a constant $c \in \mathbb{R}$ such that, for $t \in [0, T]$,

$$M_t = c + \int_0^t H_s\, dB_s \qquad \text{a.s.}$$

*Proof* The idea of the proof is to approximate $M$ with square integrable martingales, but in order to do this properly we first need to prove that $(M_t)_{0 \le t \le T}$ is itself continuous, or, to be precise, that it has a continuous modification. Let us first assume that $(M_t)_{0 \le t \le T}$ is a martingale of the filtration $(\overline{\mathscr{G}}_t)_t$ (not necessarily square integrable). As $M_T$ is integrable and $L^2(\overline{\mathscr{G}}_T)$ is dense in $L^1(\overline{\mathscr{G}}_T)$, let $(Z_n)_n$ be a sequence of r.v.'s in $L^2(\overline{\mathscr{G}}_T)$ such that

$$\|Z_n - M_T\|_1 \le 2^{-n} .$$

Let $M_n(t) = \mathrm{E}(Z_n|\overline{\mathscr{G}}_t)$; $M_n$ is a square integrable martingale for every $n$ and has a continuous modification by Theorem 12.5. $(M_t)_{0\le t\le T}$ has a right continuous modification thanks to Theorem 5.14. Then the supermartingale $(-|M_t - M_n(t)|)_t$ is right-continuous and we can apply the maximal inequality (5.16), which gives

$$\begin{aligned} \mathrm{P}\Big(\sup_{0\le t\le T} |M_t - M_n(t)| \ge \frac{1}{k}\Big) &= \mathrm{P}\Big(\inf_{0\le t\le T} -|M_t - M_n(t)| \le -\frac{1}{k}\Big) \\ &\le k\mathrm{E}(|M_T - Z_n|) \le k2^{-n} \,. \end{aligned}$$

As we have on the right-hand side the general term of a convergent series, by the Borel–Cantelli lemma,

$$\sup_{0\le t\le T} |M_t - M_n(t)| < \frac{1}{k}$$

eventually a.s. In other words, $(M_n(t))_t$ converges a.s. uniformly to $(M_t)_t$, which is therefore a.s. continuous.

If $M$ is a local martingale of the filtration $(\overline{\mathscr{G}}_t)_t$ and $(\tau_n)_n$ is a sequence of reducing stopping times, then the stopped process $M^{\tau_n}$ is a martingale. Therefore, $M$ is continuous for $t < \tau_n$ for every $n$ and, as $\lim_{n\to\infty} \tau_n = +\infty$, $M$ is continuous.

The fact that $M$ is continuous allows us to apply the argument of Remark 7.6 and we can assume that the sequence $(\tau_n)_n$ is such that the stopped processes $M^{\tau_n}$ are bounded martingales, and therefore square integrable. By Theorem 12.5, for every $n$ there exist $c_n \in \mathbb{R}$ and a process $H^{(n)} \in M^2([0,T])$ such that

$$M_t^{\tau_n} = M_{t\wedge\tau_n} = c_n + \int_0^t H_s^{(n)}\, dB_s \qquad \text{a.s.}$$

Obviously $c_n = c_{n+1} = M_0 \stackrel{def}{=} c$. Moreover, the two processes $M^{\tau_n}$ and $M^{\tau_{n+1}}$ coincide for $t \le \tau_n$, therefore

$$\int_0^{\tau_n} H_s^{(n)}\, dB_s = \int_0^{\tau_n} H_s^{(n+1)}\, dB_s \qquad \text{a.s.}$$

By the uniqueness of the representation of the r.v.'s of $L^2(\overline{\mathscr{G}}_T)$ of Theorem 12.4, it follows that the two processes $(H^{(n)}_{t\wedge\tau_n})_t$ and $(H^{(n+1)}_{t\wedge\tau_n})_t$ are modifications of one another. This allows us to define a process $H$ by setting $H_t = H_t^{(n)}$ on $\{t < \tau_n\}$, as this definition does not depend on the choice of $n$. The process thus defined is such that

$$\mathrm{E}\Big(\int_0^{\tau_n} H_s^2\, ds\Big) < +\infty$$

for every $n$; as $\tau_n \to +\infty$ as $n \to \infty$, this implies $\int_0^T H_s^2\,ds < +\infty$ a.s., and therefore $H \in M^2_{loc}([0,T])$. We still have to prove that

$$M_t = c + \int_0^t H_s\,dB_s \qquad \text{a.s.,}$$

but

$$M_{t\wedge\tau_n} = c + \int_0^{t\wedge\tau_n} H_s^{(n)}\,dB_s = c + \int_0^{t\wedge\tau_n} H_s\,dB_s \qquad \text{a.s.}$$

and we can take the limit as $n \to \infty$.

□

## 12.4 Equivalent probability measures

Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion. Girsanov's Theorem 12.1 states that if $(Z_t)_t$ is a martingale of the form

$$Z_t = \exp\Big[\int_0^t \Phi_s\,dB_s - \frac{1}{2}\int_0^t |\Phi_s|^2\,ds\Big],$$

where $(\Phi_t)_t$ is an $\mathbb{R}^m$-valued process in $M^2_{loc}([0,T])$, then $d\mathrm{Q} = Z_T\,d\mathrm{P}$ is a probability on $(\Omega, \mathscr{F}_T)$ that is equivalent to P.

Let us show now that if $\mathscr{F}_t = \overline{\mathscr{G}}_t$, then all the probabilities on $(\Omega, \mathscr{F}_T)$ that are equivalent to P are of this form.

**Theorem 12.7** Let $B = (\Omega, \mathscr{F}, (\overline{\mathscr{G}}_t)_t, (B_t)_t, \mathrm{P})$ be an $m$-dimensional Brownian motion with respect to its natural augmented filtration $(\overline{\mathscr{G}}_t)_t$. If a probability Q on $(\Omega, \overline{\mathscr{G}}_T)$ is equivalent to P then there exists a progressively measurable $m$-dimensional process $\Phi \in M^2_{loc}([0,T])$ such that $d\mathrm{Q} = Z_T\,d\mathrm{P}$, where

$$Z_t = \exp\Big[\int_0^t \Phi_s\,dB_s - \frac{1}{2}\int_0^t |\Phi_s|^2\,ds\Big], \tag{12.19}$$

is a martingale under P. Moreover, $\widetilde{B}_t = B_t - \int_0^t \Phi_s\,ds$ is a $\overline{\mathscr{G}}_t$-Brownian motion under Q.

*Proof* Let

$$Z_t = \frac{dQ_{|\overline{\mathscr{G}}_t}}{dP_{|\overline{\mathscr{G}}_t}} \cdot$$

We know that $(Z_t)_{0 \le t \le T}$ is a martingale. Obviously we have, for every $t \le T$,

$$Q(Z_t > 0) = E^P(Z_t 1_{\{Z_t > 0\}}) = E^P(Z_t) = 1 \ .$$

Therefore $Z_t > 0$ Q-a.s. and also P-a.s., as P and Q are assumed to be equivalent. By Theorem 12.6 there exists a process $(\Psi_t)_t \in M^2_{loc}([0, T])$ such that

$$Z_t = 1 + \int_0^t \Psi_s \, dB_s \ .$$

We want now to apply Ito's formula in order to compute $d \log Z_t$. This is not possible directly, as log is not a $C^2$ function on $\mathbb{R}$. Let $f$ be a function such that $f(x) = \log x$ for $x \ge \frac{1}{n}$ and then extended to $\mathbb{R}$ so that it is of class $C^2$. Then, by Ito's formula,

$$df(Z_t) = f'(Z_t) \, dZ_t + \frac{1}{2} f''(Z_t) |\Psi_t|^2 \, dt \ .$$

The derivatives of $f$ coincide with those of $\log x$ for $x \ge \frac{1}{n}$, i.e. $f'(x) = \frac{1}{x}$, $f''(x) = -\frac{1}{x^2}$, so that, if $\tau_n = \inf\{t; Z_t \le \frac{1}{n}\}$,

$$\begin{aligned} \log Z_{t \wedge \tau_n} &= \int_0^{t \wedge \tau_n} f'(Z_s) \Psi_s \, dB_s + \frac{1}{2} \int_0^{t \wedge \tau_n} f''(Z_s) |\psi_s|^2 \, ds \\ &= \int_0^{t \wedge \tau_n} \frac{\Psi_s}{Z_s} \, dB_s - \frac{1}{2} \int_0^{t \wedge \tau_n} \frac{|\Psi_s|^2}{Z_s^2} \, ds \ . \end{aligned}$$

By Exercise 5.8 b), $Z$ being a martingale, we have $P(Z_t > 0$ for every $t \le T) = 1$ for every $t \le T$ and therefore $t \mapsto Z_t(\omega)$ never vanishes a.s. Therefore, $Z$ being continuous, for every $\omega \in \Omega$ there exists an $\varepsilon > 0$ such that $Z_t(\omega) \ge \varepsilon$ for every $t \le T$. Therefore $\Phi_s = \frac{\Psi_s}{Z_s} \in M^2_{loc}([0, T])$ and

$$Z_{t \wedge \tau_n} = \exp\Big( \int_0^{t \wedge \tau_n} \Phi_s \, dB_s - \frac{1}{2} \int_0^{t \wedge \tau_n} \Phi_s^2 \, ds \Big) \ .$$

Now just let $n \to \infty$ and observe that, again as $Z_t > 0$ for every $t$ a.s., $\tau_n \to +\infty$. □

## Exercises

**12.1** (p. 603) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a Brownian motion. Let us consider the three processes

$$(B_t)_t, \qquad Y_t = B_t + ct, \qquad Z_t = \sigma B_t \ ,$$

where $c, \sigma$ are real numbers. On the canonical space $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t)$ let us consider the probabilities $\mathrm{P}^B, \mathrm{P}^Y, \mathrm{P}^Z$, respectively the law of $B$ (i.e. Wiener measure) and the laws of the processes $Y$ and $Z$. Then

a) $\mathrm{P}^B$ and $\mathrm{P}^Y$ are equivalent ($\mathrm{P}^B \ll \mathrm{P}^Y$ and $\mathrm{P}^Y \ll \mathrm{P}^B$) on $\mathscr{M}_t$ for every $t \geq 0$, but, unless $c = 0$, not on $\mathscr{M}$ where they are actually orthogonal.
b) If $|\sigma| \neq 1$ then $\mathrm{P}^B$ and $\mathrm{P}^Z$ are orthogonal on $\mathscr{M}_t$ for every $t > 0$.

a) Use Girsanov's theorem in order to find a probability Q of the form $d\mathrm{Q} = Z\,d\mathrm{P}$ with respect to which $(B_t)_t$ has the same law as $(Y_t)_t$. b) Look for an event having probability 1 for $\mathrm{P}^B$ and probability 0 for $\mathrm{P}^Z$.

**12.2** (p. 604) Given two probabilities $\mu, \nu$ on a measurable space $(E, \mathscr{E})$, the *entropy* $H(\nu;\mu)$ of $\nu$ with respect to $\mu$ is the quantity

$$H(\nu;\mu) = \begin{cases} \displaystyle\int_E \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu}\, d\mu & \text{if } \nu \ll \mu \\ +\infty & \text{otherwise} . \end{cases}$$

Note that we also have, for $\nu \ll \mu$,

$$H(\nu;\mu) = \int_E \log \frac{d\nu}{d\mu}\, d\nu \ . \tag{12.20}$$

The *chi-square discrepancy* of $\nu$ with respect to $\mu$ is the quantity

$$\chi^2(\nu;\mu) = \begin{cases} \displaystyle\int_E \Big|\frac{d\nu}{d\mu} - 1\Big|^2 d\mu & \text{if } \nu \ll \mu \\ +\infty & \text{otherwise} . \end{cases}$$

a) Show that $H(\nu;\mu) \geq 0$ and $H(\nu;\mu) = 0$ if and only if $\mu = \nu$. Show, with an example, that it is possible that $H(\nu;\mu) \neq H(\mu;\nu)$.
b) Let $\mathscr{C} = \mathscr{C}([0,T], \mathbb{R}^m)$ and let $(\mathscr{C}, \mathscr{M}_T, (\mathscr{F}_t)_{0\leq t\leq T}, (X_t)_{0\leq t\leq T}, \mathrm{P})$ be the canonical space endowed with the Wiener measure P. Let $\gamma$ be a path almost a.e. differentiable and with square integrable derivative and let $\mathrm{P}_1$ be the law of $W_t = X_t + \gamma_t$, i.e. of a Brownian motion with a deterministic drift $\gamma$.
b1) What is the value of the entropy $H(\mathrm{P}_1;\mathrm{P})$? And of $H(\mathrm{P};\mathrm{P}_1)$?
b2) Compute $\chi^2(\mathrm{P}_1;\mathrm{P})$ and $\chi^2(\mathrm{P};\mathrm{P}_1)$.

a) Use Jensen's inequality, since $x \mapsto x \log x$ is a strictly convex function. b1) Girsanov's theorem gives the density of $P_1$ with respect to P. b2) Look first for a more handy expression of the $\chi^2$ discrepancy (develop the square inside the integral).

**12.3** (p. 606) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, P)$ be a Brownian motion and let $X_t = B_t - \theta t, \theta > 0$. As $\lim_{t\to+\infty} X_t = -\infty$ (thanks, for instance, to the Iterated Logarithm Law), we know that $\sup_{t>0} X_t < +\infty$. In this exercise we use Girsanov's theorem in order to compute the law of $\sup_{t>0} X_t$. We shall find, through a different argument, the same result as in Exercise 5.17.

The idea is to compute the probability of the event $\{\sup_{t>0} X_t > R\}$ with a change of probability such that, with respect to the new probability, it has probability 1 and then to "compensate" with the density of the new probability with respect to the Wiener measure P.

a) Let $Z_t = e^{2\theta B_t - 2\theta^2 t}$. Show that $(Z_t)_t$ is an $(\mathscr{F}_t)_t$-martingale and that if, for a fixed $T > 0$,

$$dQ = Z_T \, dP$$

then Q is a probability and, if $\widetilde{B}_t = B_t - 2\theta t$, then $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\le t\le T}, (\widetilde{B}_t)_{0\le t\le T}, Q)$ is a Brownian motion.

b) Show that $(Z_t^{-1})_t$ is a Q-martingale. Note that $Z_t^{-1} = e^{-2\theta X_t}$.

c) Let $R > 0$ and $\tau_R = \inf\{t; X_t = R\}$. Show that $P(\tau_R \le T) = E^Q(1_{\{\tau_R < T\}} Z_{T\wedge\tau_R}^{-1})$ and

$$P(\tau_R \le T) = e^{-2\theta R} Q(\tau_R \le T) \,. \tag{12.21}$$

d) Show that $P(\tau_R < +\infty) = e^{-2\theta R}$, hence the r.v. $\sup_{t>0} X_t$ has an exponential law with parameter $2\theta$.

c) $\{\tau_R \le T\} \in \mathscr{F}_{T\wedge\tau_R}$ and we can apply the stopping theorem to the Q-martingale $(Z_t^{-1})_t$.

**12.4** (p. 607) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, P)$ be a standard Brownian motion. In Exercise 8.14 it is proved that

$$Z_t = \frac{1}{\sqrt{1-t}} \exp\Big(-\frac{B_t^2}{2(1-t)}\Big) = \exp\Big(-\int_0^t \frac{B_s}{1-s}\, dB_s - \frac{1}{2}\int_0^t \frac{B_s^2}{(1-s)^2}\, ds\Big)$$

is a martingale for $t \in [0, T]$ for every $T < 1$. Let Q be a new probability on $(\Omega, \mathscr{F})$ defined as $dQ = Z_T \, dP$. Show that, with respect to Q, $(B_t)_t$ is a process that we have already seen many times.

**12.5** (p. 607) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, P)$ be an $m$-dimensional Brownian motion. The aim of this exercise is to compute

$$J_\lambda = E\Big[\exp\Big(\lambda \int_0^t |X_s|^2\, ds\Big)\Big] \,.$$

It may help to first have a look at Exercise 1.12 and, for d), at the properties of the Laplace transform in Sect. 5.7.

a) Let

$$Z_t = \exp\Big(\theta \int_0^t X_s\, dX_s - \frac{\theta^2}{2} \int_0^t |X_s|^2\, ds\Big).$$

Show that $\mathrm{E}(Z_t) = 1$ for every $t \ge 0$.

b) Let Q be the probability on $(\Omega, \mathscr{F}_T)$ defined as $d\mathrm{Q} = Z_T\, d\mathrm{P}$. Prove that, with respect to Q, on the time interval $[0, T]$ $X$ is an Ornstein–Uhlenbeck process (see Sect. 9.2). What is the mean and the covariance matrix of $X_t$, $0 \le t \le T$, with respect to Q?

c) Prove that, if $\lambda = -\frac{\theta^2}{2}$ and $t \le T$,

$$J_\lambda = \mathrm{E}^{\mathrm{Q}}\big[\mathrm{e}^{-\frac{\theta}{2}(|X_t|^2 - mt)}\big]$$

and determine the value of $J_\lambda$ for $\lambda \le 0$.

d) Compute $J_\lambda$ for every $\lambda \in \mathbb{R}$.

**12.6** (p. 609) Let $X = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (X_t)_t, \mathrm{P})$ be a real Brownian motion.

a1) Let $b : \mathbb{R} \to \mathbb{R}$ be a bounded continuously differentiable function and $x$ a fixed real number. Determine a new probability Q on $(\Omega, \mathscr{F})$ such that, with respect to Q, the process $B_t = X_t - \int_0^t b(X_s + x)\, ds$ is a Brownian motion for $t \le T$. Prove that, with respect to Q, the process $Y_t = x + X_t$ is the solution of an SDE to be determined.

a2) Let $U$ be a primitive of $b$. Prove that $d\mathrm{Q} = Z_T\, d\mathrm{P}$ with

$$Z_t = \exp\Big(U(X_t + x) - U(x) - \frac{1}{2} \int_0^t [b'(X_s + x) + b^2(X_s + x)]\, ds\Big)\,. \tag{12.22}$$

b1) Let $b(z) = k \tanh(kz + c)$ for some constant $k$. Prove that $b'(z) + b^2(z) \equiv k^2$.

b2) Let $Y$ be the solution of

$$\begin{aligned} dY_t &= k \tanh(kY_t + c)\, dt + dB_t \\ Y_0 &= x\,. \end{aligned} \tag{12.23}$$

Compute the Laplace transform of $Y_t$ and show that the law of $Y_t$ is a mixture of Gaussian laws, i.e. of the form $\alpha\mu_1 + (1-\alpha)\mu_2$, where $0 < \alpha < 1$ and $\mu_1, \mu_2$ are Gaussian laws to be determined.

b3) Compute $\mathrm{E}[Y_t]$.

b2) A primitive of $z \mapsto \tanh z$ is $z \mapsto \log\cosh z$.

**12.7** (p. 611) (Wiener measure gives positive mass to every open set of $\mathscr{C}_0$) Let as usual $\mathscr{C} = \mathscr{C}([0,T],\mathbb{R}^m)$ endowed with the topology of uniform convergence and let $\mathrm{P}^W$ be the Wiener measure on $\mathscr{C}$. Let us denote by $\mathscr{C}_0$ the closed subspace of $\mathscr{C}$ formed by the paths $\gamma$ such that $\gamma_0 = 0$.

a) Show that $\mathrm{P}^W(\mathscr{C}_0) = 1$.
b) Recall that, for a real Brownian motion, if $\tau$ denotes the exit time from $[-r,r]$, $r > 0$, then the event $\{\tau > T\}$ has positive probability for every $T > 0$ (Exercise 10.5 c)). Deduce that $\mathrm{P}^W(A) > 0$ for every open set $A \subset \mathscr{C}$ containing the path $\gamma \equiv 0$.
c) Note that the paths of the form $\gamma_t = \int_0^t \Phi_s\,ds$, with $\Phi \in L^2(\mathbb{R}^+,\mathbb{R}^m)$, are dense in $\mathscr{C}_0$ (they form a subset of the paths that are twice differentiable that are dense themselves). Deduce that $\mathrm{P}^W(A) > 0$ for every open set $A \subset \mathscr{C}_0$.

b) A neighborhood of the path $\gamma \equiv 0$ is of the form $V = \{w; \sup_{0\le t\le T}|w_t| < \eta\}$ for some $T,\eta > 0$. c) Use Girsanov's formula to "translate" the open set $A$ to be a neighborhood of the origin.

**12.8** (p. 612) In this exercise we use the Feynman–Kac formula in order to find explicitly the solution of the problem

$$\begin{cases} \dfrac{1}{2}\triangle u(x,t) + \dfrac{\partial u}{\partial t}(x,t) - \lambda |x|^2 u(x,t) = 0 & \text{if } (x,t) \in \mathbb{R}^m \times [0,T[ \\ u(x,T) = 1\,, \end{cases} \tag{12.24}$$

where $\lambda \ge 0$. Let $X = (\Omega,\mathscr{F},(\mathscr{F}_t)_t,(B_t)_t,\mathrm{P})$ be a $m$-dimensional Brownian motion. By Theorem 10.6, a solution of (12.24) is given by

$$u(x,t) = \mathrm{E}^{x,t}\Big[\mathrm{e}^{-\lambda\int_t^T |X_s|^2\,ds}\Big]\,,$$

where $(\Omega,\mathscr{F},(\mathscr{F}_t)_t,(X_t)_t,(\mathrm{P}^{x,t})_{x,t})$ is the canonical diffusion associated to the operator $L = \frac{1}{2}\triangle$. We know that, with respect to $\mathrm{P}^{x,t}$, the canonical process $(X_s)_{s\ge t}$ has the same law as $(B_{s-t}+x)_t$, where $(B_t)_t$ is a Brownian motion. Hence for $x = 0$, $t = 0$ we shall recover the result of Exercise 12.5.

a) For $x \in \mathbb{R}^m$ and $\theta \in \mathbb{R}$, let

$$Z_t = \exp\Big(\theta\int_0^t (B_s + x)\,dB_s - \frac{\theta^2}{2}\int_0^t |B_s + x|^2\,ds\Big)\,.$$

Show that $\mathrm{E}(Z_T) = 1$.
b) Let Q be the probability on $(\Omega,\mathscr{F}_T)$ defined as $d\mathrm{Q} = Z_T\,d\mathrm{P}$. Prove that, with respect to Q, $Y_s = B_s + x$ is an Ornstein–Uhlenbeck process (see Sect. 9.2). What is the mean and the covariance matrix of $B_T + x$ with respect to Q?

c) Prove that, under P,

$$\int_0^T (B_s + x)\, dB_s = \frac{1}{2}\,(|B_T + x|^2 - |x|^2 - mT) \tag{12.25}$$

and deduce that, for every $\theta \geq 0$,

$$\mathrm{E}\Big[\exp\Big(-\frac{\theta^2}{2}\int_0^T |B_s + x|^2\, ds\Big)\Big] = \mathrm{e}^{\frac{\theta}{2}\,(mT+|x|^2)}\mathrm{E}^{\mathrm{Q}}\Big[\mathrm{e}^{-\frac{\theta}{2}\,|B_T+x|^2}\Big]\,.$$

d) Prove (or take as granted) that, if $W$ is an $m$-dimensional $N(b, \sigma^2 I)$-distributed r.v., then

$$\mathrm{E}(\mathrm{e}^{\theta|W|^2}) = (1 - 2\sigma^2\theta)^{-m/2} \exp\Big(\frac{\theta}{1 - 2\sigma^2\theta}|b|^2\Big) \tag{12.26}$$

for every $\theta < (2\sigma^2)^{-1}$.

e) What is the value of $\mathrm{E}\big[\exp\big(-\frac{\theta^2}{2}\int_0^T |B_s + x|^2\, ds\big)\big]$?

f) Derive from b) and e) that a solution of (12.24) is

$$u(x, t) = \cosh\big(\sqrt{2\lambda}\,(T - t)\big)^{-m/2} \exp\Big[-\frac{\sqrt{2\lambda}\,|x|^2}{2}\tanh\big(\sqrt{2\lambda}\,(T - t)\big)\Big]\,.$$

Is it unique?

a) Use the criterion of Proposition 12.2 b) and Exercise 1.12.

**12.9** (p. 615) Let $B$ be an $m$-dimensional Brownian motion, $T > 0$ and $f : \mathbb{R}^m \to \mathbb{R}$ a bounded measurable function. We know (Theorem 12.4) that there exists a process $X \in M^2$ such that

$$f(B_T) = \mathrm{E}[f(B_T)] + \int_0^T X_s\, dB_s\,. \tag{12.27}$$

In this exercise we determine $X$ explicitly (Theorem 12.4 is not necessary).

a) Determine a function $\psi : [0, T[\times\mathbb{R}^m \to \mathbb{R}$ such that

$$\mathrm{E}[f(B_T)\,|\,\mathscr{F}_t] = \psi(B_t, t)$$

and show that $\psi$ is actually differentiable infinitely many times on $\mathbb{R}^m \times [0, T[$. What is the value of $\psi(0, 0)$?

b) Write the stochastic differential of $Z_t = \psi(B_t, t)$. Because of what possible reason can you state that

$$\frac{\partial\psi}{\partial t} + \frac{1}{2}\,\triangle\psi = 0$$

without actually computing the derivatives? Prove that, for every $t < T$,

$$\psi(B_t, t) = \psi(0,0) + \int_0^t \psi'_x(B_s, s)\, dB_s\,. \tag{12.28}$$

c1) Determine a process $X \in M^2([0,T])$ such that (12.27) holds.

c2) Determine $X$ explicitly if the dimension of $B$ is $m = 1$ and $f(x) = 1_{\{x>0\}}$.

d) Prove that if, in addition, $f$ is differentiable with bounded first derivatives, then (12.27) holds with $X_s = \mathrm{E}[f'(B_T)\,|\,\mathscr{F}_s]$.

# Chapter 13
# An Application: Finance

Stochastic processes are useful as models of random phenomena, among which a particularly interesting instance is given by the evolution of the values of financial (stocks, bonds, ...) and monetary assets listed on the Stock Exchange.

In this chapter we develop some models adapted for these situation and we discuss their application to some typical problems.

## 13.1 Stocks and options

In a stock exchange, besides the more traditional stocks, bonds and commodities, there are plenty of securities or derivative securities which are quoted and traded. A *derivative security* (also called a *contingent claim*), as opposed to a primary (stock, bond,...) security, is a security whose value depends on the prices of other assets of the market.

An *option* is an example of a derivative: a call option is a contract which guarantees to its holder the right (but not the obligation) to buy a given asset (stock, commodity, currency, ...), which we shall refer to as the *underlying asset*, at a given time $T$ (the *maturity*) and at a price $K$ fixed in advance (the *strike price*).

These types of contracts have an old history and they became increasingly practised at the end of the 60s with the institution of the Chicago Board Options Exchange. A call option is obviously intended to guarantee the holder of being able to acquire the underlying asset at a price that is not larger than $K$ (and therefore being safe from market fluctuations).

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2_13

If at maturity the underlying asset has a price greater than $K$, the holder of the option will exercise his right and will obtain the asset at the price $K$. Otherwise he will choose to drop the option and buy the asset on the market at a lower price.

To be precise, assume that the call option is written on a single asset whose price $S$ is modeled by some process $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (S_t)_t, \mathrm{P})$. Then the value $Z$ that the issuer of the option has to pay at maturity is equal to 0 if $S_T \le K$ (the option is not exercised) and to $S_T - K$ if $S_T > K$. In short it is equal to $(S_T - K)^+$. This quantity $(S_T - K)^+$ is the *payoff* of the call option.

Many other kinds of options exist. For instance *put options*, which guarantee the owner the right to sell at a maturity time $T$ a certain asset at a price not lower than $K$. In this case, the issuer of the option is bound to pay an amount $K - S_T$ if the price $S_T$ at time $T$ is smaller than the strike $K$ and 0 otherwise. The payoff of the put option is therefore $(K - S_T)^+$. Other examples are considered later (see Example 13.3 and Exercise 13.4, for instance)

A problem of interest, since the appearance of these derivatives on the market, is to evaluate the right price of an option. Actually the issuer of the option faces a risk: in the case of a call option, if at time $T$ the price of the stipulated asset turns out to be greater than the strike price $K$, he would be compelled to hand to the owner the amount $S_T - K$. How much should the issuer be paid in order to compensate the risk that he is going to face?

A second important question, also connected with the determination of the price, concerns the strategy of the issuer in order to protect himself from a loss (to "hedge").

Put and call options are examples of the so-called European options: each of these is characterized by its maturity and its payoff. One can formalize this as follows

**Definition 13.1** A *European option Z with maturity T* is a pair $(Z, T)$, where $T$ is the maturity date and $Z$, the payoff, is a non-negative $\mathscr{F}_T$-measurable r.v.

In the case of calls and puts the payoff is a function of the value of the underlying asset at the maturity $T$. More generally, an option can be a functional of the whole price process up to time $T$ and there are examples of options of interest from the financial point of view which are of this kind (see again Example 13.3 and Exercise 13.4, for instance).

There are also other kinds of options, such as *American* options, which differ from the European ones in the exercise date: they can be exercised at any instant $t \le T$. Their treatment, however, requires tools that are beyond the scope of this book.

In the next sections we develop the key arguments leading to the determination of the fair price of an option. We shall also discuss which stochastic processes might be reasonable models for the evolution of the price of the underlying asset. In the last section, Sect. 13.6, we shall go deeper into the investigation of the most classical model in mathematical finance, the Black-Scholes model, and we shall be able to derive explicit formulas for the price of the options and for some related quantities.

Constructing the models, we shall take into account the following facts.

- There are no transaction costs in buying or selling assets and no taxes.
- Short-selling is admitted, i.e. it is possible to sell assets that are not actually in possession of the vendor: investors who do not own a stock can sell shares of it and arrange with the owner at some future date to be payed an amount equal to the price at that date.
- It is possible to buy or sell fractions of a security.

These facts constitute an ideal market that does not exist in real life. Hence our models must be considered as a first approximation of the real markets. In particular, models taking into account transaction costs (as in real markets) have been developed, but they introduce additional complications and it is wiser to start with our simple model in order to clarify the main ideas.

Some financial slang:

- *holder* of an option: the individual/company that has acquired the rights of the option;
- *issuer* of an option: the individual/company that has issued the option and has acquired the obligations connected to it;
- *to be short* (of something): this is said of someone who has short sold something.

## 13.2 Trading strategies and arbitrage

Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, \mathrm{P})$ be a $d$-dimensional Brownian motion.

Throughout this chapter we assume that $\mathscr{F}_t = \overline{\mathscr{G}}_t$, i.e. that the filtration $(\mathscr{F}_t)_t$ is the natural augmented filtration of $B$ (see p. 32 and Sect. 4.5)

Actually it will be clear that this assumption is crucial in many instances.

We shall consider a market where $m + 1$ assets are present and we shall denote their prices by $S_0, S_1, \dots, S_m$. We shall, by now, assume that the prices $S_1, \dots, S_m$

are Ito processes with stochastic differentials

$$dS_i(t) = A_i(t)\,dt + \sum_{j=1}^{d} G_{ij}(t)\,dB_j(t)\ , \tag{13.1}$$

where $A_i$, $G_{ij}$ are continuous adapted processes. We shall assume, moreover, that the solution $S$ is such that $S_i(t) \geq 0$ a.s for every $t \in [0, T]$. This is, of course, a condition that must be satisfied by every good model . The process $S_0$ will have differential

$$dS_0(t) = r_t S_0(t)\,dt\ , \tag{13.2}$$

where $(r_t)_t$ is a non-negative *bounded* progressively measurable process (the *spot interest rate*). If we assume $S_0(s) = 1$ then

$$S_0(t) = \mathrm{e}^{\int_s^t r_u\,du}\ . \tag{13.3}$$

Typically $S_0$ is an investment of the type of treasury bonds or money deposited into a bank account: it is the *riskless asset* (even if its evolution is however random, as here we assume that the spot interest rate is a stochastic process).

The model (13.1), (13.2) is very general and we shall see later more explicit choices for the processes $A_i$, $G_{ij}$ and $r$.

An investor may decide to invest an amount of money by acquiring shares of the $m + 1$ assets of the market.

A *trading strategy* (or simply, a strategy) over the trading interval $[0, T]$ is a progressively measurable $(m + 1)$-dimensional process $H_t = (H_0(t), H_1(t), \dots, H_m(t))$ whose general component $H_i(t)$ stands for the number of units of the $i$-th security held by an investor at time $t$. The *portfolio* associated to the strategy $H$ is the corresponding wealth process:

$$V_t(H) = \langle H_t, S_t\rangle = \sum_{i=0}^{m} H_i(t)\,S_i(t), \quad t \in [0, T]\ . \tag{13.4}$$

The initial value of the portfolio $V_0(H)$ represents the *initial investment* of the strategy $H$.

At any moment the investor may decide to move part of his wealth from one asset to another. A particularly important type of trading strategy, from our point of view, is one in which he does not add or remove capital from the portfolio. The rigorous definition is given below.

We shall assume that the trading strategy $H_t = (H_0(t), H_1(t), \dots, H_m(t))$ satisfies the condition

$$\int_0^T |H_0(t)|\, dt + \sum_{i=1}^m \int_0^T |H_i(t)|^2\, dt < \infty \quad \text{a.s.} \tag{13.5}$$

The trading strategy $H$ is said to be *self-financing* over the time interval $[0, T]$ if it satisfies (13.5) and its associated portfolio $V_t(H)$ satisfies the relation

$$dV_t(H) = \langle H_t, dS_t \rangle = \sum_{i=0}^m H_i(t)\, dS_i(t), \quad 0 \le t \le T\,. \tag{13.6}$$

Therefore a strategy is self-financing if the variations of the associated portfolio, in a small time period, depend only on increments of the asset prices $S_0, \dots, S_m$, i.e. changes in the portfolio are due to capital gains or losses and not to increase or decrease of the invested funds.

Notice that the requirements (13.5) are technical and are necessary for the differentials appearing in (13.6) to be well defined. In fact, as the processes $t \mapsto A_i(t)$, $t \mapsto G_{ij}(t)$ and $t \mapsto r_t$ are bounded on $[0, T]$ for every $\omega$ ($A_i$ and $G_{ij}$ have continuous paths), then $H_0\, r \in M^1_{loc}([0, T])$, $H_i A_i \in M^1_{loc}([0, T])$ for every $i = 1, \dots, m$, and also $H_i G_{ij} \in M^2_{loc}([0, T])$ for every $i, j$ (see also Remark 7.1 a)). Of course (13.6) can be written as

$$dV_t(H) = H_0(t)\, r_t\, S^0_t dt + \sum_{i=1}^m H_i(t) A_i(t) dt + \sum_{i=1}^m H_i(t) \sum_{j=1}^d G_{ij}(t)\, dB_j(t)\,.$$

In particular, the portfolio $V(H)$ associated to a self-financing strategy $H$ is a Ito process. Let

$$\widetilde{S}_i(t) = \frac{S_i(t)}{S_0(t)} = \mathrm{e}^{-\int_0^t r_s\, ds} S_i(t), \quad i = 1, \dots, m$$

$$\widetilde{V}_t(H) = \frac{V_t(H)}{S_0(t)} = \mathrm{e}^{-\int_0^t r_s\, ds} V_t(H)\,.$$

Notice that, thanks to (13.4), $\widetilde{V}_t(H) = H_0(t) + \sum_{i=1}^m H_i(t) \widetilde{S}_i(t)$. We shall refer to $\widetilde{S}_t = (\widetilde{S}_1(t), \dots, \widetilde{S}_m(t))$ as the *discounted price process* and to $\widetilde{V}(H)$ as the *discounted portfolio*. Intuitively, $\widetilde{S}_i(t)$ is the amount of money that must be invested at time 0 into the riskless asset in order to have the amount $S_i(t)$ at time $t$. Note that by Ito's formula, as $t \mapsto \int_0^t r_s\,ds$ has finite variation,

$$\begin{aligned} d\widetilde{S}_i(t) &= -r_t e^{-\int_0^t r_s\,ds} S_i(t)\,dt + e^{-\int_0^t r_s\,ds}\,dS_i(t) \\ &= -r_t \widetilde{S}_i(t)\,dt + e^{-\int_0^t r_s\,ds}\,dS_i(t)\;. \end{aligned} \tag{13.7}$$

The following result expresses the property of being self-financing in terms of the discounted portfolio.

**Proposition 13.1** Let $H_t = (H_0(t), H_1(t), \dots, H_m(t))$ be a trading strategy satisfying (13.5). Then $H$ is self-financing if and only if

$$\widetilde{V}_t(H) = V_0(H) + \sum_{i=1}^m \int_0^t H_i(t)\,d\widetilde{S}_i(t), \qquad t \in [0,T]\;. \tag{13.8}$$

*Proof* Suppose that $H$ is self-financing. Then, by Ito's formula and (13.7),

$$\begin{aligned} d\widetilde{V}_t(H) = d\Big(e^{-\int_0^t r_s\,ds}\,V_t(H)\Big) &= -r_t e^{-\int_0^t r_s\,ds}\,V_t(H)\,dt + e^{-\int_0^t r_s\,ds}\,dV_t(H) \\ &= e^{-\int_0^t r_s\,ds}\Big(-r_t \sum_{i=0}^m H_i(t)\,S_i(t)\,dt + \sum_{i=0}^m H_i(t)\,dS_i(t)\Big) \end{aligned}$$

but, as $H_0(t) dS_0(t) = r_t H_0(t) S_0(t)$, the terms with index $i=0$ cancel and we have

$$\begin{aligned} \dots &= e^{-\int_0^t r_s\,ds}\Big(-r_t \sum_{i=1}^m H_i(t)\,S_i(t)\,dt + \sum_{i=1}^m H_i(t)\,dS_i(t)\Big) \\ &= \sum_{i=1}^m H_i(t)\,\Big(\underbrace{-r_t e^{-\int_0^t r_s\,ds} S_i(t)\,dt + e^{-\int_0^t r_s\,ds}\,dS_i(t)}_{=d\widetilde{S}_i(t)}\Big) \end{aligned}$$

and, as $\widetilde{V}_0(H) = V_0(H)$, (13.8) holds. Conversely, if $\widetilde{V}(H)$ satisfies (13.8), again by Ito's formula and (13.7)

$$\begin{aligned} dV_t(H) = d\Big(e^{\int_0^t r_s\,ds}\widetilde{V}_t(H)\Big) &= r_t e^{\int_0^t r_s\,ds}\widetilde{V}_t(H)\,dt + e^{\int_0^t r_s\,ds}\sum_{i=1}^m H_i(t)\,d\widetilde{S}_i(t) \\ &= r_t V_t(H)\,dt + e^{\int_0^t r_s\,ds}\sum_{i=1}^m H_i(t)\,\Big(-r_t\widetilde{S}_i(t)\,dt + e^{-\int_0^t r_s\,ds}\,dS_i(t)\Big) \\ &= r_t V_t(H)\,dt - r_t \sum_{i=1}^m H_i(t)\,S_i(t)\,dt + \sum_{i=1}^m H_i(t)\,dS_i(t) \end{aligned}$$

and again using the relation $H_0(t)dS_0(t) = r_t H_0(t)S_0(t)$,

$$\cdots = r_t V_t(H)\,dt - r_t \underbrace{\sum_{i=0}^{m} H_i(t)\,S_i(t)}_{=V_t(H)}\,dt + \sum_{i=0}^{m} H_i(t)\,dS_i(t) = \sum_{i=0}^{m} H_i(t)\,dS_i(t)\;,$$

i.e. $H$ is self financing. □

Let us now introduce the following sub-class of self-financing strategies.

A self-financing strategy $H$ is said to be *admissible* if $V_t(H) \geq 0$ for every $t$, a.s.

To be precise, note that the processes $H_i$ can take negative values corresponding to short selling of the corresponding assets. In order for a strategy to be admissible it is required, however, that the overall wealth $V_t(H)$ of the portfolio remains $\geq 0$ for every $t$ (i.e. that the investor is solvable at all times).

A self-financing trading strategy $H$ over $[0, T]$ is said to be an *arbitrage* strategy if the associated portfolio $V_t(H)$ satisfies

$$\begin{aligned} &V_0(H) = 0\\ &V_t(H) \geq 0 \text{ a.s. for every } t \leq T\\ &\mathrm{P}(V_T(H) > 0) > 0\,. \end{aligned} \tag{13.9}$$

Roughly speaking, an arbitrage portfolio requires no initial capital (this is the assumption $V_0(H) = 0$), its value is never negative (i.e. its admissible) and can produce a gain with strictly positive probability.

As one can imagine, an arbitrage is an operation that is seldom possible. It can happen for instance that, keeping in mind the parity between euro and yen, the dollar is cheaper in Tokyo than in Milano. An operator might then buy dollars in Tokyo and sell them in Milano; he could pay the purchase in Japan with the money from the sale in Italy and make a gain with no risk and without the need of any capital. Such a situation is therefore an arbitrage as it satisfies the three conditions (13.9).

In the real market these situations do appear, but seldom and in any case during a very short time span: in the example above purchases in Tokyo and sales in Milano would quickly provoke a raise of the exchange rate in Japan and a drop in Italy, thus closing the possibility of arbitrage.

Therefore it is commonly assumed that in a reasonable model no arbitrage strategy should be possible.

**Definition 13.2** A market model is said to be arbitrage-free if every admissible strategy $H$ on $[0, T]$ with $V_0(H) = 0$ is such that $\mathrm{P}(V_T(H) > 0) = 0$.

We shall see in the next section that the arbitrage-free property is equivalent to an important mathematical property of the model.

## 13.3 Equivalent martingale measures

**Definition 13.3** A probability $\mathrm{P}^*$ on $(\Omega, \mathscr{F}_T)$ is called an *equivalent martingale measure* if $\mathrm{P}^*$ is equivalent to P on $\mathscr{F}_T$ and the discounted price processes $\widetilde{S}_1(t), \dots, \widetilde{S}_m(t)$ are $(\mathscr{F}_t)_t$-martingales under $\mathrm{P}^*$.

Equivalent martingale measures play a very important role in our analysis. Does an equivalent martingale measure exist for the model (13.1)? Is it unique? We shall investigate these questions later. In this section and in the next one we shall point out the consequences of the existence and uniqueness of an equivalent martingale measure.

Thanks to Theorem 12.7, if $\mathrm{P}^*$ is an equivalent martingale measure, there exists a progressively measurable process $\Phi \in M^2_{loc}([0, T])$ such that

$$\frac{d\mathrm{P}^*_{|\mathscr{F}_T}}{d\mathrm{P}_{|\mathscr{F}_T}} = \mathrm{e}^{\int_0^T \Phi_s\, dB_s - \frac{1}{2}\int_0^T |\Phi_s|^2\, ds}\,.$$

Therefore by Girsanov's Theorem 12.1 the process

$$B^*_t = B_t - \int_0^t \Phi_s\, ds$$

is a P*-Brownian motion and, recalling (13.1) and (13.7), under P* the discounted prices have a differential

$$
\begin{aligned}
&d\widetilde{S}_i(t)\\
&= \underbrace{\Big(-r_t\widetilde{S}_i(t) + \mathrm{e}^{-\int_0^t r_s\,ds}A_i(t) + \mathrm{e}^{-\int_0^t r_s\,ds}\sum_{j=1}^d G_{ij}(t)\Phi_j(t)\Big)}\,dt \\
&\qquad + \mathrm{e}^{-\int_0^t r_s\,ds}\sum_{j=1}^d G_{ij}(t)\,dB^*_j(t)\ .
\end{aligned}
\tag{13.10}
$$

Therefore the prices, which are supposed to be Ito processes under the "old" probability P, are also Ito processes under P*. Note also that properties of the trading strategies as being self-financed, admissible or arbitrage are preserved under the new probability.

The requirement that the components of $\widetilde{S}$ are martingales dictates that the quantity indicated by the brace in (13.10) must vanish. The following proposition is almost obvious.

**Proposition 13.2** If there exists an equivalent martingale measure P*, the discounted portfolio $\widetilde{V}(H)$ associated to any self-financing strategy $H$ is a local martingale under P*. Moreover, if $H$ is admissible on $[0, T]$, then $\widetilde{V}(H)$ is an $(\mathscr{F}_t)_t$-supermartingale under P* on $[0, T]$.

*Proof* Let $H$ be a self-financing strategy. Then by Proposition 13.1 and (13.10), under P*

$$
d\widetilde{V}_t(H) = \sum_{i=1}^m H_i(t)\,d\widetilde{S}_i(t) = \mathrm{e}^{-\int_0^t r_s\,ds}\sum_{i=1}^m H_i(t)\sum_{j=1}^d G_{ij}(t)\,dB^*_j(t)\ ,
$$

for $t \le T$. As $t \mapsto G_{ij}(t, \omega)$ is continuous and therefore bounded for every $i$ and $j$ on $[0, T]$, and $H_i \in M^2_{loc}([0, T])$ for every $i$ (recall that this condition is required in the definition of a self-financing strategy), it follows that $H_iG_{ij} \in M^2_{loc}([0, T])$ for every $i = 1, \dots, m$ and therefore $\widetilde{V}(H)$ is a local martingale on $[0, T]$.

If $H$ is admissible, then $V(H)$ is self-financing and such that $V_t(H) \ge 0$ a.s. for every $t$ under P. As P* is equivalent to P, then also $\widetilde{V}_t(H) \ge 0$ a.s. under P*. Hence under P*, $\widetilde{V}(H)$ is a positive local martingale and therefore a supermartingale (Proposition 7.5). □

Note that, thanks to (13.10), under $\mathrm{P}^*$ the (undiscounted) price process $S$ follows the SDE

$$\begin{aligned} dS_i(t) &= d\big(\mathrm{e}^{\int_0^t r_s\,ds}\widetilde{S}_i(t)\big) = r_t\mathrm{e}^{\int_0^t r_s\,ds}\widetilde{S}_i(t)\,dt + \mathrm{e}^{\int_0^t r_s\,ds}\,d\widetilde{S}_i(t) \\ &= r_tS_i(t)\,dt + \sum_{j=1}^d G_{ij}\,dB_j^*(t)\;. \end{aligned} \tag{13.11}$$

In particular, the drift $t \mapsto A_i(t)$ is replaced by $t \mapsto r_tS_i(t)$ and the evolution of the prices under $\mathrm{P}^*$ does not depend on the processes $A_i$.

The next statement explains the importance of equivalent martingale measures and their relation with arbitrage.

**Proposition 13.3** If there exists an equivalent martingale measure $\mathrm{P}^*$, the market model is arbitrage-free.

*Proof* We must prove that for every admissible strategy $H$ over $[0,T]$ such that $V_0(H) = 0$ a.s. we must have $V_T(H) = 0$ a.s. By Proposition 13.2, $\widetilde{V}_t(H)$ is a non-negative supermartingale under $\mathrm{P}^*$, hence

$$0 \le \mathrm{E}^*[\widetilde{V}_T(H)] \le \mathrm{E}^*[\widetilde{V}_0(H)] = 0.$$

Therefore, as $V_T(H) \ge 0$, necessarily $\mathrm{P}^*(\widetilde{V}_T(H) > 0) = 0$, and then, as $\mathrm{P}$ and $\mathrm{P}^*$ are equivalent, also $\mathrm{P}(V_T(H) > 0) = 0$. □

Often we shall need the discounted portfolio to be a true martingale and not just a local one. Thus, we introduce the following definition, which requires the existence of an equivalent martingale measure $\mathrm{P}^*$.

Let $\mathrm{P}^*$ be an equivalent martingale measure. For $T > 0$, $\mathscr{M}_T(\mathrm{P}^*)$ will denote the class of the admissible strategies $H$ on $[0,T]$ such that the associated discounted portfolio $\widetilde{V}(H)$ is an $(\mathscr{F}_t)_t$-$\mathrm{P}^*$-martingale on $[0,T]$.

## 13.4 Replicating strategies and market completeness

**Definition 13.4** Assume that an equivalent martingale measure $\mathrm{P}^*$ exists. We say that a European option $(Z,T)$ is attainable if there exists a strategy $H \in \mathscr{M}_T(\mathrm{P}^*)$ such that $V_T(H) = Z$. Such a strategy $H$ is said to *replicate* the option $(Z,T)$ in $\mathscr{M}_T(\mathrm{P}^*)$.

*Remark 13.1* If $H \in \mathscr{M}_T(\mathrm{P}^*)$ then the corresponding discounted portfolio $\widetilde{V}(H)$ is a martingale under $\mathrm{P}^*$, hence $\widetilde{V}_T(H)$ is $\mathrm{P}^*$-integrable and the same is true for $V_T(H) = \mathrm{e}^{\int_0^T r_s\,ds}\widetilde{V}_T(H)$, as the spot rate $r$ is assumed to be bounded.

Therefore the condition $Z \in L^1(\Omega, \mathscr{F}_T, \mathrm{P}^*)$ is necessary for the option $(Z, T)$ to be attainable.

From now on we denote by $\mathrm{E}^*$ the expectation with respect to $\mathrm{P}^*$.

**Proposition 13.4** Assume that an equivalent martingale measure $\mathrm{P}^*$ exists and let $(Z, T)$ be an attainable European option. Then for every replicating strategy $H \in \mathscr{M}_T(\mathrm{P}^*)$ the corresponding portfolio is given by

$$V_t(H) = \mathrm{E}^*\Big[\mathrm{e}^{-\int_t^T r_s\,ds} Z \,\Big|\, \mathscr{F}_t\Big]\,. \tag{13.12}$$

*Proof* The r.v. $\mathrm{e}^{-\int_t^T r_s\,ds} Z$ is integrable under $\mathrm{P}^*$, because $Z$ is integrable under $\mathrm{P}^*$ and $r \geq 0$. Moreover, for every replicating strategy $H \in \mathscr{M}_T(\mathrm{P}^*)$ for $Z$, $\widetilde{V}(H)$ is a $\mathrm{P}^*$-martingale, hence

$$\begin{aligned} V_t(H) &= \mathrm{e}^{\int_0^t r_s\,ds}\,\widetilde{V}_t(H) = \mathrm{e}^{\int_0^t r_s\,ds}\,\mathrm{E}^*\big[\widetilde{V}_T(H)\,|\,\mathscr{F}_t\big] = \mathrm{e}^{\int_0^t r_s\,ds}\,\mathrm{E}^*\big[\mathrm{e}^{-\int_0^T r_s\,ds} V_T(H)\,|\,\mathscr{F}_t\big] \\ &= \mathrm{e}^{\int_0^t r_s\,ds}\,\mathrm{E}^*\big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\,|\,\mathscr{F}_t\big] = \mathrm{E}^*\big[\mathrm{e}^{-\int_t^T r_s\,ds}\, Z\,|\,\mathscr{F}_t\big]\,. \end{aligned}$$

□

Let $V_t = \mathrm{E}^*[\mathrm{e}^{-\int_t^T r_s\,ds} Z \,\big|\, \mathscr{F}_t]$. Proposition 13.4 suggests that (13.12) should be the right price of an option at time $t$. In fact, whenever an equivalent martingale measure exists, the issuer of the option with an amount $V_t$ can start at time $t$ an admissible strategy covering the payoff at maturity.

From another point of view, if the price at time $t$ was not given by (13.12) there would be a possibility of arbitrage. If the price was fixed at a level $C_t > V_t$, then an investor could set up a portfolio by selling the option, thus acquiring the amount $C_t$. Part of this amount would be used in order to set up the replicating portfolio. He would invest the difference $C_t - V_t$ into the riskless asset. This operation does not require any capital to be engaged. At maturity the replicating portfolio would have the same value as the payoff of the option and the investor would remain with a strictly positive gain of $(C_t - V_t)\,\mathrm{e}^{\int_t^T r_s\,ds}$, which constitutes an arbitrage.

In quite a similar way an arbitrage portfolio can be produced if the price was fixed at a level $C_t < V_t$.

For $t < T$, the value $V_t = \mathrm{E}^*\big[\mathrm{e}^{-\int_t^T r_s\,ds} Z \,\big|\, \mathscr{F}_t\big]$ is the *no-arbitrage price* of the European option $Z$ at time $t$.

This definition of price obviously depends on the martingale measure $\mathrm{P}^*$. This might not be unique but the next statement asserts that, if many equivalent martingale measures exist, the respective prices and replicating portfolios coincide.

**Proposition 13.5** Suppose there exist two equivalent martingale measures $\mathrm{P}_1^*$ and $\mathrm{P}_2^*$. Let $(Z, T)$ be a European option attainable both in $\mathscr{M}_T(\mathrm{P}_1^*)$ and $\mathscr{M}_T(\mathrm{P}_2^*)$. Then the values at any time $t$ of the two replicating portfolios agree and are equal to

$$V_t = \mathrm{E}_1^*\Big[\mathrm{e}^{-\int_t^T r_s\,ds} Z \,\Big|\, \mathscr{F}_t\Big] = \mathrm{E}_2^*\Big[\mathrm{e}^{-\int_t^T r_s\,ds} Z \,\Big|\, \mathscr{F}_t\Big]\,,$$

where $\mathrm{E}_i^*$ denotes the expectation under $\mathrm{P}_i^*$, $i = 1, 2$. In particular, the no-arbitrage prices under $\mathrm{P}_1^*$ and $\mathrm{P}_2^*$ agree.

*Proof* Let $H_1$ and $H_2$ be replicating strategies for $(Z, T)$ in $\mathscr{M}_T(\mathrm{P}_1^*)$ and $\mathscr{M}_T(\mathrm{P}_2^*)$, respectively. In particular, they are both admissible and $Z$ is integrable both with respect to $\mathrm{P}_1^*$ and $\mathrm{P}_2^*$. Since $\mathrm{P}_1^*$ and $\mathrm{P}_2^*$ are both equivalent martingale measures, by Proposition 13.4 $\widetilde{V}(H_1)$ is a $\mathrm{P}_2^*$-supermartingale and $\widetilde{V}(H_2)$ is a $\mathrm{P}_1^*$-supermartingale.

Moreover, $V_T(H_1) = Z = V_T(H_2)$ and thus $\widetilde{V}_T(H_1) = \mathrm{e}^{-\int_0^T r_s\,ds} Z = \widetilde{V}_T(H_2)$ and by Proposition 13.4, $\widetilde{V}_t(H_1) = \mathrm{E}_1^*\big[\mathrm{e}^{-\int_0^T r_s\,ds} Z \,|\, \mathscr{F}_t\big]$ and $\widetilde{V}_t(H_2) = \mathrm{E}_2^*\big[\mathrm{e}^{-\int_0^T r_s\,ds} Z \,|\, \mathscr{F}_t\big]$. As $\widetilde{V}(H_1)$ is a $\mathrm{P}_2^*$ supermartingale we have $\mathrm{P}_2^*$-a.s.

$$\widetilde{V}_t(H_2) = \mathrm{E}_2^*\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z \,\Big|\, \mathscr{F}_t\Big] = \mathrm{E}_2^*[\widetilde{V}_T(H_1) \,|\, \mathscr{F}_t] \le \widetilde{V}_t(H_1)\,.$$

By interchanging the role of $\mathrm{P}_1^*$ and $\mathrm{P}_2^*$, we obtain $\widetilde{V}_t(H_2) \ge \widetilde{V}_t(H_1)$ $\mathrm{P}_1^*$-a.s. as well and hence, as $\mathrm{P}_1^*$, $\mathrm{P}_2^*$ and $\mathrm{P}$ are equivalent (so that $\mathrm{P}_1^*$-a.s. is the same as $\mathrm{P}_2^*$-a.s.), the two conditional expectations agree. □

In summary, the existence of an equivalent martingale measure $\mathrm{P}^*$ (i.e. a measure equivalent to $\mathrm{P}$ under which the discounted price processes are martingales) has two important consequences.

- Every discounted portfolio is a $\mathrm{P}^*$-local martingale.
- The market is arbitrage free, which allows us to determine the non-arbitrage price for attainable options and the existence of a replicating strategy.

One might ask whether the converse of the last statement also holds: is it true that absence of arbitrage implies existence of an equivalent martingale measure?

The answer is positive: this is the *fundamental theorem of asset pricing*. In the literature there are several results in this direction, according to the model chosen for the market: the interested reader can refer to Musiela and Rutkowski (2005) and the references therein.

We have seen that if an equivalent martingale measure exists, the no arbitrage-price is well defined for every attainable option. Therefore it would be nice if every option (at least under suitable integrability assumptions) were attainable.

Assume that there exists an equivalent martingale measure $\mathrm{P}^*$. The model is said to be *complete* if every European option $(Z, T)$ with $Z \in L^2(\Omega, \mathscr{F}_T, \mathrm{P}^*)$ is attainable in $\mathscr{M}_T(\mathrm{P}^*)$ for every equivalent martingale measure $\mathrm{P}^*$. Otherwise the market model is said to be *incomplete*.

We have the following fundamental result

**Theorem 13.1** If the model is complete then the equivalent martingale measure is unique.

*Proof* Suppose that there exist two equivalent martingale measures $\mathrm{P}_1^*$ and $\mathrm{P}_2^*$. Let $A \in \mathscr{F}_T$ and consider the option $(Z, T)$ defined as $Z = \mathrm{e}^{\int_0^T r_s\,ds}\, 1_A$. Notice that $Z$ is $\mathscr{F}_T$-measurable and, as $r$ is assumed to be bounded, $Z \in L^p(\Omega, \mathrm{P}_i^*)$ for every $p$ and $i = 1, 2$. Since the market is complete, $Z$ is attainable both in $\mathscr{M}_T(\mathrm{P}_1^*)$ and $\mathscr{M}_T(\mathrm{P}_2^*)$. Hence, by Proposition 13.5,

$$\mathrm{P}_1^*(A) = \mathrm{E}_1^*\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\Big] = \mathrm{E}_2^*\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\Big] = \mathrm{P}_2^*(A)\ .$$

As this holds for every $A \in \mathscr{F}_T$, $\mathrm{P}_1^* \equiv \mathrm{P}_2^*$ on $\mathscr{F}_T$, i.e. the equivalent martingale measure is unique. □

## 13.5 The generalized Black–Scholes models

In the previous sections we investigated general properties of the very general model (13.1), (13.2). In this section we introduce a more precise model, which is a particular case of (13.1) and (13.2), and we investigate the existence of an equivalent martingale measure and completeness.

Let us recall first some properties that a reasonable model should enjoy.

As remarked on p. 398, prices must remain positive at all times i.e., if we assume a situation where only one asset is present on the market, it will be necessary for its price $S_t$ to be $\geq 0$ a.s. for every $t \geq 0$.

Moreover, the increments of the price must always be considered in a multiplicative sense: reading in a financial newspaper that between time $s$ and time $t$ an increment of $p\%$ has taken place, this means that $\frac{S_t}{S_s} = 1 + \frac{p}{100}$. It is therefore wise to model the logarithm of the price rather than the price itself. These and other considerations lead to the suggestion, in the case $m = 1$ (i.e. of a single risky asset), of an SDE of the form

$$\begin{aligned} \frac{dS_t}{S_t} &= b(S_t, t)\, dt + \sigma(S_t, t)\, dB_t \\ S_s &= x\,. \end{aligned} \tag{13.13}$$

If $b$ and $\sigma$ are constants, this equation is of the same kind as (9.6) on p. 259 and we know that its solution is a geometric Brownian motion, i.e.

$$S_t = x\mathrm{e}^{(b-\frac{\sigma^2}{2})(t-s)+\sigma(B_t-B_s)}\,, \tag{13.14}$$

which, if the initial position $x$ is positive, is a process taking only positive values. More precisely, we shall consider a market where $m + 1$ assets are present with prices denoted $S_0, S_1, \dots, S_m$. We shall assume that $S_0$ is as in (13.2), i.e.

$$S_0(t) = \mathrm{e}^{\int_s^t r_u\, du}$$

for a non-negative bounded progressively measurable process $r$. As for the other assets we shall assume that they follow the SDE

$$\frac{dS_i(t)}{S_i(t)} = b_i(S_t, t)\, dt + \sum_{j=1}^{d} \sigma_{ij}(S_t, t)\, dB_j(t) \tag{13.15}$$

where $S_t = (S_1(t), \dots, S_m(t))$. Recall that in this model there are $m$ risky assets and that their evolution is driven by a $d$-dimensional Brownian motion, possibly with $m \neq d$.

We shall make the assumption

| |
|---|
| $b$ and $\sigma$ are *bounded* and locally Lipschitz continuous. |

With this assumption the price process $S$ is the solution of an SDE with coefficients satisfying Assumption (A) on p. 260. In particular (Theorem 9.1), $S \in M^2$.

This is the *generalized Black–Scholes* (or *Dupire*) model. In the financial models the diffusion coefficient is usually referred to as the *volatility*.

By Ito's formula applied to the function $f : z = (z_1, \dots, z_m) \mapsto (\log z_1, \dots, \log z_m)$, the process $\zeta_t = (\log S_1(t), \dots, \log S_m(t))$ solves the SDE

$$d\zeta_i(t) = \Big(b_i(\mathrm{e}^{\zeta_i(t)}, t)\, dt - \frac{1}{2}\, a_{ii}(\mathrm{e}^{\zeta_i(t)}, t)\Big)\, dt + \sum_{j=1}^{d} \sigma_{ik}(\mathrm{e}^{\zeta_i(t)}, t)\, dB_k(t) \;, \qquad (13.16)$$

where $a = \sigma\sigma^*$. This not a rigorous application of Ito's formula, as $f$ is not even defined on the whole of $\mathbb{R}$, but note that the assumptions stated for $b$ and $\sigma$ guarantee that the SDE (13.16) has a unique solution, $\zeta$, defined for every $t > 0$. It is then easy by Ito's formula, correctly applied to $f^{-1} : (y_1, \dots, y_m) \mapsto (\mathrm{e}^{y_1}, \dots, \mathrm{e}^{y_m})$, to check that $S_t = x\mathrm{e}^{\zeta_t} > 0$ is a solution of (13.13).

Hence if $S_i(s) = x_i > 0$ then $S_i(t) > 0$ for every $t \ge s$ a.s., which is a good thing, as remarked above.

Let us investigate the properties of this model starting from the existence and uniqueness of an equivalent martingale measure. The next theorem gives a characterization of the equivalent martingale measures in the generalize Black–Scholes model.

**Theorem 13.2** The following are equivalent.

a) An equivalent martingale measure $\mathrm{P}^*$ exists.
b) There exists a process $\gamma \in M^2_{loc}$ such that

$$\xi_t = \mathrm{e}^{\int_0^t \gamma_u\, dB_u - \frac{1}{2}\int_0^t |\gamma_u|^2\, du} \qquad (13.17)$$

is an $(\mathscr{F}_t)_t$-martingale and $\gamma$ is a solution of the system of $m$ equations (but $\gamma$ is $d$-dimensional)

$$\sigma(S_t, t)\gamma(t) = R_t - b(S_t, t) \;, \qquad (13.18)$$

where $R_t$ denotes the $m$-dimensional process having all its components equal to $r_t$.

*Proof* Let us assume that b) holds and let $\mathrm{P}^*$ be the probability having density $\xi_T$ with respect to P. By Girsanov's theorem, Theorem 12.1, the process

$$B^*_t = B_t - \int_0^t \gamma_u\, du \qquad (13.19)$$

is, for $t \le T$, a Brownian motion with respect to the new probability $d\mathrm{P}^* = \xi_T\, d\mathrm{P}$ and under $\mathrm{P}^*$ the discounted price process, thanks to (13.10), satisfies

$$\frac{d\widetilde{S}_i(t)}{\widetilde{S}_i(t)} = (b_i(S_i(t),t) - r_t)\, dt + \sum_{j=1}^d \sigma_{ij}(S_i(t),t)\, dB_j(t) = \sum_{j=1}^d \sigma_{ij}(S_i(t),t)\, dB_j^*(t)\;.$$

Let us prove that $\widetilde{S}_i$ is a martingale with respect to $\mathrm{P}^*$. Let us apply Ito's formula and compute the stochastic differential of $t \mapsto \log \widetilde{S}_i(t)$ or, to be precise, let $f_\varepsilon$ be a function coinciding with log on $]\varepsilon, +\infty[$ and extended to $\mathbb{R}$ so that it is twice differentiable. If $\tau_\varepsilon$ denotes the exit time of $\widetilde{S}_i$ from the half-line $]\varepsilon, +\infty[$, then Ito's formula gives

$$\begin{aligned}
&\log \widetilde{S}_i(t \wedge \tau_\varepsilon)\\
&= \log x_i + \int_s^{t\wedge\tau_\varepsilon} f_\varepsilon'(\widetilde{S}_i(u))\, d\widetilde{S}_i(u) + \frac{1}{2}\int_s^{t\wedge\tau_\varepsilon} f_\varepsilon''(\widetilde{S}_i(u))\, d\langle \widetilde{S}_i\rangle_u\\
&= \log x_i + \int_s^{t\wedge\tau_\varepsilon} \frac{1}{\widetilde{S}_i(u)}\, d\widetilde{S}_i(u) - \frac{1}{2}\int_s^{t\wedge\tau_\varepsilon} \frac{1}{\widetilde{S}_i(u)^2}\, d\langle \widetilde{S}_i\rangle_u\\
&= \log x_i + \int_s^{t\wedge\tau_\varepsilon} \sum_{j=1}^d \sigma_{ij}(S_i(u),u)\, dB_j^*(u) - \frac{1}{2}\int_0^{t\wedge\tau_\varepsilon} \sum_{j=1}^d \sigma_{ij}(S_i(u),u)^2\, du\;.
\end{aligned}\tag{13.20}$$

As we know that $S_i(t) > 0$, hence also $\widetilde{S}_i(t) > 0$, for every $t > 0$ a.s., we have $\tau_\varepsilon \to +\infty$ as $\varepsilon \to 0+$, so that we can take the limit as $\varepsilon \to 0+$ in (13.20), which gives for $\widetilde{S}_i$ the expression

$$\widetilde{S}_i(t) = x_i \exp\Big(\int_s^t \sum_{j=1}^d \sigma_{ij}(S_i(u),u)\, dB_j^*(u) - \frac{1}{2}\int_s^t \sum_{j=1}^d \sigma_{ij}(S_i(u),u)^2\, du\Big)\;.$$

As $\sigma$ is assumed to be bounded, $\widetilde{S}_i$ is a martingale, thanks to Proposition 12.1.

Conversely, if an equivalent martingale measure $\mathrm{P}^*$ exists, then by Theorem 12.7 there exists a process $\gamma \in M^2_{loc}$ such that the r.v. $\xi_T$ defined by (13.17) is the density of $\mathrm{P}^*$ with respect to $\mathrm{P}$ on $\mathscr{F}_T$. Then with respect to $\mathrm{P}^*$ the process $B^*$ defined in (13.19) is a Brownian motion up to time $T$ so that, for $0 \le t \le T$, $\widetilde{S}_i$ is a solution of

$$\frac{d\widetilde{S}_i(t)}{\widetilde{S}_i(t)} = \Big(b_i(S_i(t),t) - r_t + \sum_{j=1}^d \sigma_{ij}(S_i(t),t))\gamma_j(t)\Big)\, dt + \sum_{j=1}^d \sigma_{ij}(S_i(t),t)\, dB_j^*(t)\;.$$

As $\widetilde{S}_i$ is a martingale with respect to P*, then necessarily the coefficient of $dt$ in the previous differential vanishes, i.e.

$$b_i(S_i(t), t) - r_t + \sum_{j=1} \sigma_{ij}(S_i(t), t))\gamma_j(t) = 0$$

for every $i = 1, \dots, m$, i.e. (13.18). □

Recall that, as seen in the proof of Theorem 13.2, the processes $\widetilde{S}_i$ satisfy under P* the relation

$$\frac{d\widetilde{S}_i(t)}{\widetilde{S}_i(t)} = \sum_{j=1}^{d} \sigma_{ij}(S_t, t)\, dB_j^*(t) \tag{13.21}$$

whereas for the undiscounted price processes we have

$$\frac{dS_i(t)}{S_i(t)} = r_t\, dt + \sum_{j=1}^{d} \sigma_{ij}(S_t, t)\, dB_j^*(t)\ . \tag{13.22}$$

If the price processes $\widetilde{S}_i$ are martingales under the probability P*, the process $\gamma$ introduced in Theorem 13.2 is called *the market price of risk*.

Let us now investigate the existence of an equivalent martingale measure for the generalized Black–Scholes model.

**Proposition 13.6** Suppose $d \geq m$ (i.e. that the dimension of the driving Brownian motion is larger than or equal to the number of risky assets), that $\sigma(x, t)$ has maximum rank for every $(x, t)$ (i.e. $\sigma(x, t)$ has rank equal to $m$) and, moreover, that the matrix field $a = \sigma\sigma^*$ is uniformly elliptic with its smallest eigenvalue bounded from below by $\lambda > 0$ (Definition 9.5 and (9.41)).

Then, there exists at least an equivalent martingale measure P* and the corresponding market price of risk $\gamma$ is bounded a.s.

Moreover, if $d = m$ the process $\gamma$ is unique and is the unique solution of (13.18). Conversely, if $d > m$ then there is no uniqueness for the equivalent martingale measure.

*Proof* We know, see Remark 9.7, that the assumption of ellipticity implies that $a(x, t)$ is invertible for every $x, t$. Let us first consider the case $d = m$ and therefore that $\sigma(x, t)$ is invertible for every $x, t$. Equation (13.18) then has the unique solution

$$\gamma_t = \sigma(S_t, t)^{-1}(R_t - b(S_t, t))\ .$$

Such a process $\gamma$ is bounded. In fact

$$
\begin{aligned}
|\gamma_t|^2 &= \big\langle \sigma(S_t,t)^{-1}\big(R_t - b(S_t,t)\big), \sigma(S_t,t)^{-1}\big(R_t - b(S_t,t)\big)\big\rangle \\
&= \big\langle \underbrace{(\sigma\sigma^*(S_t,t))^{-1}}_{=a(x,t)^{-1}} \big(R_t - b(S_t,t)\big), R_t - b(S_t,t)\big\rangle \le \frac{1}{\lambda}\,|R_t - b(S_t,t)|^2 \le K\,.
\end{aligned}
\tag{13.23}
$$

The existence of an equivalent martingale measure is now a consequence of Theorem 13.2. If $d > m$ the argument is similar but we cannot argue in the same way, as now $\sigma(x,t)$ is not invertible. Actually $\sigma(x,t)$ has, for every $x, t$, a $(d-m)$-dimensional kernel. However, for every $x, t$, let $\rho(x,t)$ be an orthogonal $d \times d$ matrix such that its columns from the $(m+1)$-th to the $d$-th form an orthogonal basis of $\ker \sigma(x,t)$. Then (13.18) can be written as

$$
\underbrace{\sigma(S_t,t)\rho(S_t,t)}_{:=\overline{\sigma}(S_t,t)}\,\rho^{-1}(S_t,t)\gamma_t = R_t - b(S_t,t)\,.
\tag{13.24}
$$

The columns from the $(m+1)$-th to the $d$-th of the matrix $\overline{\sigma}(S_t,t) = \sigma(S_t,t)\rho(S_t,t)$ vanish, hence $\overline{\sigma}(x,t)$ is of the form

$$
\overline{\sigma}(x,t) = (\widetilde{\sigma}(x,t), 0_{m,d-m})\,,
$$

where $\widetilde{\sigma}(x,t)$ is an $m \times m$ matrix and $0_{m,d-m}$ denotes an $m \times (d-m)$ matrix of zeros. As

$$
a(x,t) = \sigma(x,t)\sigma^*(x,t) = \overline{\sigma}(x,t)\overline{\sigma}^*(x,t) = \widetilde{\sigma}(x,t)\widetilde{\sigma}^*(x,t)\,,
$$

$\widetilde{\sigma}(x,t)$ is invertible for every $x, t$. Let $\zeta_1$ be the $m$-dimensional process

$$
\zeta_1(t) = \widetilde{\sigma}(S_t,t)^{-1}\big(R_t - b(S_t,t)\big)
$$

and let $\zeta_2(t)$ be any bounded progressively measurable process taking values in $\mathbb{R}^{d-m}$. Then the process $\gamma_t = \begin{pmatrix}\zeta_1(t)\\ \zeta_2(t)\end{pmatrix}$ satisfies the relation

$$
\overline{\sigma}(S_t,t)\zeta_t = R_t - b(S_t,t)
$$

and $\gamma_t^* = \rho^{-1}(S_t,t)\zeta_t$ is a solution of (13.24). A repetition of the argument as in (13.23), proves that the process $\zeta_1$ is bounded, hence also $\gamma^*$ is bounded, thus proving the existence of an equivalent martingale measure.

Finally, note that if $d > m$ then there are many bounded solutions of (13.18) (for any possible choice of a bounded progressively measurable process $\zeta_2$). Hence if $d > m$ the equivalent martingale measure is not unique. □

In particular, if $d > m$, the market cannot be complete, thanks to Proposition 13.1 (see, however, Remark 13.2 and Example 13.1 below). Let us investigate completeness if $d = m$. The main point is contained in the next statement.

**Lemma 13.1** Suppose $d = m$ and that $\sigma\sigma^*$ is uniformly elliptic. Let $\mathrm{P}^*$ denote the unique equivalent martingale measure whose existence is proved in Proposition 13.6. Let $(Z, T)$ be a European option such that $Z \in L^1(\Omega, \mathscr{F}_T, \mathrm{P}^*)$ and let

$$\widetilde{M}_t = \mathrm{E}^*\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z \,\Big|\, \mathscr{F}_t\Big] .$$

Then there exist $m$ progressively measurable processes $H_1, \dots, H_m \in M^2_{loc}$ such that

$$d\widetilde{M}_t = \sum_{i=1}^m H_i(t)\, d\widetilde{S}_i(t) .$$

*Proof* The idea is to use the representation theorem of martingales, Theorem 12.6, recalling that in this chapter we assume that $(\mathscr{F}_t)_t$ is the natural augmented filtration of $B$. In fact, as $\widetilde{M}$ is a martingale of the Brownian filtration, we might expect to have the representation:

$$\widetilde{M}_t = \mathrm{E}^*\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\Big] + \int_0^t \widetilde{Y}_s\, dB^*_s, \quad t \in [0, T] , \tag{13.25}$$

where $\widetilde{Y}$ is a progressively measurable process in $M^2_{loc}$. In order to conclude the proof it would be sufficient to choose $H$ as the solution of

$$\sum_{i=1}^m H_i(t)\widetilde{S}_i(t)\sigma_{ik}(S_t, t) = \widetilde{Y}_k(t) , \quad k = 1, \dots, m, \tag{13.26}$$

i.e., $\sigma$ being invertible,

$$H_i(t) = \frac{[(\sigma^*)^{-1}(S_t, t)\widetilde{Y}_t]_i}{\widetilde{S}_i(t)} \quad i = 1, \dots, m . \tag{13.27}$$

The integrability requirement for $H$ is satisfied as $(\sigma^*)^{-1}$ is bounded (this has already been proved in the proof of Proposition 13.6, as a consequence of the fact that $a = \sigma\sigma^*$ is uniformly elliptic) and $\widetilde{S}_t$ is a strictly positive continuous process. Therefore $H_i \in M^2_{loc}$ for every $i$ because $\widetilde{Y}_k \in M^2_{loc}$ for every $k$.

However, the representation theorem for Brownian martingales cannot be applied in this setting. Indeed, as we are working under $\mathrm{P}^*$, thus with the Brownian motion $B^*$, the filtration to be taken into account in the representation theorem for Brownian martingales is not $(\mathscr{F}_t)_t$ but $(\mathscr{F}_t^*)_t$, i.e. the augmented filtration generated by $B^*$ and completed by the $\mathrm{P}^*$-null sets (and in general, $\mathscr{F}_t^* \subset \mathscr{F}_t$). In other words, the above argument would be correct if one had to work with $\mathrm{E}^*\big[\mathrm{e}^{-\int_t^T r_s\,ds} Z \,|\, \mathscr{F}_t^*\big]$, and not with $\mathrm{E}^*\big[\mathrm{e}^{-\int_t^T r_s\,ds} Z \,|\, \mathscr{F}_t\big]$.

The argument that follows is necessary in order to take care of this difficulty. Let $\gamma^*$ denote the market price of risk (see Proposition 13.6) and let

$$\xi_t = \mathrm{e}^{\int_0^t \gamma_s^*\,dB_s - \frac{1}{2}\int_0^t |\gamma_s^*|^2\,ds}$$

be the usual exponential martingale such that $\frac{d\mathrm{P}^*_{|\mathscr{F}_T}}{d\mathrm{P}_{|\mathscr{F}_T}} = \xi_T$. Let us consider the martingale

$$\overline{M}_t = \mathrm{E}\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\,\xi_T \,\Big|\, \mathscr{F}_t\Big]\,.$$

Note that the expectation is taken under the original measure P and not $\mathrm{P}^*$ and that $\mathrm{E}\big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\,\xi_T\big] = \mathrm{E}^*\big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\big]$, so that the r.v. $\mathrm{e}^{-\int_0^T r_s\,ds} Z\,\xi_T$ is P-integrable. Therefore $(\overline{M}_t)_{t\in[0,T]}$ is an $(\mathscr{F}_t)_t$-martingale under P and we can apply the representation theorem for the martingales of the Brownian filtration and obtain that there exists a progressively measurable process $Y \in M^2_{loc}$ such that

$$\overline{M}_t = \mathrm{E}\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\xi_T\Big] + \int_0^t Y_t\,dB_t\,.$$

Now, recalling the relation between conditional expectations under a change of probability (Exercise 4.4),

$$\widetilde{M}_t = \mathrm{E}^*\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z \,\Big|\, \mathscr{F}_t\Big] = \frac{\mathrm{E}\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\,\xi_T \,\Big|\, \mathscr{F}_t\Big]}{\mathrm{E}[\xi_T \,|\, \mathscr{F}_t]} = \frac{\mathrm{E}\Big[\mathrm{e}^{-\int_0^T r_s\,ds} Z\,\xi_T \,\Big|\, \mathscr{F}_t\Big]}{\xi_t} = \frac{\overline{M}_t}{\xi_t}\,\cdot$$

Let us compute the stochastic differential of $\widetilde{M}$. We have

$$d\overline{M}_t = Y_t\,dB_t = Y_t(dB_t^* + \gamma_t^*\,dt)$$

and $d\xi_t = \xi_t\,\gamma_t^*\,dB_t$, hence

$$d\frac{1}{\xi_t} = -\frac{1}{\xi_t^2}\,d\xi_t + \frac{1}{\xi_t^3}\,d\langle\xi\rangle_t = -\frac{1}{\xi_t}\,\gamma_t^*\,dB_t + \frac{1}{\xi_t}|\gamma_t^*|^2\,dt = -\frac{1}{\xi_t}\gamma_t^*\,dB_t^*$$

and with a final stroke of Ito's formula

$$d\widetilde{M}_t = d\frac{\overline{M}_t}{\xi_t} = \frac{1}{\xi_t}\,Y_t(dB^*_t + \gamma^*_t\,dt) - \frac{\overline{M}_t}{\xi_t}\,\gamma^*_t\,dB^*_t - \frac{1}{\xi_t}Y_t\gamma^*_t\,dt = \frac{1}{\xi_t}\,(Y_t - \overline{M}_t\gamma^*_t)\,dB^*_t\ .$$

This gives the representation formula (13.25) with $\widetilde{Y}_t = \xi_t^{-1}(Y_t - \overline{M}_t\gamma_t^*)$, which when inserted in (13.27) gives the process $H$ we are looking for. It remains to prove that $\widetilde{Y} \in M^2_{loc}$, which is easy and left to the reader. □

**Theorem 13.3** If $m = d$ and the matrix field $a(x,t) = \sigma\sigma^*(x,t)$ is uniformly elliptic then the generalized Black–Scholes model (13.15) is complete.

*Proof* Let $Z \in L^1(\Omega, \mathscr{F}_T, \mathrm{P}^*)$ be a positive r.v. We must prove that it is attainable. Let

$$\widetilde{M}_t = \mathrm{E}^*\Big[\mathrm{e}^{-\int_0^T r_s ds} Z \,\Big|\, \mathscr{F}_t\Big]\,, \qquad t \in [0,T]\ .$$

By Lemma 13.1 there exist $m$ progressively measurable processes $H_1, \dots, H_m \in M^2_{loc}$ such that

$$d\widetilde{M}_t = \sum_{i=1}^m H_i(t)d\widetilde{S}_i(t)\ .$$

Let

$$H_0(t) = \widetilde{M}_t - \sum_{i=1}^m H_i(t)\widetilde{S}_i(t)\,, \qquad t \in [0,T]\,, \tag{13.28}$$

and consider the trading strategy $H = (H_0, H_1, \dots, H_m)$ over $[0,T]$. Notice that for the corresponding portfolio we have

$$\widetilde{V}_t(H) = H_0(t) + \sum_{i=1}^m H_i(t)\widetilde{S}_i(t) = \widetilde{M}_t, \qquad t \in [0,T]\,, \tag{13.29}$$

hence $V_T(H) = Z$. Let us prove that $H$ is admissible.

First observe that $H$ satisfies (13.5). In fact, we already know that $H_1, \dots, H_m \in M^2_{loc}$. Moreover, $\widetilde{M}$ and $\widetilde{S}$ are continuous; this proves that the $H_iS_i$ are also in $M^2_{loc}$ and from (13.28) $H_0 \in M^2_{loc} \subset M^1_{loc}$. By (13.29)

$$d\widetilde{V}_t(H) = d\widetilde{M}_t = \sum_{i=1}^m H_i(t)d\widetilde{S}_i(t)\,,$$

so that $V(H)$ is self-financing thanks to Proposition 13.1. Moreover, $\widetilde{V}_t(H) = \widetilde{M}_t \geq 0$ a.s. by construction, being the conditional expectation of a positive r.v. Therefore $H$ is admissible.

Moreover, $H$ is a trading strategy in $\mathscr{M}_T(\mathrm{P}^*)$ as $\widetilde{V}(H)$ is a $(\mathscr{F}_t)_t$-martingale under $\mathrm{P}^*$: $\widetilde{V}_t(H) = \widetilde{M}_t$ and $\widetilde{M}_t$ is an $(\mathscr{F}_t)_t$-$\mathrm{P}^*$-martingale by construction.

As, finally, $V_T(H) = \mathrm{e}^{\int_0^T r_s ds}\widetilde{V}_T(H) = \mathrm{e}^{\int_0^T r_s ds}\widetilde{M}_T = Z$, $H$ replicates $Z$ at time $T$ and $Z$ is attainable. □

In the previous statements we have assumed $d \geq m$, i.e. that there are at least as many Brownian motions as there are risky assets. If, conversely, $d < m$ then the solution of the equation

$$\sigma(S_t, t)\gamma_t = b(S_t, t) - R_t$$

might not exist as $\sigma(x, t)$ is a matrix mapping $\mathbb{R}^d \to \mathbb{R}^m$. Hence an equivalent martingale measure might not exist and arbitrage is possible. Exercise 13.5 gives an example of this phenomenon.

*Remark 13.2* Let us criticize the theory developed so far.

The definition of a European option given in Definition 13.1 is not really reasonable. An option in the real market is something that depends on the value of the prices. The call and put options are of this kind, as well as the other examples that we shall see in the next sections.

Therefore a reasonable definition of a European option would be that it is a pair $(Z, T)$ (payoff and maturity) such that $Z$ *is* $\overline{\mathscr{H}}_T$*-measurable* where $\overline{\mathscr{H}}_t = \sigma(S_s, s \leq t, \mathscr{N})$ denotes the filtration *generated by the price process* $S$ and augmented. It is clear that $\overline{\mathscr{H}}_t \subset \mathscr{F}_t$ and in general $\overline{\mathscr{H}_t} \neq \mathscr{F}_t$ (recall that we assume $(\mathscr{F}_t)_t = (\overline{\mathscr{G}}_t)_t$, see also Remark 9.4).

Similarly it would be useful if the trading strategy $H_t$ at time $t$ were an $\overline{\mathscr{H}}_t$-measurable r.v., i.e. a functional of the price process.

If $m = d$ this turns out not to be a real problem: under the assumptions of Theorem 13.3, i.e. the matrix field $a(x, t) = \sigma\sigma^*(x, t)$ is uniformly elliptic, it can be proved that the two filtrations $(\mathscr{F}_t)_t$ and $(\overline{\mathscr{H}}_t)_t$ coincide.

This equality of the two filtrations is not true if $d > m$. In this case, however, one may ask whether a European option $(Z, T)$ such that $Z$ is $\overline{\mathscr{H}}_T$-measurable is attainable. The next example provides an approach to the question.

*Example 13.1* Let us consider a generalized Black–Scholes market model with one risky asset satisfying the SDE, with respect to the two-dimensional

(continued)

*Example 13.1* (continued)
Brownian motion $B$,

$$\frac{dS_t}{S_t} = b(S_t)\,dt + \sigma_1\,dB_1(t) + \sigma_2\,dB_2(t)\ , \tag{13.30}$$

where $\sigma_1, \sigma_2 > 0$. In this case, the dimension of the Brownian motion is $d = 2$, so that, by Proposition 13.6, there is no uniqueness of the equivalent martingale measure and by Theorem 13.1 the model is not complete. Hence there are options that are not attainable. But, denoting by $(\overline{\mathscr{H}}_t)_t$ the augmented filtration generated by $S$, we see now that every European option $(Z, T)$ such that $Z$ is $\overline{\mathscr{H}}_T$-measurable is actually attainable.

Indeed, consider the real Brownian motion

$$W_t = \frac{\sigma_1 B_1(t) + \sigma_2 B_2(t)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\cdot$$

Now (13.30) can be rewritten as

$$\frac{dS_t}{S_t} = b(S_t)\,dt + \sigma\,dW_t\ ,$$

where $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$ and Theorem 13.3 guarantees that every option that is $\overline{\mathscr{G}}_T^W$-measurable is attainable, where we denote by $(\overline{\mathscr{G}}_t^W)_t$ the augmented filtration of $W$ (which is strictly smaller than $\mathscr{F}_T$: try to find an example of an $\mathscr{F}_T$-measurable r.v. which is not $\mathscr{G}_T^W$-measurable). As $\overline{\mathscr{H}}_T \subset \overline{\mathscr{G}}_T^W$ (again thanks to Remark 9.4) our claim is verified.

Note also that in this case $\overline{\mathscr{H}}_T = \overline{\mathscr{G}}_T^W$ (see Exercise 9.26).

## 13.6 Pricing and hedging in the generalized Black–Scholes model

We have seen in Theorem 13.3 that under suitable conditions for the generalized Black–Scholes model a unique equivalent martingale measure P* exists and that the model is complete. Hence every European option $(Z, T)$ such that $Z$ is square integrable with respect to P* is attainable and by Proposition 13.4 its no-arbitrage price is given by

$$V_t = \mathrm{E}^*\Big[\mathrm{e}^{-\int_t^T r_s\,ds} Z \,\Big|\, \mathscr{F}_t\Big]\,. \tag{13.31}$$

The problem of the determination of the price is therefore solved.

But let us consider the problem also from the point of view of the issuer: which strategy should be taken into account in order to deliver the contract? The value $V_t$ of (13.31) is also the value of a replicating portfolio but it is also important to determine the corresponding strategy $H$. This would enable the issuer to construct the replicating portfolio, which, at maturity, will have the same value as the payoff of the option.

This is the *hedging* problem. In Theorem 13.1 we have proved the existence of a replicating strategy $H$, but we made use of the martingale representation theorem, which is not constructive. We make two additional assumptions to our model

- $(x, t) \mapsto \sigma(x, t)$ is Lipschitz continuous.
- The randomness of the spot rate $r$ is driven by the risky asset prices, i.e. is of the form

$$r_t = r(S_t, t)\ .$$

Note that this assumption contains the case where $r$ is a deterministic function. We shall require $(x, t) \mapsto r(x, t)$ to be a bounded, non-negative and Lipschitz continuous function.

Hence, recalling (13.22) and (13.21), the price process and the discounted price process solve respectively the SDEs

$$\begin{aligned} \frac{dS_i(t)}{S_i(t)} &= r(S_t, t)\, dt + \sum_{k=1}^{d} \sigma_{ik}(S_i(t), t)\, dB_k^*(t) \\ S_i(0) &= x_i \qquad\qquad i = 1, \dots, m \end{aligned} \tag{13.32}$$

and

$$\begin{aligned} \frac{d\widetilde{S}_i(t)}{\widetilde{S}_i(t)} &= \sum_{k=1}^{d} \sigma_{ik}(S_i(t), t)\, dB_k^*(t)) \\ \widetilde{S}_i(0) &= x_i \qquad\qquad i = 1, \dots, m\ . \end{aligned} \tag{13.33}$$

Note that with these assumptions on $\sigma$ and $r$, the diffusion process $S$ satisfies the hypotheses of most of the representation theorems of Chap. 10 and in particular of Theorem 10.6.

Let us consider an option $Z$ of the form $Z = h(S_T)$. Under $\mathrm{P}^*$, $S$ is an $(m + 1)$-dimensional diffusion and thanks to the Markov property, Proposition 6.1,

$$V_t = \mathrm{E}^*\Big[\mathrm{e}^{-\int_t^T r(S_s,s)\, ds}\, h(S_T)\,\Big|\, \mathscr{F}_t\Big] = P(S_t, t)\ ,$$

where

$$P(x,t) = \mathrm{E}^*\Big[\mathrm{e}^{-\int_t^T r(S_s^{x,t},s)\,ds}\, h(S_T^{x,t})\Big]\,,$$

$(S_s^{x,t})_{s\ge t}$ denoting the solution of (13.33) starting at $x$ at time $t$. The value of the replicating discounted portfolio is then

$$\widetilde{V}_t = \mathrm{e}^{-\int_0^t r(S_s,s)\,ds} V_t = \mathrm{e}^{-\int_0^t r(S_s,s)\,ds} P(S_t,t)\ .$$

Now suppose that the function $P$ is of class $C^{2,1}$ (continuous, twice differentiable in the variable $x$ and once in $t$). Ito's formula gives

$$\begin{aligned}
d\widetilde{V}_t &= -r(S_t,t)\mathrm{e}^{-\int_0^t r(S_s,s)\,ds} P(S_t,t)\,dt + \mathrm{e}^{-\int_0^t r(S_s,s)\,ds}\, dP(S_t,t)\\
&= \mathrm{e}^{-\int_0^t r(S_s,s)\,ds}\Big( -r(S_t,t)P(S_t,t) + \frac{\partial P}{\partial t}\,(S_t,t)\\
&\quad + \sum_{i=1}^m \frac{\partial P}{\partial x_i}\,(S_t,t) r(S_t,t) S_i(t)\\
&\quad + \frac{1}{2}\sum_{i,j=1}^m \frac{\partial^2 P}{\partial x_i \partial x_j}\,(S_t,t)\, a_{ij}(S_t,t)\, S_i(t)\, S_j(t)\Big)\,dt\\
&\quad + \mathrm{e}^{-\int_0^t r(S_s,s)\,ds} \sum_{i=1}^m \frac{\partial P}{\partial x_i}\,(S_t,t) S_i(t) \sum_{k=1}^d \sigma_{ik}(S_t,t)\, dB_k^*(t)\ ,
\end{aligned} \tag{13.34}$$

where $a = \sigma\sigma^*$ as usual. Now, as $\widetilde{V}$ is a $\mathrm{P}^*$-martingale, necessarily the finite variation part in the previous differential must vanish, i.e.

$$\frac{\partial P}{\partial t}(x,t) + L_t P(x,t) - r(x,t)\,P(x,t) = 0, \quad \text{on } \mathbb{R}^m \times [0,T[\ , \tag{13.35}$$

where $L_t$ is the generator of $S$ under the risk neutral measure i.e.

$$L_t = r(x,t)\sum_{i=1}^m x_i\,\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^m a_{ij}(x,t)\, x_i\, x_j\, \frac{\partial^2}{\partial x_i \partial x_j}\ \cdot$$

**Theorem 13.4** Let $h : \mathbb{R}^m \to \mathbb{R}^+$ be a continuous function such that $h(x) \le C(1+|x|^\alpha)$ for some $C, \alpha > 0$ and let

$$P(x,t) = \mathrm{E}^*\Big[\mathrm{e}^{-\int_t^T r(S_s^{x,t},s)\,ds}\, h(S_T^{x,t})\Big] \tag{13.36}$$

(continued)

**Theorem 13.4** (continued)
be the price function of the option $h(S_T)$ with maturity $T$ in the generalized Black–Scholes model. Then $P$ is the solution of the PDE problem

$$\begin{cases} \dfrac{\partial P}{\partial t}(x,t) + L_t P(x,t) - r(x,t)\,P(x,t) = 0, \quad \text{on } \mathbb{R}^m \times [0,T[ \\ P(x,T) = h(x)\ . \end{cases} \tag{13.37}$$

*Proof* In the computation above we obtained (13.37) under the assumption that $P$ is of class $C^{2,1}$, which is still to be proved. In order to achieve this point we use Theorem 10.6, which states that the PDE problem (13.37) has a solution and that it coincides with the left-hand side of (13.36).

Unfortunately Theorem 10.6 requires the diffusion coefficient to be elliptic, whereas the matrix of the second-order derivatives of $L_t$ vanishes at the origin (and becomes singular on the axes), hence it cannot be applied immediately. The somehow contorted but simple argument below is developed in order to circumvent this difficulty.

The idea is simply to consider the process of the logarithm of the prices, whose generator is elliptic, and to express the price of the option in terms of this logarithm.

For simplicity, for $x \in \mathbb{R}^m$, let us denote $(e^{x_1}, \dots, e^{x_m})$ by $e^x$. By a repetition of the argument leading to (13.16), if $\zeta_i(t) = \log S_i(t)$ then

$$d\zeta_i(t) = \Big(r(e^{\zeta(t)}, t)\,dt - \frac{1}{2}\,a_{ii}(e^{\zeta(t)}, t)\Big)\,dt + \sum_{j=1}^{d} \sigma_{ik}(e^{\zeta(t)}, t)\,dB^*_k(t)\ ,$$

where $a = \sigma\sigma^*$ and with the starting condition $\zeta_i(s) = \log x_i$. Hence $\zeta$ is a diffusion with generator

$$\overline{L}_t = \sum_{i=1}^{m} \Big(r(e^x, t) - \frac{1}{2}a_{ii}(e^x, t)\Big)\,\frac{\partial}{\partial x_i} + \sum_{i,j=m}^{m} a_{ij}(e^x, t)\,\frac{\partial^2}{\partial x_i \partial x_j}\ \cdot \tag{13.38}$$

Now for the option price we can write

$$P(e^x, t) = E^*\Big[e^{-\int_t^T r(e^{\zeta_s^{x,t}}, s)\,ds}\, h(e^{\zeta_T^{x,t}})\Big]\ .$$

As the generator $\overline{L}_t$ in (13.38) satisfies the assumptions of Theorem 10.6 (in particular, its diffusion coefficient is uniformly elliptic), the function $\overline{P}(x,t) := P(e^x, t)$ is a solution of

$$\begin{cases} \dfrac{\partial \overline{P}}{\partial t}(x,t) + \overline{L}_t \overline{P}(x,t) - r(\mathrm{e}^x,t)\,\overline{P}(x,t) = 0, & \text{on } \mathbb{R}^m \times [0,T[ \\ \overline{P}(x,T) = h(\mathrm{e}^x)\ , \end{cases}$$

which easily implies that $P$ satisfies (13.37). □

From (13.34) we now have

$$d\widetilde{V}_t = \sum_{i=1}^{m} \frac{\partial P}{\partial x_i}(S_t,t)\widetilde{S}_i(t) \sum_{k=1}^{d} \sigma_{ik}(S_t,t)\, dB_k^*(t) = \sum_{i=1}^{m} \frac{\partial P}{\partial x_i}(S_t,t)\, d\widetilde{S}_i(t)$$

so that, by Proposition 13.1, we find the replicating strategy, defined by

$$\begin{cases} H_i(t) = \dfrac{\partial}{\partial x_i} P(S_t,t) & i = 1,\dots,m \\ H_0(t) = \widetilde{V}_t - \displaystyle\sum_{i=1}^{m} H_i(t)\widetilde{S}_i(t)\ , \end{cases} \tag{13.39}$$

which is self-financing and, as $\widetilde{V}_t \geq 0$ by construction, also admissible.

Equation (13.35) is sometimes called the *fundamental PDE following from the no-arbitrage approach*. Theorem 13.4 gives a way of computing the price $P$ of the option with payoff $h(S_T)$, just by solving the PDE problem (13.37), possibly numerically.

Note the not-so-intuitive fact that the price does not depend on the drift $b$ appearing in the dynamic (13.15) of the prices in the generalized Black–Scholes model.

This means in particular that it is irrelevant, in order to consider the price of an option, whether the price of the underlying asset follows an increasing trend (corresponding to $b_i(x,t) > 0$) or decreasing ($b_i(x,t) < 0$).

Recalling that $\widetilde{V}_t = \mathrm{e}^{-\int_0^t r(S_s,s)\,ds} P(S_t,t)$, (13.39) can also be written as

$$\begin{aligned} H_i(t) &= \frac{\partial P}{\partial x_i}(S_t,t),\ i = 1,\dots,m \\ H_0(t) &= \mathrm{e}^{-\int_0^t r(S_s,s)\,ds}\Big(P(S_t,t) - \sum_{i=1}^{m} H_i(t)\,S_i(t)\Big)\ . \end{aligned} \tag{13.40}$$

The quantities $H_i(t)$, $i = 1, \dots, m$, in (13.39) are also called the *deltas* of the option and are usually denoted by $\Delta$:

$$\Delta_i(S_t, t) = \frac{\partial P}{\partial x_i}(S_t, t), \quad i = 1, \dots, m\,.$$

The delta is related to the sensitivity of the price with respect to the values of the underlying asset prices. In particular, (13.40) states that the replicating portfolio must contain a large amount in the $i$-th underlying asset if the price of the option is very sensitive to changes of the price of the $i$-th underlying.

The delta is a special case of a *Greek*.

The Greeks are quantities giving the sensitivity of the price with respect to the parameters of the model. The name "Greeks" comes from the fact that they are usually (but not all of them...) denoted by Greek letters. They are taken into special account by practitioners, because of the particular financial meanings of each of them. The most used Greeks can be summarized as follows:

- *delta*: sensitivity of the price of the option w.r.t. the initial value of the price of the underlying: $\Delta_i = \frac{\partial P}{\partial x_i}$;
- *gamma*: sensitivity of the delta w.r.t. the initial values of the price of the underlying: $\Gamma_{ij} = \frac{\partial^2 P}{\partial x_i \partial x_j}$;
- *theta*: sensitivity of the price w.r.t. the initial time: $\Theta = \frac{\partial P}{\partial t}$;
- *Rho*: sensitivity of the price w.r.t. the spot rate: $Rho = \frac{\partial P}{\partial r}$;
- *Vega*: sensitivity of the price w.r.t. the volatility: $Vega = \frac{\partial P}{\partial \sigma}$.

Obviously, in the Rho and Vega cases, the derivatives must be understood in a suitably functional way whenever $r$ and $\sigma$ are not modeled as constants.

The last two Greeks, Rho and Vega, give the behavior of the price and then of the portfolio with respect to purely financial quantities (i.e. the interest rate and the volatility), whereas the other ones (delta, gamma and theta) give information about the dependence of the portfolio with respect to parameters connected to the assets on which the European option is written (the starting time and the prices of the assets).

## 13.7 The standard Black–Scholes model

In this section we derive explicit formulas for the price of a call and put option as well as the associated Greeks in a classical one-dimensional model, the *standard Black–Scholes model*. By this we mean the particular case where there is only one risky asset and the volatility $\sigma$ and the spot rate $r$ are constant.

Under the risk-neutral measure $\mathrm{P}^*$, the price of the risk asset evolves as

$$\frac{dS_t}{S_t} = r\,dt + \sigma\,dB^*_t \tag{13.41}$$

and the price at time $t$ of the call option with maturity $T$ is given by

$$P_{call}(S_t, t) = \mathrm{E}^*\Big[\mathrm{e}^{-r(T-t)}\,(S_T - K)^+ \,\Big|\, \mathscr{F}_t\Big],$$

where $K$ stands for the strike price. If we use the notation $S^{x,t}$ to denote the solution $S$ of (13.41) starting at $x$ at time $t$, then

$$S^{x,t}_s = x\,\mathrm{e}^{(r-\frac{1}{2}\sigma^2)(s-t)+\sigma(B^*_s - B^*_t)}, \qquad s \ge t\,,$$

and we have

$$\begin{aligned} P_{call}(x,t) &= \mathrm{E}^*\Big[\mathrm{e}^{-r(T-t)}\,(S^{x,t}_T - K)^+\Big] \\ &= \mathrm{e}^{-r(T-t)}\,\mathrm{E}^*\Big[\Big(x\,\mathrm{e}^{(r-\frac{1}{2}\sigma^2)(T-t)+\sigma(B^*_T - B^*_t)} - K\Big)^+\Big]\,. \end{aligned} \tag{13.42}$$

The expectation above can be computed remarking that, with respect to $\mathrm{P}^*$, $B^*_T - B^*_t$ has the same distribution as $\sqrt{T-t}\,Z$ with $Z \sim N(0,1)$, so that

$$P_{call}(x,t) = \mathrm{e}^{-r(T-t)}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\Big(x\mathrm{e}^{(r-\frac{1}{2}\sigma^2)(T-t)+\sigma\sqrt{T-t}\,z} - K\Big)^+\,\mathrm{e}^{-z^2/2}\,dz\,.$$

The integrand vanishes if $x\mathrm{e}^{(r-\frac{1}{2}\sigma^2)(T-t)+\sigma\sqrt{T-t}\,z} \le K$, i.e. for $z \le d_0(x, T-t)$, where

$$d_0(x,t) = \frac{1}{\sigma\sqrt{t}}\Big(-\log\frac{x}{K} - \big(r - \frac{1}{2}\sigma^2\big)t\Big)$$

so that

$$\begin{aligned} &P_{call}(x,t) \\ &= \mathrm{e}^{-r(T-t)}\frac{1}{\sqrt{2\pi}}\int_{d_0(x,T-t)}^{+\infty}\Big(x\mathrm{e}^{(r-\frac{1}{2}\sigma^2)(T-t)+\sigma\sqrt{T-t}\,z} - K\Big)\,\mathrm{e}^{-z^2/2}\,dz\,. \end{aligned} \tag{13.43}$$

This integral can be computed with a simple if not amusing computation, as already developed in Exercise 1.13. If we denote by $\Phi$ the partition function of a $N(0,1)$-

distributed r.v. we have from (13.43)

$$\begin{aligned} P_{call}(x,t) &= \frac{x}{\sqrt{2\pi}} \int_{d_0(x,T-t)}^{+\infty} e^{-\frac{1}{2}(z-\sigma\sqrt{T-t})^2}\, dz - \frac{K e^{-r(T-t)}}{\sqrt{2\pi}} \int_{d_0(x,T-t)}^{+\infty} e^{-z^2/2}\, dz \\ &= \frac{x}{\sqrt{2\pi}} \int_{d_0(x,T-t)-\sigma\sqrt{T-t}}^{+\infty} e^{-z^2/2}\, dz - K e^{-r(T-t)} \Phi(-d_0(x,T-t)) \\ &= x\Phi(-d_0(x,T-t) + \sigma\sqrt{T-t}) - K e^{-r(T-t)} \Phi(-d_0(x,T-t)) \,. \end{aligned}$$

It is customary to introduce the following two quantities

$$\begin{aligned} d_1(x,t) &= -d_0(x,t) + \sigma\sqrt{t} = \frac{1}{\sigma\sqrt{t}} \Big( \log\frac{x}{K} + \big(r + \frac{1}{2}\sigma^2\big)\, t \Big) \\ d_2(x,t) &= -d_0(x,t) = \frac{1}{\sigma\sqrt{t}} \Big( \log\frac{x}{K} + \big(r - \frac{1}{2}\sigma^2\big)\, t \Big) \end{aligned}$$

so that finally the price of the call option is given by the classical *Black-Scholes formula* (see Fig. 13.1)

$$P_{call}(x,t) = x\,\Phi(d_1(x,T-t)) - K\, e^{-r(T-t)}\, \Phi(d_2(x,T-t)) \,. \qquad (13.44)$$

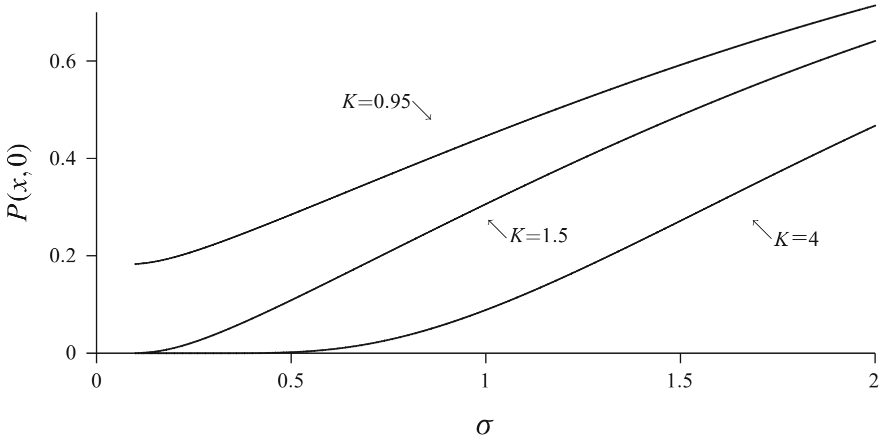


**Fig. 13.1** Behavior of the price of a call option as a function of $\sigma$, on the basis of the Black-Scholes formula for $x = 1, r = .15, T = 1$ and different values of the strike price $K$. As $\sigma \to 0$ the price tends to 0 if $\log\frac{K}{x} - rT > 0$, otherwise it tends to $x - Ke^{-rT}$

Let us determine the hedging portfolio. Again by straightforward computations, one obtains

$$\Delta_{call}(x,t) = \frac{\partial P_{call}}{\partial x}(x,t) = \Phi(d_1(x, T-t)) \,. \tag{13.45}$$

Hence, recalling formulas (13.40), a hedging portfolio for the call option in this model is given by

$$\begin{aligned} H_1(t) &= \Delta_{call}(S_t, t) = \Phi(d_1(S_t, T-t)) \\ H_0(t) &= \mathrm{e}^{-rt}\Big(P(S_t,t) - \Phi(d_1(S_t, T-t))S_t\Big) \,. \end{aligned} \tag{13.46}$$

Note that we can write

$$d_1(S_t, T-t) = \frac{1}{\sigma}\Big(\frac{1}{\sqrt{T-t}}\,\log\frac{S_t}{K} + \Big(r + \frac{1}{2}\,\sigma^2\Big)\sqrt{T-t}\Big) \,.$$

If the price $S_t$ remains $> K$ for $t$ near $T$ then $d_1(S_t, T-t)$ will approach $+\infty$ and, thanks to (13.46), $H_1(t)$ will be close to 1. This is in accordance with intuition: if the price of the underlying asset is larger than the strike it is reasonable to expect that the call will be exercised and therefore it is wise to keep in the replicating portfolio a unit of the underlying.

As for the put option, one could use similar arguments or also the call-put parity property

$$P_{call}(x,t) - P_{put}(x,t) = \mathrm{e}^{-r(T-t)}\,\mathrm{E}^*[(S_T^{x,t} - K)] = x - K\mathrm{e}^{-r(T-t)}$$

as explained later in Remark 13.3. The associated price and delta are therefore given by the formulas

$$\begin{aligned} P_{put}(x,t) &= K\,\mathrm{e}^{-r(T-t)}\,\Phi(-d_2(x,T-t)) - x\,\Phi(-d_1(x,T-t)), \\ \Delta_{put}(x,t) &= \frac{\partial P_{put}}{\partial x}(x,t) = \Phi(d_1(x,T-t)) - 1 \,. \end{aligned} \tag{13.47}$$

The other Greeks can also be explicitly written as summarized below (denoting by $\phi = \Phi'$ the standard $N(0,1)$ density):

$$\begin{aligned}
&\bullet\ \Gamma_{call}(x,t) = \frac{\phi(d_1(x,T-t))}{x\,\sigma\,\sqrt{T-t}} = \Gamma_{put}(x,t)\\
&\bullet\ \Theta_{call}(x,t) = -\frac{x\,\phi(d_1(x,T-t))\,\sigma}{2\,\sqrt{T-t}} - r\,x\,\mathrm{e}^{-r(T-t)}\,\phi(d_2(x,T-t))\\
&\bullet\ \Theta_{put}(x,t) = -\frac{x\,\phi(d_1(x,T-t))\,\sigma}{2\,\sqrt{T-t}} + r\,x\,\mathrm{e}^{-r(T-t)}\,\Phi(-d_2(x,T-t))\\
&\bullet\ Rho_{call}(x,t) = x\,(T-t)\,\mathrm{e}^{-r(T-t)}\,\Phi(d_2(x,T-t))\\
&\bullet\ Rho_{put}(x,t) = -x\,(T-t)\,\mathrm{e}^{-r(T-t)}\,\Phi(-d_2(x,T-t))\\
&\bullet\ Vega_{call}(x,t) = x\,\sqrt{T-t}\,\phi(d_1(x,T-t)) = Vega_{put}(x,t)\ .
\end{aligned}\tag{13.48}$$

These expressions of the Greeks can be obtained by the Black–Scholes formula with straightforward computations. Let us derive explicitly the formula for the Vega, whose value has some important consequences. Let us consider the case of the call option. We must take the derivative of the price with respect to $\sigma$. Instead of taking the derivative in (13.44), it is simpler to differentiate (13.43) under the integral sign. We then have, since the integrand in (13.43) vanishes at $d_0(x,T-t)$,

$$\begin{aligned}
&\frac{\partial P_{call}}{\partial\sigma}(x,t)\\
&= -\frac{\mathrm{e}^{-r(T-t)}}{\sqrt{2\pi}}\,\frac{\partial d_0}{\partial\sigma}(x,T-t)\times\underbrace{\Big(\big(x\mathrm{e}^{(r-\frac{1}{2}\sigma^2)(T-t)+\sqrt{T-t}\,\sigma z}-K\big)\,\mathrm{e}^{-z^2/2}\Big)\Big|_{z=d_0(x,T-t)}}_{=0}\\
&\qquad + \frac{\mathrm{e}^{-r(T-t)}}{\sqrt{2\pi}}\int_{d_0(x,T-t)}^{+\infty}\frac{\partial}{\partial\sigma}\Big(x\mathrm{e}^{(r-\frac{1}{2}\sigma^2)(T-t)+\sqrt{T-t}\,\sigma z}-K\Big)\,\mathrm{e}^{-z^2/2}\,dz\\
&= \frac{x}{\sqrt{2\pi}}\int_{d_0(x,T-t)}^{+\infty}\big(-\sigma(T-t)+\sqrt{T-t}\,z\big)\mathrm{e}^{-\frac{1}{2}\sigma^2(T-t)+\sqrt{T-t}\,\sigma z}\,\mathrm{e}^{-z^2/2}\,dz\\
&= \frac{x}{\sqrt{2\pi}}\int_{d_0(x,T-t)}^{+\infty}\big(-\sigma(T-t)+\sqrt{T-t}\,z\big)\,\mathrm{e}^{-\frac{1}{2}(z-\sigma\sqrt{T-t})^2}\,dz
\end{aligned}$$

and with the change of variable $y = z-\sigma\sqrt{T-t}$ we finally obtain

$$\begin{aligned}
\frac{\partial P_{call}}{\partial\sigma}(x,t) &= \frac{x}{\sqrt{2\pi}}\int_{d_0(x,T-t)-\sigma\sqrt{T-t}}^{+\infty}\sqrt{T-t}\,y\,\mathrm{e}^{-y^2/2}\,dy\\
&= \frac{x\sqrt{T-t}}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}(d_0(x,T-t)-\sigma\sqrt{T-t})^2}
\end{aligned}$$

which, recalling that $d_0(x, T-t) - \sigma\sqrt{T-t} = -d_1(x, T-t)$, is the result that was claimed in (13.48). In particular, the Vega in the standard Black–Scholes model is strictly positive, i.e.

> the price is a strictly increasing function of the volatility $\sigma$ (Fig. 13.1).

*Example 13.2* Let us assume $\sigma^2 = 0.2$, $r = 0.1$, $T = 1$, $K = 1$. What is the price of a call option at time $t = 0$ if the price of the underlying asset is $x = 0.9$? What is the composition, always at time 0, of the hedging portfolio?

This is a numerical computation, requiring appropriate software in order to perform the evaluations requested by formulas (13.44) and (13.46). The results are

$$\begin{aligned} P_{call}(0.9, 0) &= 0.16 \\ H_1(0) &= 0.58 \\ H_0(0) &= -0.37\,. \end{aligned}$$

How do we expect these quantities to change if the price at $t = 0$ were $x = 1.1$?

In this case the option is more likely to be exercised, the initial price $x$ now being greater than the strike. Hence we expect the price to be higher and the hedging portfolio to contain a larger amount of the underlying asset. The numerical computation gives

$$\begin{aligned} P_{call}(1.1, 0) &= 0.29 \\ H_1(0) &= 0.75 \\ H_0(0) &= -0.53 \end{aligned}$$

in accordance with what was expected.

And what if, still with $x = 1.1$, we assumed $\sigma^2 = 0.5$? We know that the price will increase. In fact

$$\begin{aligned} P_{call}(1.1, 0) &= 0.38 \\ H_1(0) &= 0.74 \\ H_0(0) &= -0.42\,. \end{aligned}$$

*Example 13.3 (Barrier options)* The payoff of a European option (Definition 13.1) can be any positive $\mathscr{F}_T$-measurable r.v. Puts and calls are particular cases that are functions of the value of the price of the underlying asset at time $T$. In practice more general situations are possible, in particular with payoffs that are functions of the whole path of the price. These are generally referred to as *path dependent* options.

A typical example is the payoff

$$Z = (S_T - K)^+ 1_{\{\sup_{0\le s\le T} S_s \le U\}} , \tag{13.49}$$

where $U$ is some value larger than the strike $K$. The holder of this option receives at time $T$ the amount $(S_T - K)^+$ (as for a classical call option) but under the constraint that the price has never crossed the level $U$ (the barrier) before time $T$. Many variations of this type are possible, combining the type of option (put or call as in (13.49)) with the action at the crossing of the barrier that may cancel the option as in (13.49) or activate it. In the financial jargon the payoff (13.49) is an *up and out call option*.

In this example we determine the price at time $t = 0$ of the option with payoff (13.49) under the standard Black–Scholes model. The general formula (13.12) gives the value

$$p = \mathrm{E}^*\big[\mathrm{e}^{-rT}(S_T - K)^+ 1_{\{\sup_{0\le t\le T} S_t \le U\}}\big] , \tag{13.50}$$

where, under $\mathrm{P}^*$,

$$S_t = x\mathrm{e}^{(r-\frac{1}{2}\sigma^2)t+\sigma B_t}$$

and $x = S_0$ is the price of the underlying at time 0. The computation of the expectation in (13.50) requires the knowledge of the joint distribution of $S_T$ and its running maximum $\sup_{0\le t\le T} S_t$ or, equivalently, of the Brownian motion with drift $\frac{1}{\sigma}(r - \frac{1}{2}\sigma^2)$ and of its running maximum. The idea is to first make a change of probability in order to "remove" the drift. The joint distribution of a Brownian motion and of its running maximum is actually known (Corollary 3.3). Let

$$Z_T = \exp\Big( -\frac{1}{\sigma}\big(r - \frac{1}{2}\sigma^2\big)B_T - \frac{1}{2\sigma^2}\big(r - \frac{1}{2}\sigma^2\big)^2 T\Big) ,$$

then with respect to Q, $d\mathrm{Q} = Z_T\, d\mathrm{P}^*$, the process

$$W_t = B_t + \frac{1}{\sigma}\Big(r - \frac{\sigma^2}{2}\Big)\, t$$

(continued)

*Example 13.3* (continued)

is a Brownian motion. Hence

$$\begin{aligned}&\mathrm{E}[\mathrm{e}^{-rT}(S_T-K)^+1_{\{\sup_{0\le t\le T}S_t\le U\}}]\\&=\mathrm{e}^{-rT}\mathrm{E}^{\mathrm{Q}}\Big[Z_T^{-1}(x\mathrm{e}^{\sigma W_T}-K)^+1_{\{\sup_{0\le t\le T}x\mathrm{e}^{\sigma W_t}\le U\}}\Big]\,.\end{aligned}$$

Now, as $B_T=W_T-\frac{1}{\sigma}\,(r-\frac{\sigma^2}{2})T$,

$$\begin{aligned}Z_T^{-1}&=\exp\Big[\frac{1}{\sigma}\,\big(r-\frac{1}{2}\,\sigma^2\big)B_T+\frac{1}{2\sigma^2}\,\big(r-\frac{1}{2}\,\sigma^2\big)^2T\Big]\\&=\exp\Big[\frac{1}{\sigma}\,\big(r-\frac{1}{2}\,\sigma^2\big)W_T-\frac{1}{2\sigma^2}\,\big(r-\frac{1}{2}\,\sigma^2\big)^2T\Big]\end{aligned}$$

so that the price $p$ of the option is equal to

$$\mathrm{e}^{-rT-\frac{1}{2\sigma^2}\,(r-\frac{1}{2}\,\sigma^2)^2T}\mathrm{E}^{\mathrm{Q}}\Big[\mathrm{e}^{\frac{1}{\sigma}\,(r-\frac{1}{2}\,\sigma^2)W_T}(x\mathrm{e}^{\sigma W_T}-K)^+1_{\{\sup_{0\le t\le T}x\mathrm{e}^{\sigma W_t}\le U\}}\Big]\,.$$

By Corollary 3.3 the joint law of $(W_T,\sup_{0\le t\le T}W_t)$ has density

$$f(z,y)=\Big(\frac{2}{\pi T^3}\Big)^{1/2}(2y-z)\,\mathrm{e}^{-\frac{1}{2T}\,(2y-z)^2}\,,\quad y>z,y>0\,.$$

Hence

$$\begin{aligned}&\mathrm{E}^{\mathrm{Q}}\Big[\mathrm{e}^{\frac{1}{\sigma}\,(r-\frac{1}{2}\,\sigma^2)W_T}(x\mathrm{e}^{\sigma W_T}-K)^+1_{\{\sup_{0\le t\le T}x\mathrm{e}^{\sigma W_t}\le U\}}\Big]\\&=\int_0^{\xi_2}dy\int_{-\infty}^{y}\mathrm{e}^{\frac{1}{\sigma}\,(r-\frac{1}{2}\,\sigma^2)z}(x\mathrm{e}^{\sigma z}-K)^+f(z,y)\,dz\end{aligned}$$

where $\xi_2=\frac{1}{\sigma}\,\log(\frac{U}{x})$. Note that the integrand vanishes for $z\le\xi_1:=\frac{1}{\sigma}\,\log(\frac{K}{x})$, so that we are left with the computation of

$$\begin{aligned}&\int_0^{\xi_2}dy\int_{\xi_1}^{y}\mathrm{e}^{\frac{1}{\sigma}\,(r-\frac{1}{2}\,\sigma^2)z}(x\mathrm{e}^{\sigma z}-K)f(z,y)\,dz\\&=\Big(\frac{2}{\pi T^3}\Big)^{1/2}\int_{\xi_1}^{\xi_2}dz\int_{z\vee 0}^{\xi_2}\mathrm{e}^{\frac{1}{\sigma}\,(r-\frac{1}{2}\,\sigma^2)z}(x\mathrm{e}^{\sigma z}-K)\,(2y-z)\,\mathrm{e}^{-\frac{1}{2T}\,(2y-z)^2}\,dy\,.\end{aligned}$$

(continued)

*Example 13.3* (continued)

Replacing the variable $y$ with $\zeta = 2y - z$ the integral becomes

$$\begin{aligned}
&\frac{1}{2}\int_{\xi_1}^{\xi_2} e^{\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)z}(xe^{\sigma z}-K)\,dz\int_{2(z\vee 0)-z}^{2\xi_2-z}\zeta\, e^{-\frac{1}{2T}\zeta^2}\,d\zeta\\
&=\frac{T}{2}\int_{\xi_1}^{\xi_2} e^{\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)z}(xe^{\sigma z}-K)(e^{-\frac{1}{2T}(2(z\vee 0)-z)^2}-e^{-\frac{1}{2T}(2\xi_2-z)^2})\,dz\,.
\end{aligned}$$

Finally, the requested price is equal to

$$\begin{aligned}
&e^{-rT-\frac{1}{2\sigma^2}(r-\frac{1}{2}\sigma^2)^2T}\frac{1}{\sqrt{2\pi T}}\\
&\times\int_{\xi_1}^{\xi_2} e^{\frac{1}{\sigma}(r-\frac{1}{2}\sigma^2)z}(xe^{\sigma z}-K)(e^{-\frac{1}{2T}(2(z\vee 0)-z)^2}-e^{-\frac{1}{2T}(2\xi_2-z)^2})\,dz\,.
\end{aligned}\tag{13.51}$$

From this it is possible to deduce a closed formula, having the flavor of the Black–Scholes formula, only a bit more complicated.

**Proposition 13.7** In the standard Black-Scholes model with spot interest rate $r$ and volatility $\sigma$, the price $p$ at time 0 as a function of the initial price $x$ of an up-and-out call barrier option with strike $K$, maturity $T$ and barrier $U$ is given by:

$$\begin{aligned}
&x\,(\Phi(a_1)-\Phi(a_2))-Ke^{-rT}\,(\Phi(b_1)-\Phi(b_2))\\
&-xe^{-rT-\frac{T}{2}(\frac{r}{\sigma}-\frac{\sigma}{2})^2-\frac{2}{T}\xi_2^2+\frac{1}{2}(\frac{2rT+\sigma^2T+4\sigma\xi_2}{2\sigma\sqrt{T}})^2}\,(\Phi(c_1)-\Phi(c_2))\,+\\
&+Ke^{-rT-\frac{T}{2}(\frac{r}{\sigma}-\frac{\sigma}{2})^2-\frac{2}{T}\xi_2^2+\frac{1}{2}(\frac{2rT-\sigma^2T+4\sigma\xi_2}{2\sigma\sqrt{T}})^2}\,(\Phi(d_1)-\Phi(d_2))\,,
\end{aligned}\tag{13.52}$$

where $\Phi$ is again the cumulative distribution function of a $N(0,1)$-distributed r.v. and

$$\begin{aligned}
\xi_1&=\frac{1}{\sigma}\log(\tfrac{K}{x}), & \xi_2&=\frac{1}{\sigma}\log(\tfrac{U}{x}),\\
a_1&=\frac{\xi_2}{\sqrt{T}}-\sqrt{T}\,\frac{2r+\sigma^2}{2\sigma}, & a_2&=\frac{\xi_1}{\sqrt{T}}-\sqrt{T}\,\frac{2r+\sigma^2}{2\sigma},\\
b_1&=\frac{\xi_2}{\sqrt{T}}-\sqrt{T}\,\frac{2r-\sigma^2}{2\sigma}, & b_2&=\frac{\xi_1}{\sqrt{T}}-\sqrt{T}\,\frac{2r-\sigma^2}{2\sigma},\\
c_1&=\frac{\xi_2}{\sqrt{T}}-\frac{2rT+\sigma^2T+4\sigma\xi_2}{2\sigma\sqrt{T}}, & c_2&=\frac{\xi_1}{\sqrt{T}}-\frac{2rT+\sigma^2T+4\sigma\xi_2}{2\sigma\sqrt{T}},\\
d_1&=\frac{\xi_2}{\sqrt{T}}-\frac{2rT-\sigma^2T+4\sigma\xi_2}{2\sigma\sqrt{T}}, & d_2&=\frac{\xi_1}{\sqrt{T}}-\frac{2rT-\sigma^2T+4\sigma\xi_2}{2\sigma\sqrt{T}}\,.
\end{aligned}$$

(continued)

*Example 13.3* (continued)
*Proof* The integral in (13.51) can be decomposed into the sum of four terms:

$$x\int_{\xi_1}^{\xi_2} \mathrm{e}^{(\frac{r}{\sigma}+\frac{\sigma}{2})z-\frac{1}{2T}z^2}\,dz - x\,\mathrm{e}^{-\frac{2}{T}\xi_2^2}\int_{\xi_1}^{\xi_2} \mathrm{e}^{(\frac{r}{\sigma}+\frac{\sigma}{2}+\frac{2}{T}\xi_2)z-\frac{1}{2T}z^2}\,dz$$
$$-K\int_{\xi_1}^{\xi_2} \mathrm{e}^{(\frac{r}{\sigma}-\frac{\sigma}{2})z-\frac{1}{2T}z^2}\,dz + K\mathrm{e}^{-\frac{2}{T}\xi_2^2}\int_{\xi_1}^{\xi_2} \mathrm{e}^{(\frac{r}{\sigma}-\frac{\sigma}{2}+\frac{2}{T}\xi_2)z-\frac{1}{2T}z^2}\,dz$$

and we treat them separately. The first term on the right-hand side can be rewritten as

$$x\mathrm{e}^{\frac{T}{2}(\frac{2r+\sigma^2}{2\sigma})^2}\int_{\xi_1}^{\xi_2} \mathrm{e}^{-\frac{1}{2}(\frac{z}{\sqrt{T}}-\sqrt{T}\frac{2r+\sigma^2}{2\sigma})^2}\,dz,$$

so that with the change of variable $u=\frac{z}{\sqrt{T}}-\sqrt{T}\frac{2r+\sigma^2}{2\sigma}$ we arrive at

$$x\sqrt{2\pi T}\,\mathrm{e}^{\frac{T}{2}(\frac{2r+\sigma^2}{2\sigma})^2}\int_{\frac{\xi_1}{\sqrt{T}}-\sqrt{T}\frac{2r+\sigma^2}{2\sigma}}^{\frac{\xi_2}{\sqrt{T}}-\sqrt{T}\frac{2r+\sigma^2}{2\sigma}} \frac{1}{\sqrt{2\pi T}}\mathrm{e}^{-\frac{u^2}{2}}\,du$$
$$= x\sqrt{2\pi T}\,\mathrm{e}^{\frac{T}{2}(\frac{2r+\sigma^2}{2\sigma})^2}\left(\Phi(a_1)-\Phi(a_2)\right)\ .$$

With similar changes of variables, the second term gives

$$-x\sqrt{2\pi T}\,\mathrm{e}^{-\frac{2}{T}\xi_2^2+\frac{1}{2}(\frac{2rT+\sigma^2T+4\sigma\xi_2}{2\sigma\sqrt{T}})^2}\left(\Phi(b_1)-\Phi(b_2)\right),$$

the third term

$$-K\sqrt{2\pi T}\,\mathrm{e}^{\frac{T}{2}(\frac{2r-\sigma^2}{2\sigma})^2}\left(\Phi(c_1)-\Phi(c_2)\right),$$

and the last one

$$K\sqrt{2\pi T}\,\mathrm{e}^{-\frac{2}{T}\xi_2^2+\frac{1}{2}(\frac{2rT-\sigma^2T+4\sigma\xi_2}{2\sigma\sqrt{T}})^2}\left(\Phi(d_1)-\Phi(d_2)\right).$$

Multiplying all these terms by $\mathrm{e}^{-rT-\frac{1}{2\sigma^2}(r-\frac{\sigma^2}{2})T}\frac{1}{\sqrt{2\pi T}}$ the proof is completed. □

As an example let us compare the price of a barrier option with the values $x=0.9$, $\sigma=0.2$, $r=0.1$, $T=1$, $K=1$ and with the barrier $U=2$. Hence in this case the option is canceled if before the maturity $T$ the price of the underlying becomes larger that 2.

(continued)

*Example 13.3* (continued)

We have obtained in Example 13.2 that without the barrier constraint the price of the option is 0.15. Computing with patience the value in (13.7) we find that the price of the corresponding barrier option becomes 0.09.

The barrier option is therefore considerably less expensive, but of course it does not protect from spikes of the price of the underlying asset. Taking $U = 3$ the price of the barrier option becomes 0.14, very close to the value without barrier.

*Remark 13.3 (The call-put parity)* Let $C_t$, resp. $P_t$, denote the price of a call, resp. a put, option on the same asset at time $t$. Assume that the spot rate $r$ is deterministic. Then the following call-put parity formula holds:

$$C_t = P_t + S_t - \mathrm{e}^{-\int_t^T r_s\,ds} K\,. \tag{13.53}$$

In fact, using the relation $z^+ - (-z)^+ = z$,

$$\begin{aligned} C_t - P_t &= \mathrm{E}^*\Big[\mathrm{e}^{-\int_t^T r_s\,ds}(S_T - K)^+ \,\Big|\, \mathscr{F}_t\Big] - \mathrm{E}^*\Big[\mathrm{e}^{-\int_t^T r_s\,ds}(K - S_T)^+ \,\Big|\, \mathscr{F}_t\Big] \\ &= \mathrm{E}^*\Big[\mathrm{e}^{-\int_t^T r_s\,ds}(S_T - K) \,\Big|\, \mathscr{F}_t\Big] = \mathrm{e}^{\int_0^t r_s\,ds}\mathrm{E}^*[\widetilde{S}_T | \mathscr{F}_t] - \mathrm{e}^{-\int_t^T r_s\,ds} K \\ &= \mathrm{e}^{\int_0^t r_s\,ds}\widetilde{S}_t - \mathrm{e}^{-\int_t^T r_s\,ds} K = S_t - \mathrm{e}^{-\int_t^T r_s\,ds} K\,. \end{aligned}$$

The relation (13.53) can also be obtained from the requirement of absence of arbitrage, without knowing the expression of the prices. Let us verify that a different relation between these prices would give rise to an opportunity of arbitrage.

Let us assume $C_t > P_t + S_t - K\mathrm{e}^{-\int_t^T r_s\,ds}$. One can then establish a portfolio buying a unit of the underlying asset and a put option and selling a call option. The price of the operation is $C_t - P_t - S_t$ and is covered through an investment of opposite sign in the riskless asset. This operation therefore does not require us to engage any capital. At maturity we dispose of a put option, a unit of the underlying asset and an amount of cash $-(C_t - P_t - S_t)\,\mathrm{e}^{\int_t^T r_s\,ds}$ and we have to fulfill a call.

There are two possibilities

- $S_T > K$. In this case the call is exercised; we sell the underlying, which allows us to honor the call and to collect an amount equal to $K$. The put is, of course, valueless. The global balance of the operation is

$$K - (C_t - P_t - S_t)\,\mathrm{e}^{\int_t^T r_s\,ds} > 0\,.$$

(continued)

*Remark 13.3* (continued)

- $S_T \le K$. In this case the call is not exercised. We use then the put in order to sell the unit of underlying at the price $K$. The global balance of operation is again $K - (C_t - P_t - S_t)\,\mathrm{e}^{\int_t^T r_s\,ds} > 0$.

In a similar way it is possible to establish an arbitrage portfolio if conversely $C_t < P_t + S_t - K\mathrm{e}^{-\int_t^T r_s\,ds}$.

Note that this arbitrage argument holds for very general models (the only requirement is that the spot interest rate must be deterministic).

(13.53) is useful because it allows us to derive the price of a call option from the price of the corresponding put and put options are somehow easier to deal with, as their payoff is a bounded r.v.

Note that almost all quantities appearing in the Black–Scholes formulas (13.43) and (13.47) are known in the market. The only unknown quantity is actually the volatility $\sigma$.

In practice, $\sigma$ is estimated empirically starting from the option prices already known: let us assume that in the market an option with strike $K$ and maturity $T$ is already traded with a price $z$. Let us denote by $C_{K,T}(\sigma)$ the price of a call option as a function of the volatility $\sigma$ for the given strike and volatility $K, T$. As $C_{K,T}$ is a strictly increasing function, we can determine the volatility as $\sigma = C_{K,T}^{-1}(z)$. In this way options whose price is already known allow us to determine the missing parameter $\sigma$ and in this way also the price of options not yet on the market.

Another approach to the question is to estimate $\sigma$ from the observed values of the underlying: actually nowadays the price of a financial asset is known at a high frequency. This means that, denoting the price process by $S$, the values $S_{t_1}, S_{t_2}, \dots$ at times $t_1, t_2, \dots$ are known. The question, which is mathematically interesting in itself, is whether it is possible starting from these data to estimate the value of $\sigma$, assuming that these values come from a path of a process following the Black–Scholes model.

The fact that the option price is an invertible function of the volatility also allows us to check the soundness of the Black–Scholes model. Assume that two options on the same underlying asset are present in the market, with strike and maturity $K_1, T_1$ and $K_2, T_2$ and prices $z_1, z_2$, respectively. If the Black–Scholes model was a good one, the value of the volatility computed by the inversion of the price function should be the same for the two options, i.e. the two quantities

$$C_{K_1,T_1}^{-1}(z_1) \quad \text{and} \quad C_{K_2,T_2}^{-1}(z_2)$$

should coincide. In practice, it has been observed that this is not the case. The standard Black–Scholes model, because of its assumption of constancy of the volatility, thus appears to be too rigid as a model of the real world financial markets.

Nevertheless, it constitutes an important first attempt and also the starting point of, very many, more complicated models that have been introduced in recent years.

## Exercises

**13.1** (p. 616) Assume that the prices $S_t = (S_0(t), S_1(t), \dots, S_m(t))$ follow the generalized Black–Scholes model (13.15) and that the spot rate process $(r_t)_t$ is independent of $(S_t)_t$ (this assumption is satisfied in particular if $(r_t)_t$ is deterministic). Assume, moreover, that there exists an equivalent martingale measure $\mathrm{P}^*$. Prove that, for every $i = 1, \dots, m$, and $\mathrm{E}^*[S_i(0)] < +\infty$,

$$\mathrm{E}^*[S_i(t)] = \mathrm{e}^{\int_0^t \mathrm{E}^*[r_s]\, ds}\mathrm{E}^*[S_i(0)] .$$

**13.2** (p. 617) Let us consider the generalized Black–Scholes model (13.15) which, we know, may not be complete in general, so that there are options that might not be attainable. We want to prove, however, that, if the spot interest rate $(r_t)_t$ is deterministic and there exists an equivalent martingale measure $\mathrm{P}^*$, then the option

$$Z = \int_0^T \big(\alpha S_i(s) + \beta\big)\, ds ,$$

$\alpha, \beta > 0$, is *certainly* attainable, for every $i = 1, \dots, m$.

a) Write down the differential of $S_i$ under $\mathrm{P}^*$ and compute $\mathrm{E}^*[S_i(t)\,|\,\mathscr{F}_s]$.
b) Compute $\mathrm{E}^*[\mathrm{e}^{-\int_t^T r_u\, du} Z\,|\,\mathscr{F}_t]$ and determine a replicating portfolio for the option $Z$ (check in particular that it is admissible).
c) Deduce that $Z$ is an attainable option and compute its price at time 0 as a function of the price $x = S_i(0)$.

**13.3** (p. 618) In a standard Black–Scholes model let us consider an investor that constructs a self-financing portfolio $V$ with the constraint that the value of the component invested into the risky asset is constant and equal to some value $M > 0$.

a) Assuming that the value $V_0$ at time 0 is deterministic and positive, compute the distribution of $V_t$ and compute $\mathrm{E}[V_t]$, as a function of $V_0$ and of the parameters $r, \mu, \sigma$ of the Black–Scholes model.
b) Is $V$ admissible?

**13.4** (p. 620) (Have a look at Example 12.5 first) Let us consider a standard Black–Scholes model with parameters $b, \sigma, r$. Consider an option, $Z$, that pays an amount $C$ if the price $S$ crosses a fixed level $K$, $K > 0$, before some fixed time $T$ and 0 otherwise.

a) Write down the payoff of this option.
b) Compute its no-arbitrage price at time 0.

**13.5** (p. 621) Let us consider a market with a riskless asset with price $S_0(t) = e^{rt}$ and two risky assets with prices

$$dS_i(t) = S_i(t)(\mu_i \, dt + \sigma \, dB_t), \quad S_i(0) = x_i \ ,$$

where $\mu_i \in \mathbb{R}$, $\mu_1 \neq \mu_2$, $r, \sigma > 0$ and $B$ is a real Brownian motion. Show that there are arbitrage possibilities by proving that an equivalent martingale measure does not exist and provide explicitly an arbitrage portfolio.

# Solutions of the Exercises

**1.1** If $X$ and $Y$ have the same law $\mu$, then they also have the same p.f., as

$$F_X(t) = \mathrm{P}(X \le t) = \mu(]-\infty, t]) = \mathrm{P}(Y \le t) = F_Y(t) .$$

Conversely, if $X$ and $Y$ have the same p.f. $F$, then, denoting by $\mu_X$ and $\mu_Y$ their respective laws and if $a < b$,

$$\mu_X(]a, b]) = \mathrm{P}(a < X \le b) = \mathrm{P}(X \le b) - \mathrm{P}(X \le a) = F(b) - F(a) .$$

Repeating the same argument for $\mu_Y$, $\mu_X$ and $\mu_Y$ coincide on the half-open intervals $]a, b]$, $a, b \in \mathbb{R}, a < b$. The family $\mathscr{C}$ formed by these half-open intervals is stable with respect to finite intersections (immediate) and generates the Borel $\sigma$-algebra. Actually a $\sigma$-algebra containing $\mathscr{C}$ necessarily contains any open interval $]a, b[$ (that is, the intersection of the half-open intervals $]a, b + \frac{1}{n}]$) and therefore every open set. By Carathéodory's criterion, Theorem 1.1, $\mu_X$ and $\mu_Y$ coincide on $\mathscr{B}(\mathbb{R})$.

**1.2**

a) If $x > 0$

$$F(x) = \mathrm{P}(X \le x) = \int_{-\infty}^{x} f(t)\, dt = \lambda \int_0^x \mathrm{e}^{-\lambda t}\, dt = 1 - \mathrm{e}^{-\lambda x} ,$$

whereas the same formula gives $F(x) = 0$ if $x < 0$ ($f$ vanishes on the negative real numbers). With some patience, integrating by parts, we find

$$\mathrm{E}[X] = \int_{-\infty}^{+\infty} x f(x)\, dx = \lambda \int_0^{+\infty} x\, \mathrm{e}^{-\lambda x}\, dx = \frac{1}{\lambda}$$

$$\mathrm{E}[X^2] = \int_{-\infty}^{+\infty} x^2 f(x)\, dx = \lambda \int_0^{+\infty} x^2 \mathrm{e}^{-\lambda x}\, dx = \frac{1}{\lambda} + \frac{1}{\lambda^2}$$

© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2

and $\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = \frac{1}{\lambda}$.

b) We have

$$\mathrm{E}[U] = \int_{-\infty}^{+\infty} xf(x)\,dx = \int_0^1 x\,dx = \frac{1}{2}$$
$$\mathrm{E}[U^2] = \int_{-\infty}^{+\infty} x^2 f(x)\,dx = \int_0^1 x^2\,dx = \frac{1}{3}$$

and $\mathrm{Var}(U) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$.

c1) We have

$$\mathrm{P}(Z \le z) = \mathrm{P}(\alpha U \le z) = \mathrm{P}(U \le \tfrac{z}{\alpha})\ .$$

Now

$$\mathrm{P}(U \le \tfrac{z}{\alpha}) = \begin{cases} 0 & \text{if } \frac{z}{\alpha} \le 0 \\ \frac{z}{\alpha} & \text{if } 0 \le \frac{z}{\alpha} \le 1 \\ 1 & \text{if } \frac{z}{\alpha} \ge 1 \end{cases}$$

hence

$$\mathrm{P}(Z \le z) = \begin{cases} 0 & \text{if } z \le 0 \\ \frac{z}{\alpha} & \text{if } 0 \le z \le \alpha \\ 1 & \text{if } z \ge \alpha\ . \end{cases}$$

Taking the derivative we find that the density of $Z$ is

$$f_Z(z) = \begin{cases} 0 & \text{if } z \le 0 \\ \frac{1}{\alpha} & \text{if } 0 \le z \le \alpha \\ 0 & \text{if } z \ge \alpha\ , \end{cases}$$

i.e. $Z$ is uniform on the interval $[0, \alpha]$.

c2) If $t > 0$ we have

$$F_W(t) = \mathrm{P}(-\tfrac{1}{\lambda}\log U \le t) = \mathrm{P}(U \ge \mathrm{e}^{-\lambda t}) = \int_{\mathrm{e}^{-\lambda t}}^1 dx = 1 - \mathrm{e}^{-\lambda t}$$

whereas $F_W(t) = 0$ for $t < 0$. Therefore $W$ has the same p.f. as an exponential law with parameter $\lambda$ and, by Exercise 1.1, has this law.

**1.3**

a) Let us denote by $\mu$ the law of $X$. By the integration rule with respect to an image law, Proposition 1.1, and by Fubini's theorem

$$\begin{aligned}\mathrm{E}[f(X)] &= \int_0^{+\infty} f(x)\,d\mu(x) = \int_0^{+\infty} d\mu(x)\Big(f(0) + \int_0^x f'(t)\,dt\Big)\\ &= f(0) + \int_0^{+\infty} f'(t)\,dt \int_t^{+\infty} d\mu(x) = f(0) + \int_0^{+\infty} f'(t)\mathrm{P}(X \ge t)\,dt\,.\end{aligned}$$

b) Imitating Fubini's theorem,

$$\mathrm{E}[X] = \sum_{n=1}^{\infty} n\,\mathrm{P}(X = n) = \sum_{n=1}^{\infty} \mathrm{P}(X = n)\sum_{k=1}^{n} 1 = \sum_{k=1}^{\infty}\sum_{n=k}^{\infty} \mathrm{P}(X = n) = \sum_{k=1}^{\infty} \mathrm{P}(X \ge k)\,.$$

**1.4**

a) If $\xi$ is a vector in the kernel of $C$ then, repeating the argument of (1.15),

$$0 = \langle C\xi, \xi\rangle = \mathrm{E}[\langle \xi, X\rangle^2]$$

and therefore $\langle \xi, X\rangle^2 = 0$ a.s., i.e. $X$ is orthogonal to $\xi$ a.s. As $C$ is symmetric, its image coincides with the subspace of the vectors that are orthogonal to its kernel: $z \in \mathrm{Im}\, C$ if and only if $\langle z, \xi\rangle = 0$ for every vector $\xi$ such that $C\xi = 0$. Hence, if $\xi_1, \dots, \xi_k$, $k \le m$, is a basis of the kernel of $C$, then $z \in \mathrm{Im}\, C$ if and only if $\langle z, \xi_i\rangle = 0$ for $i = 1, \dots, k$. In conclusion

$$\{X \in \mathrm{Im}\, C\} = \{\langle X, \xi_1\rangle = 0\} \cap \dots \cap \{\langle X, \xi_k\rangle = 0\}$$

and therefore the event $\{X \in \mathrm{Im}\, C\}$ has probability 1, being the intersection of a finite number of events of probability 1.

b) If $C$ is not invertible, $\mathrm{Im}\, C$ is a proper subspace of $\mathbb{R}^m$. The r.v. $X - \mathrm{E}(X)$ is centered and has the same covariance matrix as $X$ and, as we have seen, $X - \mathrm{E}(X) \in Im\ C$ with probability 1. If $X$ had a density, $f$ say, then we would have

$$1 = \mathrm{P}(X \in \mathrm{Im}\, C + \mathrm{E}(X)) = \int_{\mathrm{Im}\, C+\mathrm{E}(X)} f(x)\,dx\,.$$

But the integral on the right-hand side is equal to 0 because the hyperplane $\mathrm{Im}C+\mathrm{E}(X)$ has Lebesgue measure equal to 0 so that this is absurd.

**1.5** It is well-known that a sequence of real numbers $(a_n)_n$ converges to a limit $\ell$ if and only if from every subsequence of $(a_n)_n$ we can extract a further subsequence converging to $\ell$. Therefore $\|X_n - X\|_p \to 0$ as $n \to \infty$ if and only if from every subsequence of $(\|X_n - X\|_p)_n$ we can extract a further subsequence converging to 0. For convergence in probability the argument is similar.

For a.s. convergence this argument cannot be repeated, because to say that a subsequence converges a.s. means that there exists a negligible set $N$ *possibly depending on the subsequence* such that, on $N^c$, $X_{n_k}(\omega) \to_{k\to\infty} X(\omega)$. The argument above therefore allows us to say that the sequence $(X_n)_n$ converges to $X$ outside the union of all these negligible events. As the set of all possible subsequences of a given sequence has a cardinal larger than countable, the union of all these negligible events may have strictly positive probability.

Note also that, if this criterion was true for a.s. convergence, it would entail, using the last statement of Proposition 1.5, that convergence in probability implies a.s. convergence.

**1.6** If $X \sim N(0, I)$, then

$$\mathrm{E}[\mathrm{e}^{\langle\theta,X\rangle}] = \frac{1}{(2\pi)^{m/2}} \int_{\mathbb{R}^m} \mathrm{e}^{\langle\theta,x\rangle}\mathrm{e}^{-\frac{1}{2}|x|^2}\, dx = \mathrm{e}^{\frac{1}{2}|\theta|^2} \underbrace{\frac{1}{(2\pi)^{m/2}} \int_{\mathbb{R}^m} \mathrm{e}^{-\frac{1}{2}|x-\theta|^2}\, dx}_{=1} = \mathrm{e}^{\frac{1}{2}|\theta|^2} ,$$

as we recognized the expression of the $N(\theta, I)$ density. If, more generally, $X \sim N(b, \Gamma)$, then we know that $X = \Gamma^{1/2}Z + b$, where $Z \sim N(0, I)$. Therefore

$$\mathrm{E}[\mathrm{e}^{\langle\theta,X\rangle}] = \mathrm{E}[\mathrm{e}^{\langle\theta,\Gamma^{1/2}Z+b\rangle}] = \mathrm{e}^{\langle\theta,b\rangle}\mathrm{E}[\mathrm{e}^{\langle\Gamma^{1/2}\theta,Z\rangle}] = \mathrm{e}^{\langle\theta,b\rangle}\mathrm{e}^{\frac{1}{2}|\Gamma^{1/2}\theta|^2} .$$

But $|\Gamma^{1/2}\theta|^2 = \langle\Gamma^{1/2}\theta, \Gamma^{1/2}\theta\rangle = \langle\Gamma^{1/2}\Gamma^{1/2}\theta, \theta\rangle = \langle\Gamma\theta, \theta\rangle$.

**1.7** We have

$$\widehat{\mu}_{X+Y}(\theta) = \mathrm{E}[\mathrm{e}^{i\langle\theta,X+Y\rangle}] = \mathrm{E}[\mathrm{e}^{i\langle\theta,X\rangle}\mathrm{e}^{i\langle\theta,Y\rangle}] \underset{\uparrow}{=} \mathrm{E}[\mathrm{e}^{i\langle\theta,X\rangle}]\mathrm{E}[\mathrm{e}^{i\langle\theta,Y\rangle}] = \widehat{\mu}(\theta)\widehat{\nu}(\theta) ,$$

where the equality indicated by the arrow follows from Proposition 1.3, recalling that $X$ and $Y$ are independent.

If $X = (X_1, \dots, X_m)$ is a $\mu$-distributed $m$-dimensional r.v., then the $k$-th marginal, $\mu_k$, is nothing else than the law of $X_k$. Therefore, if we denote by $\widetilde{\theta}$ the vector of dimension $m$ whose components are all equal to 0 but for the $k$-th one that is equal to $\theta$,

$$\widehat{\mu}_k(\theta) = \mathrm{E}[\mathrm{e}^{i\theta X_k}] = \mathrm{E}[\mathrm{e}^{i\langle\widetilde{\theta},X\rangle}] = \widehat{\mu}(\widetilde{\theta}) .$$

**1.8** If $\theta \in \mathbb{R}$ then

$$\mathrm{E}[\mathrm{e}^{\theta X}] = \frac{\lambda}{2} \int_{-\infty}^{+\infty} \mathrm{e}^{-\lambda|x|}\mathrm{e}^{\theta x}\, dx = \frac{\lambda}{2}\Big( \int_0^{+\infty} \mathrm{e}^{(\theta-\lambda)x}\, dx + \int_{-\infty}^{0} \mathrm{e}^{(\theta+\lambda)x}\, dx\Big) .$$

It is apparent that in order for both integrals to converge the condition $|\theta| < \lambda$ is necessary. Once this condition is satisfied the integrals are readily computed and we obtain the expression of the Laplace transform

$$\mathrm{E}[\mathrm{e}^{\theta X}] = \frac{\lambda}{2}\Big(\frac{1}{\lambda-\theta} + \frac{1}{\lambda+\theta}\Big) = \frac{\lambda^2}{\lambda^2-\theta^2}\,\cdot$$

As for the characteristic function, since $x \mapsto \sin\theta x$ is an odd function whereas $x \mapsto \cos\theta x$ is even,

$$\mathrm{E}[\mathrm{e}^{i\theta X}] = \frac{\lambda}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{-\lambda|x|}\mathrm{e}^{i\theta x}\,dx = \frac{\lambda}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{-\lambda|x|}\cos\theta x\,dx = \lambda\int_{0}^{+\infty} \mathrm{e}^{-\lambda|x|}\cos\theta x\,dx\,.$$

The last integral can be computed by parts with some patience and we obtain

$$\mathrm{E}[\mathrm{e}^{i\theta X}] = \frac{\lambda^2}{\lambda^2+\theta^2}\,\cdot$$

We shall see in Sect. 5.7 that the characteristic function can be deduced from the Laplace transform in a simple way thanks to the property of uniqueness of the analytic continuation of holomorphic functions.

**1.9** The vector $(X, X+Y)$ can be obtained from $(X, Y)$ through the linear map associated to the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\,.$$

As $(X, Y)$ is $N(0, I)$-distributed, by the stability property of the normal laws with respect to linear-affine transformations as seen in Sect. 1.7, $(X, X+Y)$ has a normal law with mean 0 and covariance matrix

$$AA^* = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\,.$$

Similarly the vector $Z = (X, \sqrt{2}X)$ is obtained from $X$ through the linear transformation associated to the matrix

$$A = \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}$$

that allows us to conclude that $Z$ is normal, centered, with covariance matrix $AA^*$. Equivalently we might have directly computed the characteristic function of $Z$: if

$\theta = (\theta_1, \theta_2)$, then

$$\mathrm{E}[\mathrm{e}^{i\langle\theta,Z\rangle}] = \mathrm{E}[\mathrm{e}^{i(\theta_1+\sqrt{2}\,\theta_2)X}] = \exp\big(-\frac{1}{2}(\theta_1 + \sqrt{2}\,\theta_2)^2\big)$$
$$= \exp\big(-\frac{1}{2}(\theta_1^2 + 2\theta_2^2 + 2\sqrt{2}\,\theta_1\theta_2)\big)$$

whence we get that it is a normal law with mean 0 and covariance matrix

$$\begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 2 \end{pmatrix}.$$

In particular, the two vectors $(X, X+Y)$ and $(X, \sqrt{2}X)$ have the same marginals (normal centered of variance 1 and 2 respectively) but different joint laws (the covariance matrices are different).

**1.10** First observe that the r.v. $(X_1 - X_2, X_1 + X_2)$ is Gaussian, being obtained from $X = (X_1, X_2)$, which is Gaussian, through the linear transformation associated to the matrix

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

We must at this point just check that the covariance matrix, $\Gamma$, of $(X_1 - X_2, X_1 + X_2)$ is diagonal: this will imply that the two r.v.'s $X_1 - X_2, X_1 + X_2$ are uncorrelated and this, for jointly Gaussian r.v.'s, implies independence. Recalling that $(X_1, X_2)$ has covariance matrix equal to the identity, using (1.13) we find that

$$\Gamma = AA^* = 2I.$$

The same argument applies to the r.v.'s $Y_2 = \frac{1}{2}X_1 - \frac{\sqrt{3}}{2}X_2$ and $Y_2 = \frac{1}{2}X_1 + \frac{\sqrt{3}}{2}X_2$. The vector $Y = (Y_1, Y_2)$ is obtained from $X$ through the linear transformation associated to the matrix

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}. \tag{S.1}$$

Therefore $Y$ has covariance matrix $AA^* = I$ and also in this case $Y_1$ and $Y_2$ are independent. Furthermore, $Y$ has the same law, $N(0, I)$, as $X$.

Taking a closer look at this computation, we have proved something more general: if $X \sim N(0, I)$, then, if $A$ is an orthogonal matrix (i.e. such that $A^{-1} = A^*$), $AX$ also has law $N(0, I)$. The matrix $A$ in (S.1) describes a rotation of the plane by an angle equal to $\frac{\pi}{3}$.

**1.11**

a) Let us compute the law of $e^X$ with the method of the partition function. Let us denote by $\Phi_{\mu,\sigma}$ and $f_{\mu,\sigma}$, respectively, the partition function and the density of an $N(\mu,\sigma^2)$ law; for $y>0$ we have

$$\mathrm{P}(e^X \le y) = \mathrm{P}(X \le \log y) = \Phi_{\mu,\sigma}(\log y)\ .$$

By taking the derivative we obtain the density of $e^X$:

$$g_{\mu,\sigma}(y) = \frac{d}{dy}\,\Phi_{\mu,\sigma}(\log y) = \frac{1}{y}\,f_{\mu,\sigma}(\log y) = \frac{1}{\sqrt{2\pi}\,\sigma y}\exp\Big(-\frac{1}{2\sigma^2}(\log y-\mu)^2\Big)\ .$$

b) If $X\sim N(\mu,\sigma^2)$, then we can write $X=\mu+\sigma Z$, where $Z\sim N(0,1)$. Therefore, using Exercise 1.6, for every $p\ge 0$,

$$\mathrm{E}[(e^X)^p] = \mathrm{E}[e^{p\sigma Z+p\mu}] = e^{p\mu}e^{\frac{1}{2}p^2\sigma^2}.$$

For $p=1$ this gives the mean, which is therefore equal to $e^{\mu}e^{\sigma^2/2}$; as for the variance

$$\mathrm{Var}(e^X) = \mathrm{E}[e^{2X}]-\mathrm{E}[e^X]^2 = e^{2\mu}e^{2\sigma^2}-\big(e^{\mu}e^{\sigma^2/2}\big)^2 = e^{2\mu}e^{\sigma^2}\big(e^{\sigma^2}-1\big)\ .$$

Note that the computations would have been more complicated if we tried to compute the moments by integrating the density of the lognormal law, which leads to the nasty looking integral

$$\int_0^{+\infty} y^p g_{\mu,\sigma}(y)\,dy\ .$$

**1.12** Let us assume first that $X\sim N(0,1)$, then

$$\mathrm{E}[e^{tX^2}] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{tx^2}e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2(1-2t)}\,dx\ .$$

The integral diverges if $t\ge\frac{1}{2}$. If, conversely, $t<\frac{1}{2}$ we recognize in the integrand

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2(1-2t)}\,dx\ ,$$

up to a constant, the density of a normal law with mean 0 and variance $(1-2t)^{-1}$. Therefore the integral is equal to $\sqrt{2\pi}\,(1-2t)^{-1/2}$, hence $\mathrm{E}[e^{tX^2}]=+\infty$ if $t\ge\frac{1}{2}$

and $\mathrm{E}[e^{tX^2}] = (1-2t)^{-1/2}$ if $t < \frac{1}{2}$. Recalling that if $Z \sim N(0,1)$ then $X = \sigma Z \sim N(0,\sigma^2)$, we have $\mathrm{E}[e^{tX^2}] = \mathrm{E}[e^{t\sigma^2 Z^2}]$. In conclusion, if $X =\sim N(0,\sigma^2)$,

$$\mathrm{E}[e^{tZ^2}] = \begin{cases} +\infty & \text{if } t \geq \frac{1}{2\sigma^2} \\ \frac{1}{\sqrt{1-2\sigma^2 t}} & \text{if } t < \frac{1}{2\sigma^2}\,. \end{cases}$$

**1.13** Thanks to Proposition 1.1

$$\mathrm{E}[(xe^{b+\sigma X} - K)^+] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \big(e^{b+\sigma z} - K\big)^+ e^{-z^2/2}\, dz\,.$$

Note that the integrand vanishes if $xe^{b+\sigma z} - K < 0$, i.e. if

$$z \leq \zeta := \frac{1}{\sigma}\Big(\log \frac{K}{x} - b\Big)\cdot$$

Hence, with a few standard changes of variable,

$$\begin{aligned}
\mathrm{E}[(xe^{b+\sigma X} - K)^+] &= \frac{1}{\sqrt{2\pi}} \int_{\zeta}^{+\infty} \big(xe^{b+\sigma z} - K\big) e^{-z^2/2}\, dz \\
&= \frac{x}{\sqrt{2\pi}} \int_{\zeta}^{+\infty} e^{b+\sigma z - z^2/2}\, dz - \frac{1}{\sqrt{2\pi}} \int_{\zeta}^{+\infty} K\, e^{-z^2/2}\, dz \\
&= \frac{xe^{b+\frac{1}{2}\sigma^2}}{\sqrt{2\pi}} \int_{\zeta}^{+\infty} e^{-\frac{1}{2}(z-\sigma)^2}\, dz - K\Phi(-\zeta) \\
&= \frac{xe^{b+\frac{1}{2}\sigma^2}}{\sqrt{2\pi}} \int_{\zeta-\sigma}^{+\infty} e^{-z^2/2}\, dz - K\Phi(-\zeta) \\
&= xe^{b+\frac{1}{2}\sigma^2}\Phi(-\zeta+\sigma) - K\Phi(-\zeta)\,.
\end{aligned}$$

**1.14**

a) It is immediate to compute the characteristic functions of the r.v.'s $X_n$ and their limit as $n \to \infty$: for every $\theta \in \mathbb{R}^m$ we have

$$\phi_{X_n}(\theta) = e^{i\langle b_n,\theta\rangle} e^{-\frac{1}{2}\langle \Gamma_n\theta,\theta\rangle} \underset{n\to\infty}{\to} e^{i\langle b,\theta\rangle} e^{-\frac{1}{2}\langle \Gamma\theta,\theta\rangle}$$

and we recognize on the right-hand side the characteristic function of an $N(b,\Gamma)$-distributed r.v.

b1) $X_1 = \alpha x + Z_1$ has a normal law (it is a linear-affine function of the normal r.v. $Z_1$) of mean $\alpha x$ and variance $\sigma^2$.

$X_2 = \alpha X_1 + Z_2$ is also normal, being the sum of the two r.v.'s $\alpha X_1$ and $Z_2$, which are normal and independent. As

$$\mathrm{E}[X_2] = \mathrm{E}[\alpha X_1] + \underbrace{\mathrm{E}[Z_2]}_{=0} = \alpha^2 x$$

$$\mathrm{Var}(X_2) = \mathrm{Var}(\alpha X_1) + \mathrm{Var}(Z_2) = \alpha^2\sigma^2 + \sigma^2$$

we derive that $X_2 \sim N(\alpha^2 x, (1+\alpha^2)\sigma^2)$. By recurrence

$$X_n \sim N(\alpha^n x, \sigma^2(1+\alpha^2+\dots+\alpha^{2(n-1)}))\ .$$

Indeed let us assume that this relation is true for a value $n$ and let us prove that it holds also for $n+1$. As $X_{n+1} = \alpha X_n + Z_{n+1}$ and the two r.v.'s $X_n$ and $Z_{n+1}$ are independent and both normally distributed, $X_{n+1}$ is also normally distributed. We still have to check the values of the mean and the variance of $X_{n+1}$:

$$\begin{aligned}\mathrm{E}[X_{n+1}] &= \mathrm{E}[\alpha X_n] + \mathrm{E}[Z_{n+1}] = \alpha\cdot\alpha^n x = \alpha^{n+1}x\\ \mathrm{Var}(X_{n+1}) &= \mathrm{Var}(\alpha X_n) + \mathrm{Var}(Z_{n+1})\\ &= \alpha^2\cdot\sigma^2(1+\alpha^2+\dots+\alpha^{2(n-1)}) + \sigma^2\\ &= \sigma^2(1+\alpha^2+\dots+\alpha^{2n})\ .\end{aligned}$$

As $|\alpha| < 1$ we have $\alpha^n x \to 0$ as $n\to\infty$ and

$$\sigma^2(1+\alpha^2+\dots+\alpha^{2(n-1)}) \underset{n\to\infty}{\to} \frac{\sigma^2}{1-\alpha^2}\ \cdot$$

Thanks to a) $(X_n)_n$ converges in law to an $N(0, \frac{\sigma^2}{1-\alpha^2})$-distributed r.v.

b2) The vector $(X_n, X_{n+1})$ is Gaussian as a linear transformation of the vector $(X_n, Z_{n+1})$, which is Gaussian itself, $X_n$ and $Z_{n+1}$ being independent and Gaussian. In order to compute the limit in law we just need to compute the limit of the covariance matrices $\Gamma_n$ (we know already that the means converge to 0). Now

$$\begin{aligned}\mathrm{Cov}(X_n, X_{n+1}) &= \mathrm{Cov}(X_n, \alpha X_n + Z_{n+1}) = \alpha\mathrm{Cov}(X_n, X_n) + \mathrm{Cov}(X_n, Z_{n+1})\\ &= \alpha\sigma^2(1+\alpha^2+\dots+\alpha^{2(n-1)})\ .\end{aligned}$$

As $n\to\infty$ this quantity converges to $\frac{\alpha\sigma^2}{1-\alpha^2}$. We already know the value of the limit of the variances (and therefore of the elements on the diagonal of $\Gamma_n$), and we obtain

$$\lim_{n\to\infty} \Gamma_n = \frac{\sigma^2}{1-\alpha^2}\begin{pmatrix}1 & \alpha\\ \alpha & 1\end{pmatrix}.$$

The limit law is therefore Gaussian, centered and with this covariance matrix.

**1.15**

a) The result is obvious for $p \le 2$ (even without the assumption of Gaussianity), thanks to the inequality between $L^p$ norms (1.5). Let us assume therefore that $p \ge 2$ and let us consider first the case $m = 1$. If $X$ is centered and $E[X^2] = \sigma^2$, we can write $X = \sigma Z$ with $Z \sim N(0, 1)$. Therefore

$$E[|X|^p] = \sigma^p \underbrace{E[|Z|^p]}_{=c_p} = c_p E[|X|^2]^{p/2}.$$

For $m \ge 2$, let $X = (X_1, \dots, X_m)$. Using the result obtained for $m = 1$ and as $\frac{p}{2} \ge 1$, thanks to both the inequalities of the hint,

$$\begin{aligned} E[|X|^p] &\le m^{\frac{p-2}{2}} \sum_{i=1}^m E[|X_i|^p] \le c_p m^{\frac{p-2}{2}} \sum_{i=1}^m E[|X_i|^2]^{p/2}\\ &\le c_p m^{\frac{p-2}{2}} \Big(\sum_{i=1}^m E[|X_i|^2]\Big)^{p/2} = c_p m^{\frac{p-2}{2}} E[|X|^2]^{p/2}. \end{aligned}$$

b) First of all $L^2$-convergence implies $L^1$-convergence and therefore the convergence of the means: if $m_n = E[X_n]$, then $\lim_{n\to\infty} m_n = 0$. Let $\widetilde{X}_n = X_n - m_n$. Then $\widetilde{X}_n$ is centered and also $\widetilde{X}_n \to_{n\to\infty} 0$ in $L^2$. Thanks to a),

$$\begin{aligned} E[|X_n|^p] &= E[|\widetilde{X}_n + m_n|^p] \le 2^{p-1}\big(|m_n|^p + E[|\widetilde{X}_n|^p]\big)\\ &= 2^{p-1}\big(|m_n|^p + c_{p,m} E[|\widetilde{X}_n|^2]^{p/2}\big) \underset{n\to\infty}{\to} 0. \end{aligned}$$

**1.16**

a) For the partition function $F_Y$ of $Y$ we have, $X$ and $Z$ being independent,

$$\begin{aligned} F_Y(t) &= P(Y \le t) = P(XZ \le t) = P(XZ \le t, Z = 1) + P(XZ \le t, Z = -1)\\ &= P(X \le t, Z = 1) + P(-X \le t, Z = -1) = \frac{1}{2} P(X \le t) + \frac{1}{2} P(X \ge -t). \end{aligned}$$

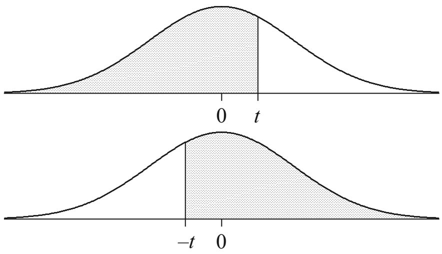


**Fig. S.1** The two shaded surfaces have the same area

But, with an obvious change of variable,

$$P(X \geq -t) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+\infty} e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2}\,dx = P(X \leq t)$$

(see also Fig. S.1, for an intuitive explanation). Therefore, going back to the computation of $F_Y$, $P(Y \leq t) = P(X \leq t)$ for every $t \in \mathbb{R}$ and $X$ and $Y$ have the same law. It is also possible to compute the characteristic function of $Y$: always using the independence of $X$ and $Z$,

$$\begin{aligned} E[e^{i\theta Y}] &= E[e^{i\theta XZ}] = E[e^{i\theta XZ} 1_{\{Z=1\}}] + E[e^{i\theta XZ} 1_{\{Z=-1\}}] \\ &= E[e^{i\theta X} 1_{\{Z=1\}}] + E[e^{-i\theta X} 1_{\{Z=-1\}}] = \frac{1}{2} E[e^{i\theta X}] + \frac{1}{2} E[e^{-i\theta X}] \\ &= \frac{1}{2} e^{-\theta^2/2} + \frac{1}{2} e^{-\theta^2/2} = e^{-\theta^2/2}. \end{aligned}$$

b) Let us compute the characteristic function of $X + Y$: we have $E[e^{i\theta(X+Y)}] = E[e^{i\theta X(1+Z)}]$ and repeating the argument above for the characteristic function of $Y$,

$$E[e^{i\theta(X+Y)}] = E[1_{\{Z=-1\}}] + E[e^{i2\theta X} 1_{\{Z=1\}}] = \frac{1}{2} + \frac{1}{2} E[e^{i2\theta X}] = \frac{1}{2} + \frac{1}{2} e^{-2\theta^2}\,.$$

It is easy to see that this cannot be the characteristic function of a normal r.v.: for instance, note that $X + Y$ has mean 0 and variance $\sigma^2 = 2$ (taking the derivatives of the characteristic function at 0) and if it was Gaussian its characteristic function would be $\theta \mapsto e^{-\theta^2}$. The pair $(X, Y)$ cannot therefore be jointly normal: if it was, then $X + Y$ would also be Gaussian, being a linear function of $(X, Y)$.

**1.17**

a) By the Borel–Cantelli lemma, Proposition 1.7, (1.29) holds if

$$\sum_{n=1}^{\infty} P(X_n \geq (\alpha \log n)^{1/2}) < +\infty\,. \tag{S.2}$$

Thanks to the inequality of the hint (the one on the right-hand side)

$$\mathrm{P}(X_n \geq (\alpha \log n)^{1/2}) = \frac{1}{\sqrt{2\pi}} \int_{(\alpha \log n)^{1/2}}^{+\infty} \mathrm{e}^{-y^2/2}\, dy$$
$$\leq \frac{1}{\sqrt{2\pi\alpha \log n}} \mathrm{e}^{-\frac{1}{2}\alpha \log n} = \frac{1}{\sqrt{2\pi\alpha \log n}\, n^{\alpha/2}} \cdot$$

As $\frac{\alpha}{2} > 1$, the series in (S.2) is summable.

b) Again in view of the Borel–Cantelli lemma we have to investigate the summability of $\mathrm{P}(X_n \geq (2\log n)^{1/2})$. By the other inequality of the hint

$$\mathrm{P}(X_n \geq (2\log n)^{1/2}) \geq \frac{1}{\sqrt{2\pi}}\big((2\log n)^{1/2} + (2\log n)^{-1/2}\big)^{-1} \mathrm{e}^{-\log n} \sim \frac{1}{2n\sqrt{\pi \log n}}$$

which is the term of a divergent series. Hence, by the Borel–Cantelli lemma

$$\mathrm{P}(X_n \geq (2\log n)^{1/2} \text{ infinitely many times}) = 1\,.$$

c) The r.v.'s $(\log n)^{-1/2}X_n$ have zero mean and variance, $(\log n)^{-1}$, tending to 0 as $n \to \infty$. By Chebyshev's inequality, for every $\alpha > 0$,

$$\mathrm{P}\Big(\Big|\frac{X_n}{\sqrt{\log n}}\Big| \leq \alpha\Big) \leq \frac{1}{\alpha^2 \log n} \underset{n\to\infty}{\to} 0\,.$$

Therefore $(\log n)^{-1/2}X_n \to_{n\to\infty} 0$ in probability. On the other hand, by b), we have, with probability 1,

$$\frac{X_n}{\sqrt{\log n}} \geq \sqrt{2} \qquad \text{for infinitely many indices } n$$

and therefore it is not possible for the sequence to converge to 0 a.s.

**1.18** Let us first check the formula for $f = 1_A$, with $A \in \mathscr{E}$. The left-hand side is obviously equal to $\mu_X(A)$.

Note that $1_A(X) = 1_{\{X\in A\}}$: $X(\omega) \in A$ if and only if $\omega \in \{X \in A\}$. Therefore the right-hand side is equal to $\mathrm{P}(X \in A)$ and, by the definition of image probability, (1.7) is true for every function that is the indicator of an event.

By linearity (1.7) is also true for every function which is a linear combination of indicator functions, i.e. for every elementary function.

Let now $f$ be a positive measurable function on $E$. By Proposition 1.11 there exists an increasing sequence $(f_n)_n$ of elementary functions converging to $f$. Applying Beppo Levi's theorem twice we find

$$\int_E f\, d\mu_X = \lim_{n\to\infty} \int_E f_n\, d\mu_X = \lim_{n\to\infty} \int_\Omega f_n(X)\, d\mathrm{P} = \int_\Omega f(X)\, d\mathrm{P} \tag{S.3}$$

so that (1.7) is satisfied for every positive function $f$ and, by taking its decomposition into positive and negative parts, for every measurable function $f$.

**1.19** Let us consider the family, $\mathscr{E}'$ say, of the sets $A \in \mathscr{E}$ such that $X^{-1}(A) \in \mathscr{F}$. Using the relations

$$X^{-1}(A)^c = X^{-1}(A^c), \quad X^{-1}\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \bigcup_{n=1}^{\infty} X^{-1}(A_n)$$

it is easy to check that $\mathscr{E}'$ is a $\sigma$-algebra contained in $\mathscr{E}$. It contains the class $\mathscr{D}$ hence also $\mathscr{E}$. Therefore $X^{-1}(A) \in \mathscr{F}$ for every $A \in \mathscr{E}$ and $X$ is measurable.

**1.20**

a) As $Z_n \to Z$ a.s. we also have $H_n \to Z$ a.s. and, as $0 \le H_n \le Z$, one can apply Lebesgue's theorem.
b) Note that if $Z \ge Z_n$ then $|Z - Z_n| = Z - Z_n = Z - H_n$ whereas $Z_n - H_n = 0$ and the relation is proved. Repeating the argument changing the roles of $Z$ and $Z_n$ we see that the relation also holds if $Z_n \ge Z$. We now have

$$\mathrm{E}[|Z - Z_n|] = \mathrm{E}[Z - H_n] + \mathrm{E}[Z_n - H_n] = 2\mathrm{E}[Z - H_n] \underset{n\to\infty}{\to} 0\,.$$

**1.21**

a) By the Central Limit Theorem the sequence

$$Y_n = \frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n}\,\sigma}$$

converges in law to an $N(0, 1)$-distributed r.v., where $\mu$ and $\sigma^2$ are respectively the mean and the variance of $X_1$. Obviously $\mu = \frac{1}{2}$, whereas

$$\mathrm{E}[X_1^2] = \int_0^1 x^2\,dx = \frac{1}{3}$$

and therefore $\sigma^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$. $W$ is nothing else than $Y_{12}$. We still have to see whether $n = 12$ is a number large enough for $Y_n$ to be approximatively $N(0, 1)$ …

b) We have, integrating by parts,

$$\mathrm{E}[X^4] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^4 \mathrm{e}^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}} \Big( -x^3 \mathrm{e}^{-x^2/2}\Big|_{-\infty}^{+\infty} + 3 \int_{-\infty}^{+\infty} x^2 \mathrm{e}^{-x^2/2}\,dx \Big)$$
$$= 3 \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 \mathrm{e}^{-x^2/2}\,dx}_{=\mathrm{Var}(X)=1} = 3\,.$$

The computation of the moment of order 4 of $W$ is a bit more complicated. If $Z_i = X_i - \frac{1}{2}$, then the r.v.'s $Z_i$ are independent and uniform on $[-\frac{1}{2}, \frac{1}{2}]$ and

$$\mathrm{E}[W^4] = \mathrm{E}[(Z_1 + \cdots + Z_{12})^4] .$$

As $\mathrm{E}[Z_i] = \mathrm{E}[Z_i^3] = 0$ (the $Z_i$'s are symmetric with respect to the origin), the expectation of many terms appearing in the expansion of $(Z_1 + \cdots + Z_{12})^4$ vanishes. For instance, as the r.v.'s $Z_i$, $i = 1, \dots, 12$, are independent,

$$\mathrm{E}[Z_1^3 Z_2] = \mathrm{E}[Z_1^3]\mathrm{E}[Z_2] = 0 .$$

A moment of reflection shows that a non-zero contribution is given only by the terms of the form $\mathrm{E}[Z_i^2 Z_j^2] = \mathrm{E}[Z_i^2]\mathrm{E}[Z_j^2]$ with $i \neq j$ and those of the form $\mathrm{E}(Z_i^4)$. The term $Z_i^4$ clearly has a coefficient $= 1$ in the expansion. In order to determine the coefficient of $Z_i^2 Z_j^2$, $i \neq j$, we can observe that in the power series expansion of the function

$$\phi(x_1, \dots, x_{12}) = (x_1 + \cdots + x_{12})^4$$

the monomial $x_i^2 x_j^2$, for $i \neq j$, has the coefficient

$$\frac{1}{2!2!} \frac{\partial^4 \phi}{\partial x_i^2 \partial x_j^2}(0) = \frac{1}{4} \times 24 = 6 .$$

Let us compute now

$$\mathrm{E}[Z_i^2] = \int_{-1/2}^{1/2} x^2 \, dx = \frac{1}{12}, \qquad \mathrm{E}[Z_i^4] = \int_{-1/2}^{1/2} x^4 \, dx = \frac{1}{80} \cdot$$

As, by symmetry, the terms of the form $\mathrm{E}[Z_i^2 Z_j^2]$, $i \neq j$, are all equal and there are $11 + 10 + \cdots + 1 = \frac{1}{2} \times 12 \times 11$ of them, their contribution is

$$6 \times \frac{1}{2} \times 12 \times 11 \times \frac{1}{144} = \frac{11}{4} \cdot$$

The contribution of the terms of the form $\mathrm{E}[Z_i^4]$ (there are 12 of them), is conversely $\frac{12}{80}$. In conclusion

$$\mathrm{E}[W^4] = \frac{11}{4} + \frac{12}{80} = 2.9 .$$

In practice $W$ turns out to have a law quite close to an $N(0, 1)$. It is possible to compute its density and to draw its graph, which is almost indistinguishable from the graph of the Gaussian density.

However it has some drawbacks: for instance $W$ cannot take values outside the interval $[-6, 6]$ whereas an $N(0, 1)$ can, even if with a very small probability. In practice $W$ can be used as a substitute for the Box–Müller algorithm of Proposition 1.10 for tasks that require a moderate number of random numbers.

**2.1**

a) $X$ and $Y$ are equivalent because two r.v.'s that are a.s. equal have the same law: if $t_1, \dots, t_n \in T$, then

$$\{(X_{t_1}, \dots, X_{t_n}) \neq (Y_{t_1}, \dots, Y_{t_n})\} = \bigcup_{i=1}^n \{X_{t_i} \neq Y_{t_i}\} .$$

These are negligible events, being finite unions of negligible events. Therefore, for every $A \in \mathscr{E}^{\otimes n}$, the two events $\{(X_{t_1}, \dots, X_{t_n}) \in A\}$ and $\{(Y_{t_1}, \dots, Y_{t_n}) \in A\}$ can differ at most by a negligible event and thus have the same probability.

b) As the paths of the two processes are a.s. continuous but for a negligible event, if they coincide at the times of a dense subset $D \subset T$, they necessarily coincide on the whole of $T$. Let $D = \{t_1, t_2, \dots\}$ be a sequence of times which is dense in $T$ ($T \cap \mathbb{Q}$, e.g.). Then

$$\{X_t = Y_t \text{ for every } t\} = \bigcap_{t \in T} \{X_t = Y_t\} = \bigcap_{t_i \in D} \{X_{t_i} = Y_{t_i}\} .$$

As $\mathrm{P}(X_{t_i} = Y_{t_i}) = 1$ for every $i$, it follows that

$$\mathrm{P}\Big(\bigcap_{t_i \in D} \{X_{t_i} = Y_{t_i}\}\Big) = 1$$

so that also $\mathrm{P}(X_t = Y_t \text{ for every } t) = 1$ and the two processes are indistinguishable.

**2.2**

a) First it is clear that $\sigma(X_t, t \le T) \supset \sigma(X_t, t \in D)$: every r.v. $X_s$, $s \in D$, is obviously $\sigma(X_t, t \le T)$-measurable and $\sigma(X_t, t \in D)$ is by definition the smallest $\sigma$-algebra that makes the r.v.'s $X_s$, $s \in D$, measurable.

In order to show the converse inclusion, $\sigma(X_t, t \le T) \subset \sigma(X_t, t \in D)$, we need only show that every r.v. $X_t, t \le T$, is measurable with respect to $\sigma(X_t, t \in D)$. But if $t \le T$ there exists a sequence $(s_n)_n \subset D$ such that $s_n \to t$ as $n \to \infty$. As the process is continuous, $X_{s_n} \to X_t$ and therefore $X_t$ is the limit of $\sigma(X_t, t \in D)$-measurable r.v.'s and is therefore $\sigma(X_t, t \in D)$-measurable itself.

b) The argument of a) can be repeated as is, but now we must choose the sequence $(s_n)_n$ *decreasing* to $t$ and then use the fact that the process is right-continuous. This implies that all the r.v.'s $X_t$ with $t < T$ are $\sigma(X_t, t \in D)$-measurable, but this argument does not apply to $t = T$, as there are no times $s \in D$ larger than $T$.

Therefore if the process is only assumed to be right-continuous, it is necessary to make the additional assumption $T \in D$.

**2.3** By hypothesis, for every $u > 0$, the map $(([0,u] \times \Omega), \mathscr{B}([0,u] \times \mathscr{F}_u) \to (E, \mathscr{E}))$ defined as $(t,\omega) \to X_t(\omega)$ is measurable. Now note that, as the composition of measurable function is also measurable, $(t,\omega) \to \Psi(X_t(\omega))$ is measurable $(([0,u] \times \Omega), \mathscr{B}([0,u] \times \mathscr{F}_u) \to (G, \mathscr{G}))$, i.e. $t \mapsto \psi(X_t)$ is progressively measurable.

**2.4**

a) To say that the sequence $(Z_n(\omega))_n$ does not converge to $Z_\infty(\omega)$ is equivalent to saying that there exists an $m \geq 1$ such that for every $n_0 \geq 1$ there exists an $n \geq n_0$ such that $|Z_n - Z_\infty| \geq \frac{1}{m}$, which is exactly the event (2.2).

b1) Thanks to a) we know that

$$\Big\{\lim_{n\to\infty} X_{t_n} \neq \ell\Big\} = \bigcup_{m=1}^{\infty} \bigcap_{n_0=1}^{\infty} \bigcup_{n \geq n_0} \{|X_{t_n} - \ell| \geq \tfrac{1}{m}\} \in \mathscr{F} \tag{S.4}$$

is negligible. Let us prove that the corresponding event for $\widetilde{X}$ is also negligible. As the r.v.'s $(X_{t_n}, \dots, X_{t_{n+k}})$ and $(\widetilde{X}_{t_n}, \dots, \widetilde{X}_{t_{n+k}})$ have the same distribution, the events (belonging to different probability spaces)

$$\bigcup_{n=n_0}^{k} \{|X_{t_n} - \ell| \geq \tfrac{1}{m}\} \quad \text{and} \quad \bigcup_{n=n_0}^{k} \{|\widetilde{X}_{t_n} - \ell| \geq \tfrac{1}{m}\}$$

have the same probability. As these events are increasing in $k$, we have that

$$\bigcup_{n=n_0}^{\infty} \{|X_{t_n} - \ell| \geq \tfrac{1}{m}\} \quad \text{and} \quad \bigcup_{n=n_0}^{\infty} \{|\widetilde{X}_{t_n} - \ell| \geq \tfrac{1}{m}\} \tag{S.5}$$

also have the same probability. As the event in (S.4) is negligible and observing that the events in (S.5) are decreasing in $m, n_0$, for $m, n_0$ large

$$\mathrm{P}\Big(\bigcup_{n=n_0}^{\infty} \{|X_{t_n} - \ell| \geq \tfrac{1}{m}\}\Big) \leq \varepsilon$$

therefore also

$$\widetilde{\mathrm{P}}\Big(\bigcup_{n=n_0}^{\infty} \{|\widetilde{X}_{t_n} - \ell| \geq \tfrac{1}{m}\}\Big) \leq \varepsilon\,,$$

which allows us to conclude the proof.

b2) A repetition of the arguments of a) allows us to state that the event of the $\widetilde{\omega} \in \widetilde{\Omega}$ such that the limit (2.3) does not exist is

$$\bigcup_{m=1}^{\infty} \bigcap_{k=1}^{\infty} \bigcup_{q \in \mathbb{Q}, |q-t| \le \frac{1}{k}} \{|\widetilde{X}_q - \ell| \ge \tfrac{1}{m}\}$$

and, by a repetition of the argument of b1), this event has the same probability as its analogue for $X$, i.e. 0.

**2.5**

a) We have $\mathrm{E}(X_s X_t) = \mathrm{E}(X_s)\mathrm{E}(X_t) = 0$ for every $s \neq t$, the two r.v.'s being independent and centered. Therefore the function $(s,t) \mapsto \mathrm{E}(X_s X_t)$ vanishes but on the diagonal $s = t$, which is a subset of Lebesgue measure 0 of $[a,b]^2$. Hence the integral in (2.4) vanishes.

b) As $(X_t)_t$ is assumed to be measurable, the map $\omega \mapsto \int_a^b X_s(\omega)\, ds$ is a r.v. (recall Example 2.2). By Fubini's theorem

$$\mathrm{E}\Big(\int_a^b X_s\, ds\Big) = \int_a^b \mathrm{E}[X_s]\, ds = 0\,. \tag{S.6}$$

Also the map $(s,t,\omega) \mapsto X_s(\omega)X_t(\omega)$ is measurable and again by Fubini's theorem and a)

$$\mathrm{E}\Big[\Big(\int_a^b X_s\, ds\Big)^2\Big] = \mathrm{E}\Big[\int_a^b X_s\, ds \int_a^b X_t\, dt\Big] = \int_a^b \int_a^b \mathrm{E}(X_s X_t)\, ds\, dt = 0\,. \tag{S.7}$$

The r.v. $\int_a^b X_s\, ds$, which is centered by (S.6), has variance 0 by (S.7). Hence it is equal to 0 a.s.

c) From b) we have that, for every $a, b \in [0,1]$, $a \le b$, $\int_a^b X_s\, ds = 0$ a.s. Therefore, for almost every $\omega$, the function $t \mapsto X_t(\omega)$ is such that its integral on a subinterval $[a,b]$ vanishes for every $a \le b$ and it is well-known that such a function is necessarily $\equiv 0$ a.e.

Actually the previous argument is not completely correct as, if $\int_a^b X_s(\omega)\, ds = 0$ but for a negligible event, this event $N_{a,b} \subset \Omega$ might depend on $a$ and $b$, whereas in the previous argument we needed a negligible event $N$ such that $\int_a^b X_s(\omega)\, ds = 0$ for every $a, b$. In order to deal with this question just set

$$N = \bigcup_{a,b \in \mathbb{Q}} N_{a,b}\,.$$

$N$ is still negligible, being the union of a countable family of negligible events. For every $\omega \notin N$ therefore

$$\int_a^b X_s(\omega)\, ds = 0, \qquad \text{for every } a, b \in \mathbb{Q}$$

and a function whose integral on all the intervals with rational endpoints vanishes is necessarily $= 0$ almost everywhere.

- To be rigorous, in order to apply Fubini's theorem in (S.7) we should first prove that $s, t, \omega \mapsto X_s(\omega)X_t(\omega)$ is integrable. But this is a consequence of Fubini's theorem itself applied to the positive measurable function $s, t, \omega \mapsto |X_s(\omega)X_t(\omega)|$. Indeed, as $\mathrm{E}[|X_t|] \le \mathrm{E}[X_t^2] = \sqrt{c}$,

$$\begin{aligned}\mathrm{E}\Big[\int_a^b ds \int_a^b |X_s X_t|\, dt\Big] &= \int_a^b \int_a^b \mathrm{E}(|X_s X_t|)\, ds\, dt \\ &= \int_a^b \int_a^b \mathrm{E}(|X_s|)\mathrm{E}(|X_t|)\, ds\, dt \le c(b-a)^2\end{aligned}$$

hence $s, t, \omega \mapsto X_s(\omega)X_t(\omega)$ is integrable.

**2.6**

a) We have

$$\psi_Z^{-1}(A_{\gamma,t,\varepsilon}) = \{\omega; |\gamma_t - Z_t(\omega)| \le \varepsilon\}.$$

Therefore $\psi_Z^{-1}(A_{\gamma,t,\varepsilon}) \in \mathscr{F}$, as this set is the inverse image through $Z_t$ of the closed ball of $\mathbb{R}^m$ centered at $\gamma_t$ and with radius $\varepsilon$.

b) As the paths are continuous,

$$\begin{aligned}\psi_Z^{-1}(\overline{U}_{\gamma,T,\varepsilon}) &= \big\{\omega; |\gamma_t - Z_t(\omega)| \le \varepsilon \text{ for every } t \in [0,T]\big\} \\ &= \big\{\omega; |\gamma_r - Z_r(\omega)| \le \varepsilon \text{ for every } r \in [0,T] \cap \mathbb{Q}\big\} \\ &= \bigcap_{r \in [0,T]\cap\mathbb{Q}} \{\omega; |\gamma_r - Z_r(\omega)| \le \varepsilon\} = \bigcap_{r \in [0,T]\cap\mathbb{Q}} \psi_Z^{-1}(A_{\gamma,r,\varepsilon})\,.\end{aligned}$$

Therefore, thanks to a), $\psi_Z^{-1}(\overline{U}_{\gamma,T,\varepsilon}) \in \mathscr{F}$, as a countable intersection of events of $\mathscr{F}$.

c) Keeping in mind Exercise 1.19 we must prove that the sets (2.5) generate the Borel $\sigma$-algebra $\mathscr{B}(\mathscr{C})$. For every $\gamma \in \mathscr{C}$ the sets $\overline{U}_{\gamma,T,\varepsilon}$ form a basis of neighborhoods of $\gamma$ and, as $\mathscr{C}$ is separable, every open set of $\mathscr{C}$ is the countable union of sets of the form (2.5). Hence the $\sigma$-algebra generated by the sets (2.5) contains the Borel $\sigma$-algebra.

**3.1**

a) With the usual trick of separating the actual position from the increment

$$\mathrm{E}[B_sB_t^2] = \mathrm{E}[B_s(B_t-B_s+B_s)^2] = \mathrm{E}[B_s(B_t-B_s)^2]+2\mathrm{E}[B_s^2(B_t-B_s)]+\mathrm{E}[B_s^3] = 0\,.$$

More easily the clever reader might have argued that $B_sB_t^2$ has the same law as $-B_sB_t^2$ ($(-B_t)_t$ is again a Brownian motion), from which $\mathrm{E}[B_sB_t^2] = 0$ (true even if $t < s$).

b) Again

$$\begin{aligned}\mathrm{E}[B_s^2B_t^2] &= \mathrm{E}[B_s^2(B_t-B_s+B_s)^2]\\ &= \underbrace{\mathrm{E}[B_s^2(B_t-B_s)^2]}_{=s(t-s)} + 2\underbrace{\mathrm{E}[B_s^3(B_t-B_s)]}_{=0} + \underbrace{\mathrm{E}[B_s^4]}_{=3s^2} = s(t-s)+3s^2\,.\end{aligned}$$

c) We know that, if $Z$ denotes an $N(0,1)$-distributed r.v., then $B_s \sim \sqrt{s}\,Z$. Hence, integrating by parts and recalling the expression for the Laplace transform of the Gaussian r.v.'s (Exercise 1.6),

$$\begin{aligned}\mathrm{E}[B_s\mathrm{e}^{B_s}] &= \mathrm{E}[\sqrt{s}\,Z\mathrm{e}^{\sqrt{s}\,Z}] = \frac{\sqrt{s}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} z\mathrm{e}^{\sqrt{s}\,z}\mathrm{e}^{-z^2/2}\,dz\\ &= \underbrace{-\frac{\sqrt{s}}{\sqrt{2\pi}}\mathrm{e}^{\sqrt{s}\,z}\mathrm{e}^{-z^2/2}\Big|_{-\infty}^{+\infty}}_{=0} + \frac{s}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{\sqrt{s}\,z}\mathrm{e}^{-z^2/2}\,dz = s\mathrm{e}^{s/2}\,.\end{aligned}$$

d) We have $\mathrm{E}[B_s\mathrm{e}^{B_t}] = \mathrm{E}[B_s\mathrm{e}^{B_s}\mathrm{e}^{B_t-B_s}]$ and as $B_s$ and $B_t-B_s$ are independent

$$\mathrm{E}[B_s\mathrm{e}^{B_t}] = \mathrm{E}[B_s\mathrm{e}^{B_s}]\mathrm{E}[\mathrm{e}^{B_t-B_s}] = s\mathrm{e}^{s/2}\mathrm{e}^{(t-s)/2} = s\mathrm{e}^{t/2}\,.$$

**3.2**

a) We have

$$\mathrm{E}[1_{\{B_t\le a\}}] = \mathrm{P}(B_t\le a) = \mathrm{P}(\sqrt{t}B_1\le a) = \mathrm{P}\Big(B_1\le\frac{a}{\sqrt{t}}\Big)$$

and therefore

$$\lim_{t\to+\infty}\mathrm{E}[1_{\{B_t\le a\}}] = \mathrm{P}(B_1\le 0) = \frac{1}{2}\,\cdot$$

b)

$$\mathrm{E}[B_t 1_{\{B_t \le a\}}] = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{a} x e^{-\frac{x^2}{2t}}\, dx = -\frac{t}{\sqrt{2\pi t}}\, e^{-\frac{x^2}{2t}} \Big|_{-\infty}^{a}$$
$$= -\frac{\sqrt{t}}{\sqrt{2\pi}}\, e^{-\frac{a^2}{2t}} \underset{t\to+\infty}{\to} -\infty\, .$$

**3.3** Recalling that $B_t \sim \sqrt{t}\, B_1$, we are led to the computation of

$$\lim_{t\to+\infty} \sqrt{t}\, \mathrm{E}[tZ^2 e^{-tZ^2}] = \lim_{t\to+\infty} \frac{t^{3/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-tx^2} e^{-\frac{1}{2}x^2}\, dx$$
$$= \lim_{t\to+\infty} \frac{t^{3/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{1}{2}(2t+1)x^2}\, dx = \lim_{t\to+\infty} \frac{t^{3/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{1}{2\sigma^2}x^2}\, dx\, ,$$

where $Z$ denotes an $N(0,1)$-distributed r.v. and we have set $\sigma^2 = \frac{1}{2t+1}$. With this position we are led back to the expression of the variance of a centered Gaussian r.v.:

$$\cdots = \lim_{t\to+\infty} \sigma t^{3/2} \underbrace{\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x^2 e^{-\frac{1}{2\sigma^2}x^2}\, dx}_{=\sigma^2} = \lim_{t\to+\infty} \sigma^3 t^{3/2} = \lim_{t\to+\infty} \frac{t^{3/2}}{(2t+1)^{3/2}}$$
$$= 2^{-3/2} = \frac{1}{\sqrt{8}}\, \cdot$$

**3.4**

a) As $\{B_{t_m} - B_{t_{m-1}} \in \Gamma_m\}$ is independent of $\mathscr{F}_{t_{m-1}}$ whereas all the other events are $\mathscr{F}_{t_{m-1}}$-measurable,

$$\mathrm{P}(B_{t_m} - B_{t_{m-1}} \in \Gamma_m, \dots, B_{t_1} - B_s \in \Gamma_1, A)$$
$$= \mathrm{P}(B_{t_m} - B_{t_{m-1}} \in \Gamma_m)\mathrm{P}(B_{t_{m-1}} - B_{t_{m-2}} \in \Gamma_{m-1}, \dots, B_{t_1} - B_s \in \Gamma_1, A)\, .$$

Iterating this procedure $m$ times we have

$$\mathrm{P}(B_{t_m} - B_{t_{m-1}} \in \Gamma_m, \dots, B_{t_1} - B_s \in \Gamma_1, A)$$
$$= \mathrm{P}(B_{t_m} - B_{t_{m-1}} \in \Gamma_m)\mathrm{P}(B_{t_{m-1}} - B_{t_{m-2}} \in \Gamma_{m-1}) \dots \mathrm{P}(B_{t_1} - B_s \in \Gamma_1)\mathrm{P}(A)$$
$$= \mathrm{P}(B_{t_m} - B_{t_{m-1}} \in \Gamma_m, \dots, B_{t_1} - B_s \in \Gamma_1)\mathrm{P}(A)\, .$$

b) The r.v.'s $B_{t_m} - B_s, \dots, B_{t_1} - B_s$ are functions of $B_{t_m} - B_{t_{m-1}}$, $B_{t_{m-1}} - B_{t_{m-2}}, \dots$, $B_{t_1} - B_s$, so that they are $\sigma(B_{t_m} - B_{t_{m-1}}, B_{t_{m-1}} - B_{t_{m-2}}, \dots, B_{t_1} - B_s)$-measurable

and therefore

$$\sigma(B_{t_m} - B_s, \dots, B_{t_1} - B_s) \subset \sigma(B_{t_m} - B_{t_{m-1}}, B_{t_{m-1}} - B_{t_{m-2}}, \dots, B_{t_1} - B_s) .$$

But $B_{t_m} - B_{t_{m-1}}, B_{t_{m-1}} - B_{t_{m-2}}, \dots, B_{t_1} - B_s$ are also functions of $B_{t_m} - B_s, \dots, B_{t_1} - B_s$ and by the same argument we have the opposite inclusion.

c) Thanks to a) and b) we have, for $s \le t_1 < \dots < t_m$ and $\Gamma_1, \dots, \Gamma_m \in \mathscr{B}(\mathbb{R}^d)$,

$$\mathrm{P}(B_{t_m} - B_s \in \Gamma_m, \dots, B_{t_1} - B_s \in \Gamma_1, A) = \mathrm{P}(B_{t_m} - B_s \in \Gamma_m, \dots, B_{t_1} - B_s \in \Gamma_1) \cdot \mathrm{P}(A)$$

if $s \le t_1 < \dots < t_m$. But the events

$$\{B_{t_m} - B_s \in \Gamma_m, \dots, B_{t_1} - B_s \in \Gamma_1\}$$

form a class that is stable with respect to finite intersections and generates $\sigma(B_t - B_s, t \ge s))$ and we can conclude the argument using Remark 1.1.

**3.5**

a) It is immediate that $\mathscr{C}$ is stable with respect to finite intersections. Also $\mathscr{C}$ contains $\mathscr{F}_s$ (just choose $G = \Omega$) and $\mathscr{G}$ (choose $A = \Omega$). Therefore the $\sigma$-algebra generated by $\mathscr{C}$ also contains $\widetilde{\mathscr{F}}_s = \mathscr{F}_s \vee \mathscr{G}$ (which, by definition, is the smallest $\sigma$-algebra containing $\mathscr{F}_s$ and $\mathscr{G}$). The converse inclusion $\sigma(\mathscr{C}) \subset \mathscr{F}_s \vee \mathscr{G}$ is obvious.

b) We must prove that, for $s \le t$, $B_t - B_s$ is independent of $\widetilde{\mathscr{F}}_s$. By Remark 1.1 and a) it is enough to prove that, for every Borel set $\Gamma \in \mathscr{B}(\mathbb{R})$ and for every $A \in \mathscr{F}_s, G \in \mathscr{G}$,

$$\mathrm{P}(\{B_t - B_s \in \Gamma\} \cap A \cap G) = \mathrm{P}(B_t - B_s \in \Gamma)\mathrm{P}(A \cap G) .$$

Now $\{B_t - B_s \in \Gamma\} \cap A \in \mathscr{F}_t$ is independent of $G$ ($\mathscr{G}$ and $\mathscr{F}_t$ are independent). Therefore

$$\begin{aligned}\mathrm{P}(\{B_t - B_s \in \Gamma\} \cap A \cap G) &= \mathrm{P}(\{B_t - B_s \in \Gamma\} \cap A)\mathrm{P}(G)\\ = \mathrm{P}(\{B_t - B_s \in \Gamma\})\mathrm{P}(A)\mathrm{P}(G) &= \mathrm{P}(B_t - B_s \in \Gamma)\mathrm{P}(A \cap G)\end{aligned}$$

where we used the fact that $\{B_t - B_s \in \Gamma\}$ and $A$ are independent.

**3.6**

a) The joint law of $(B_s, B_t)$ is a centered Gaussian distribution with covariance matrix

$$C = \begin{pmatrix} s & s \\ s & t \end{pmatrix} .$$

We have

$$C^{-1}=\frac{1}{s(t-s)}\begin{pmatrix} t & -s\\ -s & s\end{pmatrix}$$

so that, defining $z=(x,y)$, the joint density of $(B_s,B_t)$ is

$$\begin{aligned}f_{s,t}(z)=\frac{1}{2\pi\sqrt{s(t-s)}}\,\mathrm{e}^{-\frac12\langle C^{-1}z,z,\rangle}&=\frac{1}{2\pi\sqrt{s(t-s)}}\,\mathrm{e}^{-\frac{1}{2s(t-s)}(tx^2+sy^2-2sxy)}\\ &=\frac{1}{2\pi\sqrt{s(t-s)}}\,\mathrm{e}^{-\frac{1}{2s(t-s)}((t-s)x^2+(sx^2+sy^2-2sxy))}\\ &=\frac{1}{2\pi\sqrt{s(t-s)}}\,\mathrm{e}^{-\frac{1}{2s(t-s)}((t-s)x^2+s(x-y)^2)}\end{aligned}$$

from which (3.19) follows.

b) Of course

$$\begin{aligned}\mathrm{P}(B_s<0,B_{2s}>0)&=\int_{-\infty}^0 dx\int_0^{+\infty} f_{s,2s}(x,y)\,dy\\ &=\int_{-\infty}^0\frac{1}{\sqrt{2\pi s}}\mathrm{e}^{-\frac{1}{2s}x^2}dx\int_0^{+\infty}\frac{1}{\sqrt{2\pi s}}\mathrm{e}^{-\frac{1}{2s}(y-x)^2}\,dy\end{aligned}$$

and with the change of variable $z=x-y$ in the inner integral,

$$\mathrm{P}(B_s<0,B_{2s}>0)=\int_{-\infty}^0\frac{1}{\sqrt{2\pi s}}\mathrm{e}^{-\frac{1}{2s}x^2}dx\int_{-\infty}^x\frac{1}{\sqrt{2\pi s}}\mathrm{e}^{-\frac{1}{2s}z^2}\,dz\,.$$

Let

$$\Phi_s(x)=\frac{1}{\sqrt{2\pi s}}\int_{-\infty}^x \mathrm{e}^{-\frac{1}{2s}z^2}\,dz$$

be the partition function of the $N(0,s)$ distribution, then the previous relation can be written as

$$\mathrm{P}(B_s<0,B_{2s}>0)=\int_{-\infty}^0\Phi_s'(x)\Phi_s(x)\,dx=\frac12\,\Phi_s(x)^2\Big|_{-\infty}^0=\frac18$$

since $\Phi_s(0)=\frac12$ for every $s>0$.

**3.7**

a1) $X$ is a Gaussian process, as the random vector $(X_{t_1}, \dots, X_{t_m})$ is a linear function of the vector $(B_{e^{2t_1}}, \dots, B_{e^{2t_m}})$ which is Gaussian itself. If $s \le t$

$$\mathrm{Cov}(X_t, X_s) = \mathrm{E}[X_t X_s] = e^{-t} e^{-s} \mathrm{E}[B_{e^{2t}} B_{e^{2s}}] = e^{-(t+s)} e^{2s} = e^{-(t-s)} = e^{-|t-s|} .$$

a2) We have, for $1 \le i,j \le m$, $\mathrm{Cov}(X_{t_i+h} X_{t_j+h}) = e^{-|t_i+h-(t_j+h)|} = e^{-|t_i-t_j|} = \mathrm{Cov}(X_{t_i} X_{t_j})$. Therefore the two centered Gaussian random vectors $(X_{t_1+h}, \dots, X_{t_m+h})$ and $(X_{t_1}, \dots, X_{t_m})$ have the same covariance matrix, hence the same law.

a3) $K$, being the covariance kernel of a stochastic process, is necessarily positive definite as explained in Remark 2.5.

b1) A repetition of the argument of a1) gives that $W$ is a centered Gaussian process. Moreover, if $v \le u$,

$$\mathrm{E}[W_u W_v] = \mathrm{E}[\sqrt{u}\, X_{\frac{1}{2}\log u} \sqrt{v}\, X_{\frac{1}{2}\log v}] = \sqrt{uv}\, e^{-\frac{1}{2}(\log u - \log v)} = \sqrt{uv}\,\sqrt{\frac{v}{u}} = v$$

so that $W$ is a natural Brownian motion thanks to Proposition 3.1.

b2) Two possibilities. First one may try to apply Theorem 2.1 (Kolmogorov's continuity theorem): the r.v. $X_t - X_s$ is centered Gaussian with covariance

$$\mathrm{Var}(X_t - X_s) = \mathrm{Var}(X_t) + \mathrm{Var}(X_s) - 2\mathrm{Cov}(X_t, X_s) = 2(1 - e^{-(t-s)}) .$$

Therefore $X_t - X_s \sim \sqrt{2(1 - e^{-(t-s)})}\, Z$, where $Z \sim N(0,1)$. Hence

$$\mathrm{E}[|X_t - X_s|^\beta] = \big(2(1 - e^{-(t-s)})\big)^{\beta/2} \mathrm{E}[|Z|^\beta] .$$

Now, for every $\alpha > 0$, $|1 - e^{-\alpha}| \le \alpha$, as the function $x \mapsto e^{-x}$ has a derivative that is $\le 1$ in absolute value for $x \ge 0$. We have then

$$\mathrm{E}[|X_t - X_s|^\beta] \le c_\beta |t-s|^{\beta/2} = c_\beta |t-s|^{1+(\beta/2-1)} .$$

Hence, by choosing $\beta$ large enough, $X$ has a modification that is Hölder continuous with exponent $\gamma$ for every $\gamma < \frac{1}{2}$, very much similarly as for the Brownian motion.

More quickly just note that $W$, being a Brownian motion, has a continuous modification. If we denote it by $\widetilde{W}$, then $\widetilde{X}_t = e^{-t} \widetilde{W}_{e^{2t}}$ is a continuous modification of $X$.

**3.8**

a) It is immediate that $X$ is a Gaussian process (for every $t_1 < t_2, \dots t_m$ the r.v. $(X_{t_1}, \dots, X_{t_m})$ is a linear function of $(B_{t_1}, \dots, B_{t_m})$ which is Gaussian) and that

$$\mathrm{Cov}(X_1(t), X_2(s)) = \rho(t \wedge s) .$$

Moreover, $X_2$ is obviously a Brownian motion and as

$$\mathrm{Cov}(X_1(t), X_1(s)) = \rho^2(t \wedge s) + (1-\rho^2)(t \wedge s) = t \wedge s$$

the same is true for $X_1$.

b) As $X$ is a Gaussian process, then $B$ is also Gaussian (same argument as in a)). Let $|\rho| < 1$. $B_2$ being already a real Brownian motion we must prove that $B_1$ is a Brownian motion and that $B_1$ and $B_2$ are independent. Now, for $s \le t$,

$$\begin{aligned}
&\mathrm{Cov}(B_1(t), B_1(s))\\
&= \frac{1}{1-\rho^2}\,\mathrm{Cov}(X_1(t), X_1(s)) - \frac{\rho}{1-\rho^2}\,\mathrm{Cov}(X_1(t), X_2(s))\\
&\qquad - \frac{\rho}{1-\rho^2}\,\mathrm{Cov}(X_2(t), X_1(s)) + \frac{\rho^2}{1-\rho^2}\,\mathrm{Cov}(X_2(t), X_2(s))\\
&\qquad = \frac{s}{1-\rho^2}\bigl(1-\rho^2-\rho^2+\rho^2\bigr) = s
\end{aligned}$$

and

$$\mathrm{Cov}(B_1(t), B_2(s)) = \frac{1}{\sqrt{1-\rho^2}}\,\mathrm{Cov}(X_1(t), X_2(s)) - \frac{\rho}{\sqrt{1-\rho^2}}\,\mathrm{Cov}(X_2(t), X_2(s)) = 0\,.$$

- Note that in b) the condition $\rho \ne 1, -1$ is needed.

**3.9** If $X$ is a Brownian motion then $X_t - X_s = A(B_t - B_s)$ must be $N(0, (t-s)I)$-distributed. It is Gaussian, being a linear function of a jointly Gaussian r.v., and its covariance matrix is $C = (t-s)AA^*$ (see (1.13)). We therefore have the condition $AA^* = I$, or $A^* = A^{-1}$. $A$ must therefore be orthogonal. On the other hand, if $A$ is orthogonal, $X$ is a Brownian motion with respect to the filtration $(\mathscr{F}_t)_t$, as $X_t - X_s = A(B_t - B_s)$ is independent of $\mathscr{F}_s$ and a) of Definition 3.2 is immediate.

**3.10** We have

$$\lambda(S_A) = \int_0^{+\infty} 1_{\{B_t \in A\}}\, dt\,.$$

Therefore, with two strokes of Fubini's theorem,

$$\begin{aligned}
\mathrm{E}[\lambda(S_A)] &= \mathrm{E}\Big[\int_0^{+\infty} 1_{\{B_t \in A\}}\, dt\Big] = \int_0^{+\infty} \mathrm{P}(B_t \in A)\, dt\\
&= \int_0^{+\infty} \frac{1}{(2\pi t)^{m/2}}\, dt \int_A e^{-\frac{|x|^2}{2t}}\, dx = \int_A dx \int_0^{+\infty} \frac{1}{(2\pi t)^{m/2}}\, e^{-\frac{|x|^2}{2t}}\, dt\,.
\end{aligned}$$

With the changes of variable $s = \frac{1}{t}$ and subsequently $u = \frac{1}{2}|x|^2 s$,

$$\int_0^{+\infty} \frac{1}{(2\pi t)^{m/2}} \, e^{-\frac{|x|^2}{2t}} \, dt = \frac{1}{(2\pi)^{m/2}} \int_0^{+\infty} s^{-2+\frac{m}{2}} \, e^{-\frac{1}{2}|x|^2 s} \, ds$$

$$= \frac{1}{(2\pi)^{m/2}} \int_0^{+\infty} \Big(\frac{2u}{|x|^2}\Big)^{-2+\frac{m}{2}} \frac{2}{|x|^2} \, e^{-u} \, du = \frac{1}{2\pi^{m/2}} |x|^{2-m} \int_0^{+\infty} u^{-2+\frac{m}{2}} \, e^{-u} \, du \,.$$

The last integral diverges (at 0+) if $-2 + \frac{m}{2} \le -1$, i.e. if $m \le 2$. On the other hand, if $m \ge 3$, it is equal to $\Gamma(\frac{m}{2} - 1)$.

**3.11**

a) If $\gamma \in G$, as the paths of $X$ are continuous, the integral in (3.20) is the limit, for every $\omega$, of its Riemann sums, i.e.

$$X_\gamma = \int X_s \, d\gamma(s) = \lim_{n\to\infty} \underbrace{\sum_{i\ge 0} X_{i/n} \gamma([\tfrac{i}{n}, \tfrac{(i+1)}{n}[)}_{=I_n(\gamma)} \,.$$

Now if $\gamma_1, \dots, \gamma_m \in G$, then the vector $I_n = (I_n(\gamma_1), \dots, I_n(\gamma_m))$ is Gaussian, being a linear function of the r.v.'s $X_{i/n}, i = 0, 1, \dots$, which are jointly Gaussian.

On the other hand $\lim_{n\to\infty} I_n = (X_{\gamma_1}, \dots, X_{\gamma_m})$ a.s.; as a.s. convergence implies convergence in law, by Proposition 1.9 the r.v. $(X_{\gamma_1}, \dots, X_{\gamma_m})$ is Gaussian and therefore $(X_\gamma)_{\gamma\in G}$ is a Gaussian family.

b1) In order to show that $Y$ is a Gaussian process, just observe that we can write

$$Y_t = \int_0^t X_s \, \mu(ds) = \int X_s \, \gamma(ds) \,,$$

where $d\gamma = 1_{[0,t]} \, d\mu$. One can therefore apply what we have already seen in a).

b2) $Y_t$ is centered for every $t$ as, by Fubini's theorem,

$$\mathrm{E}(Y_t) = \mathrm{E}\Big(\int_0^t X_u \, d\mu(u)\Big) = \int_0^t \mathrm{E}(X_u) \, d\mu(u) = 0 \,.$$

If $K_{s,t} = \mathrm{E}(Y_t Y_s)$ denotes its covariance function, then

$$\begin{aligned} K_{s,t} = \mathrm{E}\Big(\int_0^t X_u \, d\mu(u) \cdot \int_0^s X_v \, d\mu(v)\Big) &= \int_0^t d\mu(u) \int_0^s \mathrm{E}(X_u X_v) \, d\mu(v) \\ &= \int_0^t d\mu(u) \int_0^s u \wedge v \, d\mu(v) \,. \end{aligned} \tag{S.8}$$

Therefore

$$\sigma_t^2 = K_{t,t} = \int_0^t d\mu(u) \int_0^t u \wedge v \, d\mu(v)$$
$$= \int_0^t d\mu(u) \int_0^u v \, d\mu(v) + \int_0^t d\mu(u) \int_u^t u \, d\mu(v) = I_1 + I_2 \; .$$

By Fubini's theorem

$$I_2 = \int_0^t d\mu(v) \int_0^v u \, d\mu(u) = I_1 \; .$$

Moreover,

$$\int_0^u v \, d\mu(v) = \int_0^u d\mu(v) \int_0^v dr = \int_0^u dr \int_r^u d\mu(v) = \int_0^u \mu(]r,u]) \, dr$$

and therefore, using (3.21),

$$I_1 = \int_0^t d\mu(u) \int_0^u \mu(]r,u]) \, dr = \int_0^t dr \int_r^t \mu(]r,u]) \, d\mu(u) = \frac{1}{2} \int_0^t \mu(]r,t])^2 \, dr \; ,$$

which concludes the computation of the variance of $Y_t$. In order to compute the covariance $\mathrm{Cov}(Y_s, Y_t) = K_{s,t}$, we have, starting from (S.8) and assuming $s \le t$,

$$K_{s,t} = \int_0^s d\mu(u) \int_0^s u \wedge v \, d\mu(v) + \int_s^t d\mu(u) \int_0^s u \wedge v \, d\mu(v)$$
$$= \int_0^s \mu(]r,s])^2 \, dr + \int_s^t d\mu(u) \int_0^s v \, d\mu(v)$$
$$= \int_0^s \mu(]r,s])^2 \, dr + \mu(]s,t]) \int_0^s \mu(]r,s]) \, dr = \int_0^s \mu(]r,t])\mu(]r,s]) \, dr \; .$$

Only for completeness let us justify (3.21). $\mu(]r,t])^2$ is nothing else than the measure, with respect to $\mu \otimes \mu$, of the square $]r,t] \times ]r,t]$, whereas the integral on the right-hand side is the measure of the shaded triangle in Fig. S.2. The rigorous proof can be done easily with Fubini's theorem.

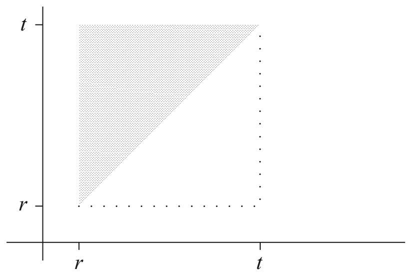


**Fig. S.2** The integration domain in (3.21)

**3.12**

a) The integral that appears in the definition of $Z_t$ is convergent for almost every $\omega$ as, by the Iterated Logarithm Law, with probability 1, $|B_t| \le ((2+\varepsilon)t \log\log \frac{1}{t})^{1/2}$ for $t$ in a neighborhood of 0 and the function $t \mapsto t^{-1/2}(\log\log \frac{1}{t})^{1/2}$ is integrable at $0+$. Let us prove that $Z$ is a Gaussian process, i.e. that, for every choice of $t_1, \dots, t_m$, the r.v. $\widetilde{Z} = (Z_{t_1}, \dots, Z_{t_m})$ is Gaussian. We cannot apply Exercise 3.11 b) immediately because $\frac{1}{u}\,du$ is not a Borel measure (it gives infinite mass to every interval containing 0). But, by Exercise 3.11 a), the r.v. $\widetilde{Z}^{(n)} = (Z^{(n)}_{t_1}, \dots, Z^{(n)}_{t_m})$ is indeed Gaussian, where $Z^{(n)}_{t_i} = 0$ if $t_i < \frac{1}{n}$ and

$$Z^{(n)}_{t_i} = B_{t_i} - \int_{1/n}^{t_i} \frac{B_u}{u}\,du \qquad i = 1, \dots, m$$

otherwise. Then just observe that $\lim_{n\to\infty} \widetilde{Z}^{(n)} = \widetilde{Z}$ a.s. and that a.s. convergence implies convergence in law. We deduce that $\widetilde{Z}$ is Gaussian, by Proposition 1.9. Clearly $Z_0 = 0$ and $Z_t$ is centered. We only need to compute the covariance function of $Z$ and verify that it coincides with that of the Brownian motion. Repeatedly using Fubini's theorem we have for $s \le t$,

$$\mathrm{E}(Z_s Z_t) = \mathrm{E}\Big[\Big(B_s - \int_0^s \frac{B_v}{v}\,dv\Big)\Big(B_t - \int_0^t \frac{B_u}{u}\,du\Big)\Big]$$

$$= \mathrm{E}(B_s B_t) - \int_0^s \frac{\mathrm{E}(B_t B_v)}{v}\,dv - \int_0^t \frac{\mathrm{E}(B_s B_u)}{u}\,du + \int_0^s dv \int_0^t \frac{\mathrm{E}(B_v B_u)}{uv}\,du$$

$$= s - \int_0^s dv - \int_0^s du - \int_s^t \frac{s}{u}\,du + \underbrace{\int_0^s dv \int_0^t \frac{v \wedge u}{uv}\,du}_{=I}$$

$$= -s - s(\log t - \log s) + I\,.$$

Let us compute the double integral:

$$I=\int_0^s dv\int_0^v \frac{1}{v}\,du+\int_0^s dv\int_v^t \frac{1}{u}\,du$$
$$=s+\int_0^s(\log t-\log v)\,dv=s+s\log t-s\log s+s=2s+s(\log t-\log s)$$

so that, finally, $\mathrm{E}(Z_sZ_t)=s=s\wedge t$, which, thanks to Proposition 3.1, completes the proof that $(Z_t)_t$ is a natural Brownian motion.

b) $(Z_t)_t$ is clearly adapted to $(\mathscr{F}_t)_t$ (see Exercise 2.2 a)). In order to show that it is not a Brownian motion with respect to $(\mathscr{F}_t)_t$, there are many possible approaches. For instance, if it was Brownian, then $Z_t-Z_s$ would be independent of $B_s$, that is an $\mathscr{F}_s$-measurable r.v. Instead we have

$$\mathrm{E}[(Z_t-Z_s)B_s]=\mathrm{E}[(B_t-B_s)B_s]-\int_s^t\frac{\mathrm{E}(B_sB_u)}{u}\,du=-\int_s^t\frac{s}{u}\,du=-s\log\frac{t}{s}\neq 0\,.$$

c) Arguing as in a) $Z_s$ and $B_t$ are jointly Gaussian. Hence, in order to prove that $B_t$ is independent of $\mathscr{G}_t$, by Remark 1.1 it is enough to prove that $B_t$ is independent of $Z_s$ for every $s\le t$, i.e. that they are uncorrelated. Actually

$$\mathrm{E}(Z_sB_t)=\mathrm{E}(B_sB_t)-\int_0^s\frac{\mathrm{E}(B_uB_t)}{u}\,du=s-\int_0^s du=0\,.$$

**3.13** We have

$$\{B_t\le a\sqrt{t}\text{ for every }t\le T\}$$
$$=\Big\{\frac{B_t}{\big(2t\log\log\frac{1}{t}\big)^{1/2}}\le\frac{a}{\sqrt{2\log\log\frac{1}{t}}}\text{ for every }t\le T\Big\}$$

but by the Iterated Logarithm Law, with probability 1

$$\overline{\lim_{t\to0+}}\frac{B_t}{\sqrt{2t\log\log\frac{1}{t}}}=1$$

hence, with probability 1 there exists a sequence of times $(t_n)_n$ such that

$$\frac{B_{t_n}}{\sqrt{2t_n\log\log\frac{1}{t_n}}}\ \underset{n\to\infty}{\to}\ 1$$

whereas

$$\frac{a}{\sqrt{2\log\log\frac{1}{t_n}}} \underset{n\to\infty}{\to} 0$$

and therefore the event in question has probability 0.

**3.14**

a) By the Iterated Logarithm Law, see also (3.8), $|B_t| \le (1+\varepsilon)\sqrt{2t\log\log t}$ for $t$ large. Therefore if $b < 0$

$$\lim_{t\to+\infty} X_t = 0$$

whatever the value of $\sigma$. The same arguments give $\lim_{t\to+\infty} X_t = +\infty$ if $b > 0$ (again for every $\sigma$). If $b = 0$ and $\sigma > 0$ we know, by the behavior of the Brownian motion as $t \to +\infty$ as described in Remark 3.4, that $\overline{\lim}_{t\to+\infty} X_t = +\infty$, $\underline{\lim}_{t\to+\infty} X_t = 0$.

b) We have

$$\mathrm{E}[X_t] = \mathrm{e}^{bt}\mathrm{E}[\mathrm{e}^{\sigma B_t}] = \mathrm{e}^{(b+\frac{\sigma^2}{2})t}$$

so that the limit (3.22) is finite if and only if $b \le -\frac{\sigma^2}{2}$ and equal to $+\infty$ otherwise. Observe the apparent contradiction: in the range $b \in ]-\frac{\sigma^2}{2}, 0[$ we have $\lim_{t\to+\infty} X_t = 0$ a.s., but $\lim_{t\to+\infty} \mathrm{E}[X_t] = +\infty$.

**3.15**

a) By the Iterated Logarithm Law $|B_u| \le (1+\varepsilon)\sqrt{2u\log\log u}$ for $t$ large. Hence, if $b > 0$, $\mathrm{e}^{bu+\sigma B_u} \to_{u\to+\infty} +\infty$ (this is also Exercise 3.14 a)) and in this case the integrand itself diverges, hence also the integral. If $b < 0$, conversely, we have, for $t$ large,

$$\mathrm{e}^{bu+\sigma B_u} \le \exp\big(bu + (1+\varepsilon)\sqrt{2u\log\log u}\big) \le \mathrm{e}^{bu/2}$$

and the integral converges to a finite r.v.

b1) The integral can vanish only if the integrand, which is $\ge 0$, vanishes a.s. However, we know by the Iterated Logarithm Law (see Remark 3.4) that the Brownian path takes a.s. strictly positive values in every neighborhood of 0. As the paths are continuous they are therefore strictly positive on a set of times of strictly positive Lebesgue measure a.s.

b2) By a change of variable and using the scaling properties of the Brownian motion, as $v \mapsto \frac{1}{\sqrt{t}} B_{tv}$ is also a Brownian motion,

$$\int_0^t 1_{\{B_u>0\}}\,du = t\int_0^1 1_{\{B_{tv}>0\}}\,dv = t\int_0^1 1_{\{\frac{1}{\sqrt{t}}B_{tv}>0\}}\,dv \overset{\mathscr{L}}{\sim} t\int_0^1 1_{\{B_v>0\}}\,dv\,.$$

Now

$$\lim_{t\to+\infty} t\int_0^1 1_{\{B_u>0\}}\,du = +\infty$$

as we have seen in b1) that the r.v. $\int_0^1 1_{\{B_u>0\}}\,du$ is strictly positive a.s. Hence, as the two r.v.'s $\int_0^t 1_{\{B_u>0\}}\,du$ and $t\int_0^1 1_{\{B_u>0\}}\,du$ have the same distribution for every $t$, we have

$$\lim_{t\to+\infty} \int_0^t 1_{\{B_u>0\}}\,du = +\infty \qquad \text{in probability}\,.$$

In order to prove the a.s. convergence, it suffices to observe that the limit

$$\lim_{t\to+\infty} \int_0^t 1_{\{B_u>0\}}\,du$$

exists a.s. as the integral is an increasing function of $t$. Hence the proof is complete, because the a.s. limit and the limit in probability necessarily coincide.

b3) It suffices to observe that

$$\mathrm{e}^{\sigma B_t} \ge 1_{\{B_t\ge 0\}}$$

and then to apply b2).

c) By Fubini's theorem and recalling the expression of the Laplace transform of the Gaussian distributions,

$$\mathrm{E}\Big[\int_0^{+\infty} \mathrm{e}^{bu+\sigma B_u}\,du\Big] = \int_0^{+\infty} \mathrm{E}[\mathrm{e}^{bu+\sigma B_u}]\,du = \int_0^{+\infty} \mathrm{e}^{(b+\frac{\sigma^2}{2})u}\,du\,.$$

The expectation is therefore finite if and only if $b<-\frac{\sigma^2}{2}$. The integral is then easily computed giving, in conclusion,

$$\mathrm{E}\Big[\int_0^{+\infty} \mathrm{e}^{bu+\sigma B_u}\,du\Big] = \begin{cases} -\frac{1}{b+\frac{\sigma^2}{2}} & \text{if } b<-\frac{\sigma^2}{2}\\ +\infty & \text{otherwise}\,.\end{cases}$$

**3.16** The sequence $(\tau_n)_n$ is increasing and therefore converges to some r.v. $\sigma$. Obviously, as $\tau_n\le\tau$, also $\sigma\le\tau$. Let us prove that actually $\sigma=\tau$.

Let us consider first the event $\{\sigma<+\infty\}$. On it of course $\tau_n<+\infty$ and $X_{\tau_n}\in\partial D_n$. $X$ being continuous, $X_{\tau_n}\to X_\sigma$ and $d(X_\sigma,\partial D)=\lim_{n\to\infty} d(X_{\tau_n},\partial D)=0$, so that $X_\sigma\in\partial D$ and $\sigma\ge\tau$. Conversely, on $\{\sigma=+\infty\}$ there is nothing to prove.

The argument above also implies that $X_{\tau_n}\to X_\tau$ as $n\to\infty$.

Note that this argument only works for an open set $D$ because otherwise the condition $X_\sigma \in \partial D$ does not imply $\sigma \geq \tau$. Actually the statement is not true if $D$ is closed (try to find a counterexample...).

**3.17** If $X_s = \rho B_{s/\rho^2}$, then $X$ is also a Brownian motion by Proposition 3.2 so that, if

$$\tau_\rho^X = \inf\{u; X_u \notin \rho D\} ,$$

we have $\tau_\rho^X \sim \tau_\rho$. As $\rho B_s = X_{s\rho^2}$

$$\begin{aligned}\tau &= \inf\{s; B_s \notin D\} = \inf\{s; \rho B_s \notin \rho D\} = \inf\{s; X_{s\rho^2} \notin \rho D\} \\ &= \tfrac{1}{\rho^2} \inf\{u; X_u \notin \rho D\} \sim \tfrac{1}{\rho^2}\, \tau_\rho .\end{aligned}$$

**3.18**

a) By the Iterated Logarithm Law, with probability 1 there exist values of $t$ such that $X_1(t) \geq (1-\varepsilon)\sqrt{2t \log\log t}$ ($X_1$ is the first component of the Brownian motion $X$). There exists therefore, with probability 1, a time $t$ such that $|X_t| > 1$ and so $\tau < +\infty$ a.s.

   Let us denote by $\mu$ the law of $X_\tau$ and let $O$ be an orthogonal matrix; if $Y_t = OX_t$, then (Exercise 3.9) $Y$ is also a Brownian motion. Moreover, as $|Y_t| = |X_t|$ for every $t$, $\tau$ is also the exit time of $Y$ from $S$. Therefore the law of $Y_\tau = OX_\tau$ coincides with the law of $X_\tau$, i.e. with $\mu$. Therefore the image law of $\mu$ through $O$ (that defines a transformation of the surface of the sphere, $\partial S$, into itself) is still equal to $\mu$.

   This allows us to conclude the proof since, as indicated in the hint, the only probability on $\partial S$ with this property is the normalized $(m-1)$-dimensional Lebesgue measure. Figures S.3 and S.4 show the positions of some simulated exit points.

b) Let $\Gamma$ and $A$ be Borel sets respectively of $\partial S$ and of $\mathbb{R}^+$; we must show that $\mathrm{P}(X_\tau \in \Gamma, \tau \in A) = \mathrm{P}(X_\tau \in \Gamma)\mathrm{P}(\tau \in A)$. Repeating the arguments developed in a), we have, for every orthogonal matrix $O$,

$$\mathrm{P}(X_\tau \in \Gamma, \tau \in A) = \mathrm{P}(X_\tau \in O\Gamma, \tau \in A) .$$

   Therefore, for every fixed $A \in \mathscr{B}(\mathbb{R}^+)$, the measure $\mu_A$ on $\partial S$ defined as $\mu_A(\Gamma) = \mathrm{P}(X_\tau \in \Gamma, \tau \in A)$ is rotationally invariant. It is therefore of the form $\mu_A = c \cdot \lambda$ and obviously the constant $c$ is determined by $c\lambda(\partial S) = \mathrm{P}(\tau \in A)$. Therefore

$$\mathrm{P}(X_\tau \in \Gamma, \tau \in A) = c \cdot \lambda(\Gamma) = \frac{\mathrm{P}(\tau \in A)}{\lambda(\partial S)}\, \lambda(\Gamma) = \mathrm{P}(X_\tau \in \Gamma)\mathrm{P}(\tau \in A) .$$

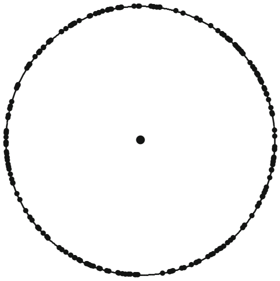

**Fig. S.3** The exit positions of 200 simulated paths of a two-dimensional Brownian motion from the unit ball. The exit distribution appears to be uniform on the boundary

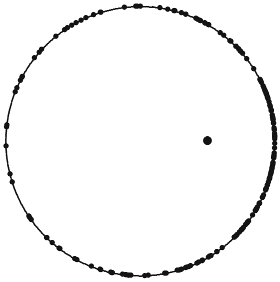

**Fig. S.4** The exit positions of 200 simulated paths from the unit ball for a two-dimensional Brownian motion starting at $(\frac{1}{2}, 0)$ (denoted by a black small circle). Of course the exit distribution does no longer appears to be uniform and seems to be more concentrated on the part of the boundary that is closer to the starting position; wait until Chap. 10 in order to determine this distribution

## 3.19

a) Immediate as

$$\|e^f\|_\infty = \sup_{0\le s\le 1} e^{f(s)} = e^{\sup_{0\le s\le 1} f(s)}\ .$$

b1) The clever reader has certainly sensed the imminent application of the scaling properties of Brownian motion. Replacing in the left-hand side the Brownian motion $B$ with $s \mapsto \sqrt{t}B_{s/t}$ we have, with the substitution $u = s/t$,

$$\int_0^t e^{B_s}\, ds \ \stackrel{\mathcal{L}}{\sim}\ \int_0^t e^{\sqrt{t}B_{s/t}}\, ds = t\int_0^1 e^{\sqrt{t}B_u}\, du\ .$$

b2) The previous relation gives

$$\log \int_0^t e^{B_s}\, ds \ \stackrel{\mathcal{L}}{\sim}\ \log\Big(t\int_0^1 e^{\sqrt{t}B_s}\, ds\Big)$$

and by the property of the $L^p$ norms mentioned in the hint we have, in distribution,

$$\lim_{t\to+\infty} \frac{1}{\sqrt{t}} \log \int_0^t e^{B_s}\,ds = \lim_{t\to+\infty} \frac{1}{\sqrt{t}} \log t + \lim_{t\to+\infty} \log\Big[\Big(\int_0^1 e^{\sqrt{t}B_s}\,ds\Big)^{1/\sqrt{t}}\Big]$$
$$= \log \|e^B\|_\infty = \sup_{0\le s\le 1} B(s)\ .$$

b3) Taking the log and dividing by $\sqrt{t}$

$$\lim_{t\to+\infty} P\Big(\int_0^t e^{B_s}\,ds\le 1\Big)=\lim_{t\to+\infty} P\Big(\frac{1}{\sqrt{t}} \log \int_0^t e^{B_s}\,ds \le 0\Big)=P\Big(\sup_{s\le 1} B_s \le 0\Big) = 0\ .$$

Similarly, by the Reflection Principle,

$$\lim_{t\to+\infty} P\Big(\int_0^t e^{B_s}\,ds \le e^{.77337\cdot\sqrt{t}}\Big) = \lim_{t\to+\infty} P\Big(\frac{1}{\sqrt{t}} \log \int_0^t e^{B_s}\,ds \le .77337\Big)$$
$$= P\Big(\sup_{0\le s\le 1} B_s \le .77337\Big) = 1 - 2P(B_1 > .77337) = \frac{1}{2}\,\cdot$$

**3.20**

a) By the reflection principle the partition function $F_a$ of $\tau_a$ is, for $t > 0$,

$$F_a(t) = P(\tau_a \le t) = P\Big(\sup_{0\le s\le t} B_s > a\Big) = 2P(B_t > a) = 2P(B_1 > at^{-1/2})$$
$$= \frac{2}{\sqrt{2\pi}} \int_{at^{-1/2}}^{+\infty} e^{-x^2/2}\,dx\ .$$

$F_a$ is differentiable, so that $\tau_a$ has density, for $t > 0$,

$$f_a(t) = F_a'(t) = -\sqrt{\frac{2}{\pi}}\, e^{-a^2/2t}\, \frac{d}{dt} \frac{a}{\sqrt{t}} = \frac{a}{\sqrt{2\pi}\, t^{3/2}}\, e^{-a^2/2t}\ .$$

We have

$$E(\sqrt{\tau_a}) = \int_0^{+\infty} \sqrt{t} f_a(t)\,dt = \int_0^{+\infty} \frac{a}{\sqrt{2\pi}\, t}\, e^{-a^2/2t}\,dt = +\infty\ ,$$

as the integrand behaves as $\frac{1}{t}$ towards infinity. A fortiori $\mathrm{E}(\tau_a) = +\infty$. Again by the reflection principle, as $t \to +\infty$,

$$\begin{aligned} \mathrm{P}(\tau_a > t) &= 1 - 2\mathrm{P}(B_t > a) = 1 - 2\mathrm{P}(B_1 > at^{-1/2}) \\ &= 1 - \mathrm{P}(B_1 > at^{-1/2}) - \mathrm{P}(B_1 < -at^{-1/2}) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-at^{-1/2}}^{at^{-1/2}} \mathrm{e}^{-x^2/2}\, dx \sim \frac{1}{\sqrt{2\pi}} \cdot 2at^{-1/2} = \frac{2a}{\sqrt{2\pi t}} \cdot \end{aligned}$$

b) The events $\{T_j \le t\}$, $j = 1, \dots, N$ are independent, each of them having probability

$$\mathrm{P}(T_j \le t) = 2\mathrm{P}(B_t \ge 1) = 2\mathrm{P}\Big(Z \ge \frac{1}{\sqrt{t}}\Big) ,$$

where $Z \sim N(0, 1)$. Hence

$$\begin{aligned} &\mathrm{P}(T_j > t \text{ for at least one index } j, j = 1, \dots, N) \\ &= 1 - \mathrm{P}\Big(\bigcap_{j=1}^{N} \{T_j \le t\}\Big) = 1 - \mathrm{P}(T_j \le t)^N = 1 - \Big(2\mathrm{P}\Big(Z \ge \frac{1}{\sqrt{t}}\Big)\Big)^N . \end{aligned}$$

For $N = 10000$ and $t = 10^8$ we have $2\mathrm{P}\big(Z \ge 10^{-4}\big) = 0.9999202$ and

$$1 - \big(2\mathrm{P}\big(Z \ge 10^{-4}\big)\big)^{10000} = 1 - .55 = 45\%$$

whereas similar computations for $t = 10^{10}$ give

$$1 - \big(2\mathrm{P}\big(Z \ge 10^{-5}\big)\big)^{10000} = .076 = 7.6\% .$$

Therefore the program has the drawback that it can remain stuck on a single path for a very very long time. We shall see some remedies to this problem later on (see Example 12.4).

c) By Theorem 3.3, $\widetilde{B}_t = B_{\tau_a + t} - B_{\tau_a}$ is a Brownian motion independent of $\mathscr{F}_{\tau_a}$. If we denote by $\widetilde{\tau}_a$ the passage time at $a$ of $\widetilde{B}$, the two r.v.'s $\tau_a$ and $\widetilde{\tau}_a$ have the same law and are independent. Moreover, it is clear that $\tau_{2a} = \tau_a + \widetilde{\tau}_a$. By recurrence therefore the sum of $n$ independent r.v.'s $X_1, \dots, X_n$ each having a law equal to that of $\tau_a$ has the same law as $\tau_{na}$. Therefore the density of $\frac{1}{n^2}(X_1 + \dots + X_n)$ is

$$n^2 f_{na}(n^2 t) = n^2 \frac{na}{\sqrt{2\pi}\,(n^2 t)^{3/2}} \exp\Big(-\frac{a^2 n^2}{2tn^2}\Big) = \frac{a}{\sqrt{2\pi}\, t^{3/2}}\, \mathrm{e}^{-a^2/2t} = f_a(t) .$$

**3.21**

a) We have

$$\sup_{0\le s\le t} B_s = \sup_{u\le 1} B_{ut} = \sqrt{t}\,\sup_{u\le 1}\frac{1}{\sqrt{t}}\,B_{ut}\;.$$

As $(\frac{1}{\sqrt{t}}\,B_{ut})_u$ is itself a Brownian motion, this means that

$$\sup_{0\le s\le t} B_s \sim \sqrt{t}\,\sup_{u\le 1} B_u$$

(recall that $\sim$ means equality in law). Of course this implies (3.24).

b) From Exercise 1.3 and thanks to the reflection principle

$$\begin{aligned}
\mathrm{E}\Big[\sup_{0\le s\le t} B_s\Big] &= \int_0^{+\infty} \mathrm{P}\Big(\sup_{0\le s\le t} B_s > x\Big)\,dx = 2\int_0^{+\infty}\mathrm{P}(B_t > x)\,dx\\
&= \frac{2}{\sqrt{2\pi t}}\int_0^{+\infty} dx\int_x^{+\infty}\mathrm{e}^{-\frac{z^2}{2t}}\,dz = \frac{2}{\sqrt{2\pi t}}\int_0^{+\infty}\mathrm{e}^{-\frac{z^2}{2t}}\,dz\int_0^z dx\\
&= \frac{2}{\sqrt{2\pi t}}\int_0^{+\infty} z\mathrm{e}^{-\frac{z^2}{2t}}\,dz = \frac{2}{\sqrt{2\pi t}}\,t\mathrm{e}^{-\frac{z^2}{2t}}\Big|_0^{+\infty} = \sqrt{\frac{2t}{\pi}}\;\cdot
\end{aligned}$$

**3.22**

a1) The vector $\frac{1}{|z|}\,z$ has modulus equal to 1, therefore $X$ is a Brownian motion thanks to Remark 3.1. We have

$$\tau = \inf\{t;\langle z, B_t\rangle \ge k\} = \inf\{t;|z|X_t \ge k\} = \inf\{t; X_t \ge \tfrac{k}{|z|}\}\;.$$

a2) As $\tau$ coincides with the passage time of $X$ at $a=\frac{k}{|z|}$ it has infinite expectation, as seen in Exercise 3.20, and its density is given by (3.23).

b1) $X$ is a centered Gaussian process starting at 0 whatever the value of $\alpha$. In order to check whether it is a Brownian motion we have to determine for which values of $\alpha > 0$ condition 3) of Proposition 3.1 is satisfied. We have, for $s\le t$,

$$\begin{aligned}
\mathrm{E}[X_tX_s] &= \mathrm{E}[\alpha(B_1(t)+B_2(t))\alpha(B_1(s)+B_2(s))]\\
&= \alpha^2\big(\mathrm{E}[B_1(t)B_1(s)]+\mathrm{E}[B_2(t)B_2(s)]+\mathrm{E}[B_1(t)B_2(s)]+\mathrm{E}[B_1(s)B_2(t)]\big)\\
&= \alpha^2\big(s+s+\rho s+\rho s\big) = 2\alpha^2(1+\rho)s
\end{aligned}$$

hence, if $\rho > -1$, we have the condition

$$\alpha = \frac{1}{\sqrt{2(1+\rho)}} \cdot \tag{S.9}$$

If $\rho = -1$ such an $\alpha$ does not exist. Actually, if $\rho = -1$ we have $B_1(t)+B_2(t) = 0$ a.s.

b2) A repetition of the argument of a2): $\tau$ coincides with the passage time of the Brownian motion $X$ with the choice of $\alpha$ as in (S.9) at $a = (2+2\rho)^{-1/2}$ so that $\tau$ has the density (3.23) with $a = (2+2\rho)^{-1/2}$ and has infinite expectation. By the reflection principle (Proposition 3.4)

$$\mathrm{P}(\tau \le 1) = 2\mathrm{P}(X_1 \ge (2+2\rho)^{-1/2}) ,$$

which is maximum for $\rho = 1$.

**3.23**

a) We know that $t \mapsto \langle w, B_t \rangle$ is a Brownian motion if $w$ is a vector having modulus equal to 1 (Example 3.1). Hence $v = \frac{1}{|z|}$ satisfies the requested condition.

b1) Of course

$$\tau = \inf\{t; B_1(t) + B_2(t) = 1\} = \inf\{t; \langle z, B_t \rangle = 1\} .$$

b2) We have $|z|^2 = 2$, so that, thanks to a), $W_t := \frac{1}{\sqrt{2}} \langle z, B_t \rangle$ is a Brownian motion. Hence

$$\tau = \inf\{t; B_1(t) + B_2(t) = 1\} = \inf\{t; \sqrt{2}\, W_t = 1\} = \inf\{t; W_t = \tfrac{1}{\sqrt{2}}\}$$

so that $\tau$ has the same distribution as the passage time of $W$ at $a = \frac{1}{\sqrt{2}}$ and therefore has density

$$f(t) = \frac{a}{(2\pi)^{1/2}\, t^{3/2}} \mathrm{e}^{-a^2/2t} = \frac{1}{2\pi^{1/2}\, t^{3/2}} \mathrm{e}^{-1/4t}$$

and, by the reflection principle and using tables of the partition function $\Phi$ of an $N(0,1)$-distributed r.v.,

$$\mathrm{P}(\tau \le 1) = 2\mathrm{P}\Big(W_1 \ge \frac{1}{\sqrt{2}}\Big) = 0.48 .$$

c1) We have $\tau_1 = \inf\{t; 2B_1(t) + B_2(t) = 1\}$. We can repeat the arguments of b) for $2B_1(t) + B_2(t) = \langle z, B(t) \rangle$ with $z = (2,1)$; here $|z| = \sqrt{5}$ and $2B_1(t) + B_2(t) = \sqrt{5}\, W_1(t)$, where $W_1$ is a Brownian motion. Therefore

$$\tau_1 = \inf\{t; 2B_1(t) + B_2(t) = 1\} = \inf\{t; \sqrt{5}\, W_1(t) = 1\} = \inf\{t; W_1(t) = \tfrac{1}{\sqrt{5}}\}$$

so that $\tau_1$ has density, for $t \geq 0$,

$$f_1(t) = \frac{1}{(10\pi)^{1/2}\, t^{3/2}} \mathrm{e}^{-1/10t} \,.$$

The same argument, for the new Brownian motion $W_2(t) = \frac{1}{\sqrt{5}} B_1(t) - \frac{2}{\sqrt{5}} B_2(t)$, gives

$$\tau_2 = \inf\{t; \tfrac{1}{2} B_1(t) - B_2(t) = 1\} = \inf\{t; \tfrac{\sqrt{5}}{2} W_2(t) = 1\} = \inf\{t; W_2(t) = \tfrac{2}{\sqrt{5}}\}$$

with density

$$f_2(t) = \frac{2}{(10\pi)^{1/2}\, t^{3/2}} \mathrm{e}^{-1/5t} \,.$$

c2) In order to prove that $\tau_1$ and $\tau_2$ are independent, it suffices to show that the two Brownian motions $W_1, W_2$ are independent. We have

$$W_1(t) = \frac{2}{\sqrt{5}} B_1(t) + \frac{1}{\sqrt{5}} B_2(t), \quad W_2(s) = \frac{1}{\sqrt{5}} B_1(s) - \frac{2}{\sqrt{5}} B_2(s) \,.$$

Therefore, assuming $s \leq t$,

$$\mathrm{Cov}(W_1(t), W_2(s)) = \mathrm{E}[W_1(t)W_2(s)] = \frac{2}{\sqrt{5}} \mathrm{E}[B_1(t)B_1(s)] - \frac{2}{\sqrt{5}} \mathrm{E}[B_2(t)B_2(s)] = 0.$$

Hence, as the r.v.'s $W_1(t), W_2(s)$ are jointly Gaussian, they are independent for every $t, s$. Alternatively, just observe that the two-dimensional process $(W_1, W_2)$ is obtained from $(B_1, B_2)$ through the linear transformation associated to the matrix

$$\begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix}$$

which is orthogonal. Hence by Exercise 3.9 $(W_1, W_2)$ is a two-dimensional Brownian motion which implies the independence of $W_1$ and $W_2$.

c3) The event $\{\sigma \leq 1\}$ is the union of the events $\{\tau_1 \leq 1\}$ and $\{\tau_2 \leq 1\}$, which are independent. Therefore

$$\mathrm{P}(\sigma \leq 1) = \mathrm{P}(\tau_1 \leq 1) + \mathrm{P}(\tau_2 \leq 1) - \mathrm{P}(\tau_1 \leq 1)\mathrm{P}(\tau_2 \leq 1) \,.$$

Now, by the reflection principle and again denoting by $\Phi$ the partition function of the $N(0,1)$ distribution,

$$\begin{aligned}\mathrm{P}(\tau_1 \le 1) &= 2\mathrm{P}\Big(W_1 \ge \tfrac{1}{\sqrt{5}}\Big) = 2\Big(1 - \Phi\Big(\tfrac{1}{\sqrt{5}}\Big)\Big) = 2(1-0.67)) = 0.65\,,\\ \mathrm{P}(\tau_2 \le 1) &= 2\mathrm{P}\Big(W_2 \ge \tfrac{2}{\sqrt{5}}\Big) = 2\Big(1 - \Phi\Big(\tfrac{2}{\sqrt{5}}\Big)\Big) = 2(1-0.81)) = 0.37\,,\end{aligned}$$

which gives

$$\mathrm{P}(\sigma \le 1) = 0.65 + 0.37 - 0.65 \cdot 0.37 = 0.78\,.$$

c4) The important thing is to observe that the event $\{\tau_1 \le \tau_2\}$ coincides with the event "the pair $(\tau_1, \tau_2)$ takes its values above the diagonal", i.e., as $\tau_1$ and $\tau_2$ are independent,

$$\mathrm{P}(\tau_1 \le \tau_2) = \int_0^{+\infty} f_2(t)\,dt \int_0^t f_1(s)\,ds\,. \tag{S.10}$$

Let us compute this integral. To simplify the notations let us set $a_1 = \frac{1}{\sqrt{5}}$, $a_2 = \frac{2}{\sqrt{5}}$. By the reflection principle

$$\int_0^t f_1(s)\,ds = \mathrm{P}(\tau_1 \le t) = 2\mathrm{P}(W_1(t) \ge a_1) = \frac{2}{\sqrt{2\pi t}} \int_{a_1}^{+\infty} \mathrm{e}^{-\frac{x^2}{2t}}\,dx$$

and substituting into (S.10)

$$\begin{aligned}\mathrm{P}(\tau_1 \le \tau_2) &= \int_{a_1}^{+\infty} dx \int_0^{+\infty} \frac{a_2}{(2\pi)^{1/2}\,t^{3/2}}\,\mathrm{e}^{-a_2^2/2t}\,\frac{2}{\sqrt{2\pi t}}\,\mathrm{e}^{-x^2/2t}\,dt\\ &= \frac{a_2}{\pi} \int_{a_1}^{+\infty} dx \int_0^{+\infty} \frac{1}{t^2}\,\mathrm{e}^{-\frac{1}{2t}(a_2^2+x^2)}\,dt = \frac{a_2}{\pi}\int_{a_1}^{+\infty} dx \Big(\frac{2}{a_2^2+x^2}\,\mathrm{e}^{-\frac{1}{2t}(a_2^2+x^2)}\Big|_{t=0}^{t=+\infty}\Big)\\ &= \frac{a_2}{\pi}\int_{a_1}^{+\infty} \frac{2}{a_2^2+x^2}\,dx = \frac{2}{\pi a_2}\int_{a_1}^{+\infty} \frac{1}{1+(\frac{x}{a_2})^2}\,dx = \frac{2}{\pi}\arctan\frac{x}{a_2}\Big|_{a_1}^{+\infty}\\ &= \frac{2}{\pi}\Big(\frac{\pi}{2} - \arctan\frac{a_1}{a_2}\Big) = 1 - \frac{2}{\pi}\arctan\frac{1}{2} \simeq 0.7\,.\end{aligned}$$

- Note that the independence of $\tau_1$ and $\tau_2$ is a consequence of the orthogonality of the two straight lines $2x + 1 = 1$ and $\frac{1}{2}\,x - y = 1$.

**4.1** Statement a) looks intuitive: adding the information $\mathscr{D}$, which is independent of $X$ and of $\mathscr{G}$, should not provide any additional information useful for the prediction of $X$. But the formulation of the exercise has certainly led the reader to surmise that things are not really this way. Let us therefore start proving b); we shall then look for a counterexample showing that the answer to a) is negative.

b) The events $G \cap D$, $G \in \mathscr{G}, D \in \mathscr{D}$ form a class which is stable with respect to finite intersections, generating $\mathscr{G} \vee \mathscr{D}$ and containing $\Omega$. Let us prove that

$$\mathrm{E}[\mathrm{E}(X|\mathscr{G})1_{G\cap D}] = \mathrm{E}(X1_{G\cap D})$$

for every $G \in \mathscr{G}, D \in \mathscr{D}$. As the r.v. $\mathrm{E}[X|\mathscr{G}]$ is obviously $\mathscr{G} \vee \mathscr{D}$-measurable, by Remark 4.2 this will prove that $\mathrm{E}(X|\mathscr{G}) = \mathrm{E}(X|\mathscr{G} \vee \mathscr{D})$. As $D$ is independent of $\sigma(X) \vee \mathscr{G}$ and therefore also of $\mathscr{G}$,

$$\begin{aligned}\mathrm{E}[\mathrm{E}(X|\mathscr{G})1_{G\cap D}] &= \mathrm{E}[\mathrm{E}(X1_G|\mathscr{G})1_D] = \mathrm{E}[\mathrm{E}(X1_G|\mathscr{G})]\mathrm{E}[1_D]\\ &= \mathrm{E}[X1_G]\mathrm{E}[1_D] \underset{\uparrow}{=} \mathrm{E}(X1_G1_D) = \mathrm{E}(X1_{G\cap D})\;,\end{aligned}$$

where $\uparrow$ denotes the place where independence of $D$ and $\sigma(X) \vee \mathscr{G}$ is used.

a) The counterexample is based on the fact that it is possible to construct three r.v.'s $X, Y, Z$ such that the pairs $(X, Y)$, $(Y, Z)$ and $(Z, X)$ are each formed by independent r.v.'s but such that $X, Y, Z$ are not independent globally.

An example is given by $\Omega = \{1, 2, 3, 4\}$, with the uniform probability $\mathrm{P}(k) = \frac{1}{4}$, $k = 1, \dots, 4$, and the $\sigma$-algebra $\mathscr{F}$ of all subsets of $\Omega$. Let $X = 1_{\{1,2\}}$, $Y = 1_{\{2,4\}}$ and $Z = 1_{\{3,4\}}$. Then we have

$$\mathrm{P}(X = 1, Y = 1) = \mathrm{P}(\{1, 2\} \cap \{2, 4\}) = \mathrm{P}(2) = \frac{1}{4} = \mathrm{P}(X = 1)\mathrm{P}(Y = 1)\;.$$

Similarly it can be shown that $\mathrm{P}(X = i, Y = j) = \mathrm{P}(X = i)\mathrm{P}(Y = j)$ for every possible value of $i, j \in \{0, 1\}$, which implies that $X$ and $Y$ are independent. Similarly we prove that $(X, Z)$ and $(Y, Z)$ are also pairs of independent r.v.'s. Let $\mathscr{G} = \sigma(Y)$ and $\mathscr{D} = \sigma(Z)$. Then the $\sigma$-algebra $\mathscr{G} \vee \mathscr{D}$ contains the events

$$\begin{aligned}\{1\} &= \{Y = 0, Z = 0\}, &\qquad \{2\} &= \{Y = 1, Z = 0\}\;,\\ \{3\} &= \{Y = 0, Z = 1\}, &\qquad \{4\} &= \{Y = 1, Z = 1\}\end{aligned}$$

and therefore $\mathscr{G} \vee \mathscr{D} = \mathscr{F}$. Hence $\mathrm{E}[X|\mathscr{G} \vee \mathscr{D}] = X$, whereas $\mathrm{E}[X|\mathscr{G}] = \mathrm{E}(X) = \frac{1}{2}$ a.s. as $X$ and $\mathscr{G}$ are independent.

**4.2** Let us first assume that $Y$ is integrable. Then if $Y$ were $\mathscr{G}$-measurable we would have

$$Y = \mathrm{E}[Y|\mathscr{G}] \qquad \text{a.s.}$$

But on the other hand, as $Y$ is independent of $\mathscr{G}$,

$$\mathrm{E}[Y|\mathscr{G}] = \mathrm{E}[Y] \qquad \text{a.s.}$$

and these two relations can both be true only if $Y = \mathrm{E}[Y]$ a.s. If $Y$ is not integrable let us approximate it with integrable r.v.'s. If $Y_n = Y \vee (-n) \wedge n$ then $Y_n$ is still independent of $\mathscr{G}$ and $\mathscr{G}$-measurable. Moreover, as $|Y_n| \le n$, $Y_n$ is integrable. Therefore by the first part of the proof $Y_n$ is necessarily a.s. constant and, taking the limit as $n \to \infty$, the same must hold for $Y$.

**4.3** Recall (p. 88) that $f(x) = \mathrm{E}[Z|X = x]$ means that $f(X) = \mathrm{E}[Z|\sigma(X)]$ a.s., i.e. that $\mathrm{E}[f(X)\psi(X)] = \mathrm{E}[Z\psi(X)]$ for every bounded measurable function $\psi : E \to \mathbb{R}$. Let $A \in \mathscr{E}$. Then we have $\{X \in A\} \in \sigma(X)$ and

$$\mu_{\mathrm{Q}}(A) = \mathrm{Q}(X \in A) = \mathrm{E}[Z1_{\{X\in A\}}] = \mathrm{E}[f(X)1_{\{X\in A\}}] = \int_A f(y)\, d\mu_{\mathrm{P}}(y)\;,$$

which proves simultaneously that $\mu_{\mathrm{Q}} \ll \mu_{\mathrm{P}}$ and that $\frac{d\mu_{\mathrm{Q}}}{d\mu_{\mathrm{P}}} = f$.

**4.4**

a) If $G = \{\mathrm{E}(Z|\mathscr{G}) = 0\}$, then $G \in \mathscr{G}$ and therefore, as $\mathrm{E}(Z|\mathscr{G})1_G = 0$,

$$\mathrm{E}(Z1_G) = \mathrm{E}[\mathrm{E}(Z1_G|\mathscr{G})] = \mathrm{E}[\mathrm{E}(Z|\mathscr{G})1_G] = 0\;. \tag{S.11}$$

As $Z \ge 0$, this implies $Z = 0$ a.s. on $G$, i.e. $\{Z = 0\} \supset \{\mathrm{E}(Z|\mathscr{G}) = 0\}$ a.s. Taking the complements, $\{Z > 0\} \subset \{\mathrm{E}(Z|\mathscr{G}) > 0\}$ and $1_{\{Z>0\}} \le 1_{\{\mathrm{E}(Z|\mathscr{G})>0\}}$. Hence if $Y \ge 0$ we have a.s.

$$\mathrm{E}[ZY|\mathscr{G}] = \mathrm{E}[Z1_{\{Z>0\}}Y|\mathscr{G}] \le \mathrm{E}[Z1_{\{\mathrm{E}(Z|\mathscr{G})>0\}}Y|\mathscr{G}] = 1_{\{\mathrm{E}(Z|\mathscr{G})>0\}}\mathrm{E}[ZY|\mathscr{G}]\;.$$

As the opposite inequality is obvious, (4.28) is proved for $Y \ge 0$. We obtain the case of a general r.v. $Y$ splitting into positive and negative parts.

b) We have, with a repetition of the argument of (S.11),

$$\mathrm{Q}(\mathrm{E}[Z|\mathscr{G}] = 0) = \mathrm{E}[Z1_{\{\mathrm{E}[Z|\mathscr{G}]=0\}}] = 0\;,$$

hence $\mathrm{E}[Z|\mathscr{G}] > 0$ Q-a.s. Moreover, as the right-hand side of (4.29) is clearly $\mathscr{G}$-measurable, in order to verify (4.29) we must prove that, for every $\mathscr{G}$-measurable bounded r.v. $W$,

$$\mathrm{E}^{\mathrm{Q}}\Big[W\,\frac{\mathrm{E}(YZ|\mathscr{G})}{\mathrm{E}(Z|\mathscr{G})}\Big] = \mathrm{E}^{\mathrm{Q}}[YW]\;.$$

But

$$\mathrm{E}^{\mathrm{Q}}\Big[W\,\frac{\mathrm{E}(YZ|\mathscr{G})}{\mathrm{E}(Z|\mathscr{G})}\Big] = \mathrm{E}\Big[ZW\,\frac{\mathrm{E}(YZ|\mathscr{G})}{\mathrm{E}(Z|\mathscr{G})}\Big]$$

and, as inside the expectation on the right-hand side $Z$ is the only r.v. that is not $\mathscr{G}$-measurable,

$$\cdots = \mathrm{E}\Big[\mathrm{E}(Z|\mathscr{G})W\,\frac{\mathrm{E}(YZ|\mathscr{G})}{\mathrm{E}(Z|\mathscr{G})}\Big] = \mathrm{E}[\mathrm{E}(YZ|\mathscr{G})W] = \mathrm{E}[YZW] = \mathrm{E}^{\mathrm{Q}}[YW]$$

and (4.29) is satisfied.

- In the solution of Exercise 4.4 we left in the background a delicate point which deserves some attention. Always remember that a conditional expectation (with respect to a probability P) is *not* a r.v., but a family of r.v.'s, only differing from each other by P-negligible events. Therefore the quantity $\mathrm{E}[Z|\mathscr{G}]$ must be considered with care when arguing with respect to a probability Q different from P, as it might happen that a P-negligible event is not Q-negligible. In the case of this exercise there are no difficulties as $\mathrm{P} \gg \mathrm{Q}$, so that negligible events for P are also negligible for Q.

**4.5** Let us prove that every $\mathscr{D}$-measurable real r.v. $W$ is independent of $X$. The characteristic function of the pair $Z = (X, W)$, computed at $\theta = (\lambda, t)$, $\lambda \in \mathbb{R}^m, t \in \mathbb{R}$, is equal to

$$\mathrm{E}[\mathrm{e}^{i\langle\theta,Z\rangle}] = \mathrm{E}[\mathrm{e}^{i\langle\lambda,X\rangle}\mathrm{e}^{itW}] = \mathrm{E}[\mathrm{e}^{itW}\mathrm{E}(\mathrm{e}^{i\langle\lambda,X\rangle}\,|\,\mathscr{D})] = \mathrm{E}[\mathrm{e}^{itW}]\mathrm{E}[\mathrm{e}^{i\langle\lambda,X\rangle}]$$

so that $X$ and $W$ are independent by criterion 7 of Sect. 1.6. This entails the independence of $X$ and $\mathscr{D}$.

**4.6**

a) Thanks to Example 4.5 and particularly (4.12), the requested characteristic function is

$$\mathrm{E}[\mathrm{e}^{i\langle\theta,B_\xi\rangle}] = \lambda\int_0^{+\infty} \mathrm{e}^{-\frac{t}{2}\,|\theta|^2}\mathrm{e}^{-\lambda t}\,dt = \frac{\lambda}{\lambda+\frac{1}{2}\,|\theta|^2} = \frac{2\lambda}{2\lambda+|\theta|^2}\,\cdot$$

b1) The characteristic function of $X$ is

$$\phi_X(\theta) = \frac{\mu}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{i\theta x}\mathrm{e}^{-\mu|x|}\,dx\,.$$

Now $x \mapsto \sin(\theta x)\,\mathrm{e}^{-\mu|x|}$ is an odd function so that the imaginary part in the integral above vanishes. Conversely, $x \mapsto \cos(\theta x)\,\mathrm{e}^{-\mu|x|}$ is an even function so that

$$\begin{aligned}\phi_X(\theta) = \mu\int_0^{+\infty}\cos(\theta x)\,\mathrm{e}^{-\mu x}\,dx &= \Re\Big(\mu\int_0^{+\infty}\mathrm{e}^{i\theta x}\mathrm{e}^{-\mu x}\,dx\Big)\\ &= \Re\Big(\frac{\mu}{\mu-i\theta}\Big) = \frac{\mu^2}{\mu^2+\theta^2}\,\cdot\end{aligned}$$

b2) Comparing the characteristic function of $B_\zeta$ computed in a) with that of a Laplace distribution computed in b2) we see that if $m = 1$ then $B_\zeta$ has a Laplace distribution with parameter $\sqrt{2\lambda}$.

**4.7** There are two possible methods: the best is the second one below...

a) First method. Let us check directly that $X$ has the same finite-dimensional distributions as a Brownian motion. Let $t_1 < t_2 < \dots < t_m$, $\theta = (\theta_1, \dots, \theta_m) \in \mathbb{R}^m$.

Thanks to the freezing lemma,

$$\begin{aligned}\mathrm{E}[\mathrm{e}^{i\theta_1 X_{t_1}+\dots+i\theta_m X_{t_m}}] &= \mathrm{E}[\mathrm{e}^{i\theta_1(B_{\zeta+t_1}-B_\zeta)+\dots+i\theta_m(B_{\zeta+t_m}-B_\zeta)}]\\ &= \mathrm{E}\big[\mathrm{E}[\mathrm{e}^{i\theta_1(B_{\zeta+t_1}-B_\zeta)+\dots+i\theta_m(B_{\zeta+t_m}-B_\zeta)}|\sigma(\zeta)]\big] = \mathrm{E}[\Phi(\zeta)]\ ,\end{aligned}$$

where

$$\Phi(s) = \mathrm{E}[\mathrm{e}^{i\theta_1(B_{s+t_1}-B_s)+\dots+i\theta_m(B_{s+t_m}-B_s)}]\ .$$

As we know that the increments $(B_{s+t} - B_s)_t$ form a Brownian motion (Proposition 3.2), $\Phi$ does not depend on $s$ and we have that

$$\mathrm{E}[\mathrm{e}^{i\theta_1 X_{t_1}+\dots+i\theta_m X_{t_m}}] = \mathrm{E}[\mathrm{e}^{i\theta_1 B_{t_1}+\dots+i\theta_m B_{t_m}}]\ .$$

The process $X$ has the same finite-dimensional distributions as $B$ so that it is itself a *natural* Brownian motion.

b) Second method. Let $\widetilde{\mathscr{F}}_s = \sigma(\mathscr{F}_s, \sigma(\zeta))$ be the filtration that is obtained by adding to $(\mathscr{F}_t)_t$ the $\sigma$-algebra generated by $\zeta$. Thanks to Exercise 3.5, $B$ is also a Brownian motion with respect to $(\widetilde{\mathscr{F}}_t)_t$ and now $\zeta$ is a stopping time for this larger filtration as $\{\zeta \le t\} \in \sigma(\zeta)$ and therefore $\{\zeta \le t\} \in \widetilde{\mathscr{F}}_t$ for every $t$. Then the stopping theorem, Theorem 3.3, allows us to conclude that $X$ is an $(\widetilde{\mathscr{F}}_t)_t$-Brownian motion.

**4.8**

a) Recall that the $\sigma$-algebras $\mathscr{G}_1 = \sigma(B_1(u), u \ge 0)$ and $\mathscr{G}_2 = \sigma(B_2(u), u \ge 0)$ are independent (Remark 3.2 b)) and note that $\tau$ is $\mathscr{G}_2$-measurable.

b) Recalling Remark 4.5 and particularly (4.11), the density of $B_1(\tau)$ is given by

$$g(x) = \int_0^{+\infty} \frac{1}{\sqrt{2\pi t}}\,\mathrm{e}^{-x^2/2t}\,d\nu(t)\ ,$$

where $\nu$ denotes the law of $\tau$. From Exercise 3.20 we know that $\tau$ has density

$$f(t) = \frac{a}{\sqrt{2\pi}\,t^{3/2}}\,\mathrm{e}^{-a^2/2t}$$

for $t > 0$ hence the density of $B_1(\tau)$ is

$$g(x) = \int_0^{+\infty} \frac{1}{\sqrt{2\pi t}} \frac{a}{\sqrt{2\pi}\, t^{3/2}}\, e^{-a^2/2t} e^{-x^2/2t}\, dt = \frac{a}{2\pi} \int_0^{+\infty} \frac{1}{t^2}\, e^{-(a^2+x^2)/2t}\, dt\,.$$

With the change of variable $s = \frac{1}{t}$ we obtain

$$g(x) = \frac{a}{2\pi} \int_0^{+\infty} e^{-\frac{1}{2}(a^2+x^2)s}\, ds = \frac{a}{\pi(a^2+x^2)}\,,$$

which is a *Cauchy law*.

**4.9** The idea is always to split $B_u$ into the sum of $B_s$ and of the increment $B_u - B_s$. As $B_u^2 = (B_u - B_s + B_s)^2 = (B_u - B_s)^2 + B_s^2 + 2B_s(B_u - B_s)$, we have

$$\int_s^t B_u^2\, du = (t-s)B_s + \int_s^t (B_u - B_s)^2\, du + 2B_s \int_s^t (B_u - B_s)\, du\,.$$

$B_s$ is $\mathscr{F}_s$-measurable whereas (recall Remark 4.5)

$$\mathrm{E}\Big(\int_s^t (B_u - B_s)^2\, du \,\Big|\, \mathscr{F}_s\Big) = \int_s^t \mathrm{E}\big[(B_u - B_s)^2 \,\big|\, \mathscr{F}_s\big]\, du = \int_s^t (u-s)\, du = \frac{1}{2}\,(t-s)^2\,,$$

the r.v.'s $B_u - B_s$ being independent of $\mathscr{F}_s$. By the same argument

$$\mathrm{E}\Big(2B_s \int_s^t (B_u - B_s)\, du \,\Big|\, \mathscr{F}_s\Big) = 2B_s \int_s^t \mathrm{E}\big(B_u - B_s \,\big|\, \mathscr{F}_s\big)\, du = 0$$

so that finally

$$\mathrm{E}\Big(\int_s^t B_u^2\, du \,\Big|\, \mathscr{F}_s\Big) = (t-s)B_s + \frac{1}{2}\,(t-s)^2\,.$$

As this quantity is already $\sigma(B_s)$-measurable,

$$\mathrm{E}\Big(\int_s^t B_u^2\, du | B_s\Big) = \mathrm{E}\Big[\mathrm{E}\Big(\int_s^t B_u^2\, du \,\Big|\, \mathscr{F}_s\Big) \,\Big|\, B_s\Big] = (t-s)B_s + \frac{1}{2}\,(t-s)^2\,.$$

The meaning of the equality between these two conditional expectations will become clearer in the light of the Markov property in Chap. 6.

**4.10** Let us denote by $\nu_Y$, $\mu_Z$, respectively, the laws of $Y$ and $Z$ and by $\mu_y$ the law of $\phi(y) + Z$. We must prove that for every pair of bounded measurable functions

$f : E \to \mathbb{R}$ and $g : G \to \mathbb{R}$ we have

$$\mathrm{E}[f(X)g(Y)] = \int_G g(y)\, d\nu_Y(y) \int_E f(x)\, d\mu_y(x)\ .$$

But we have

$$\int_E f(x)\, d\mu_y(x) = \int_E f(\phi(y) + z)\, d\mu_Z(z)$$

and

$$\begin{aligned}\mathrm{E}[f(X)g(Y)] = \mathrm{E}[f(\phi(Y) + Z)g(Y)] &= \int_G g(y)\, d\nu_Y(y) \int_E f(\phi(y) + z)\, d\mu_Z(z)\\ &= \int_G g(y)\, d\nu_Y(y) \int_E f(x)\, d\mu_y(x)\ .\end{aligned}$$

**4.11**

a) We must find a function of the observation, $Y$, that is a good approximation of $X$. We know (see Remark 4.3) that the r.v. $\phi(Y)$ minimizing the squared $L^2$ distance $\mathrm{E}[(\phi(Y) - X)^2]$ is the conditional expectation $\phi(Y) = \mathrm{E}(X|Y)$. Therefore, if we measure the quality of the approximation of $X$ by $\phi(Y)$ in the $L^2$ norm, the best approximation of $X$ with a function of $Y$ is $\mathrm{E}(X|Y)$. Let us go back to formulas (4.23) and (4.24) concerning the mean and the variance of the conditional laws of a Gaussian r.v.'s: here

$$m_X = m_Y = 0, \quad \mathrm{Cov}(X, Y) = \mathrm{Cov}(X + W, X) = 1, \quad \mathrm{Var}(Y) = 1 + \sigma^2\ ,$$

so that

$$\mathrm{E}(X|Y = y) = m_X + \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(Y)}\,(y - m_Y) = \frac{y}{1 + \sigma^2}\ \cdot$$

The required best approximation is

$$\frac{Y}{1 + \sigma^2}\ \cdot$$

Note that the variance of the conditional distribution of $X$ given $Y = y$ is

$$\mathrm{Var}(X) - \frac{\mathrm{Cov}(X, Y)^2}{\mathrm{Var}(Y)} = 1 - \frac{1}{1 + \sigma^2} = \frac{\sigma^2}{1 + \sigma^2}\ \cdot \tag{S.12}$$

b) The computation follows the same line of reasoning as in a) but now $Y$ is two-dimensional and we shall use the more complicated relations (4.21) and (4.22).

If $Y = (Y_1, Y_2)$, $y = (y_1, y_2)$, we have $m_X = 0$, $m_Y = 0$, $C_X = 1$. As $\mathrm{Var}(Y_1) = \mathrm{Var}(Y_2) = 1 + \sigma^2$ and $\mathrm{Cov}(Y_1, Y_2) = \mathrm{Cov}(X + W_1, X + W_2) = 1$,

$$C_Y = \begin{pmatrix} 1+\sigma^2 & 1 \\ 1 & 1+\sigma^2 \end{pmatrix}, \qquad C_{X,Y} = \begin{pmatrix} 1 & 1 \end{pmatrix} .$$

A quick computation gives

$$C_Y^{-1} = \frac{1}{(1+\sigma^2)^2 - 1} \begin{pmatrix} 1+\sigma^2 & -1 \\ -1 & 1+\sigma^2 \end{pmatrix} .$$

Therefore the conditional law of $X$ given $Y = y$ has mean

$$C_{X,Y} C_Y^{-1} y = \frac{1}{2\sigma^2 + \sigma^4} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} (1+\sigma^2)y_1 - y_2 \\ (1+\sigma^2)y_2 - y_1 \end{pmatrix} = \frac{y_1 + y_2}{2 + \sigma^2} \cdot$$

The best approximation of $X$ given $Y = y$ now is

$$\frac{Y_1 + Y_2}{2 + \sigma^2} \cdot$$

The variance of the conditional law is now

$$\begin{aligned} 1 - C_{X,Y} C_Y^{-1} C_{X,Y}^* &= 1 - \frac{1}{2\sigma^2 + \sigma^4} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1+\sigma^2 & -1 \\ -1 & 1+\sigma^2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= 1 - \frac{1}{2\sigma^2 + \sigma^4} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \sigma^2 \\ \sigma^2 \end{pmatrix} = 1 - \frac{2}{2+\sigma^2} = \frac{\sigma^2}{2+\sigma^2} \\ &, \end{aligned}$$

which is smaller than the value of the conditional variance given a single observation, as computed in (S.12).

**4.12**

a) We apply formulas (4.21) and (4.22) to $X = (B_{t_1}, \dots, B_{t_m})$, $Y = B_1$. We have $C_Y = 1$ whereas, as $\mathrm{Cov}(B_{t_i}, B_1) = t_i \wedge 1 = t_i$,

$$C_{X,Y} = \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} .$$

Therefore as

$$C_{X,Y}C_Y^{-1}C_{X,Y}^* = \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix} \begin{pmatrix} t_1 & \dots & t_m \end{pmatrix}$$

the matrix $C_{X,Y}C_Y^{-1}C_{X,Y}^*$ has $t_it_j$ as its $(i,j)$-th entry. As $C_X$ is the matrix $(t_i \wedge t_j)_{i,j}$, the conditional law is Gaussian with covariance matrix $C_X - C_{X,Y}C_Y^{-1}C_{X,Y} = (t_i \wedge t_j - t_it_j)_{i,j}$, i.e.

$$\begin{pmatrix} t_1(1-t_1) & t_1(1-t_2) & \dots & t_1(1-t_m) \\ t_1(1-t_2) & t_2(1-t_2) & \dots & t_2(1-t_m) \\ & & \ddots & \\ t_1(1-t_m) & t_2(1-t_m) & \dots & t_m(1-t_m) \end{pmatrix}$$

and mean (here $\mathrm{E}[X] = 0, \mathrm{E}[Y] = 0$)

$$C_{X,Y}C_Y^{-1}y = \begin{pmatrix} t_1 y \\ \vdots \\ t_m y \end{pmatrix} .$$

b) Now $Y = (B_1, B_v)$ and

$$C_Y = \begin{pmatrix} 1 & 1 \\ 1 & v \end{pmatrix}, \qquad C_Y^{-1} = \frac{1}{v-1}\begin{pmatrix} v & -1 \\ -1 & 1 \end{pmatrix}, \qquad C_{X,Y} = \begin{pmatrix} t_1 & t_1 \\ \vdots & \vdots \\ t_m & t_m \end{pmatrix}$$

so that

$$C_{X,Y}C_Y^{-1} == \frac{1}{v-1}\begin{pmatrix} t_1 & t_1 \\ \vdots & \vdots \\ t_m & t_m \end{pmatrix}\begin{pmatrix} v & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} t_1 & 0 \\ \vdots & \vdots \\ t_m & 0 \end{pmatrix}$$

and

$$C_{X,Y}C_Y^{-1}C_{X,Y}^* = \begin{pmatrix} t_1 & 0 \\ \vdots & \vdots \\ t_m & 0 \end{pmatrix}\begin{pmatrix} t_1 & \dots & t_m \\ t_1 & \dots & t_m \end{pmatrix} ,$$

which is still the matrix with $t_i t_j$ entries. Hence the covariance matrix of the conditional law is the same as in a). The same holds for the mean which is equal to

$$C_{X,Y} C_Y^{-1} \begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} t_1 y \\ \vdots \\ t_m y \end{pmatrix} .$$

In conclusion, to be repositioned the conditional law neither depends on $x$ nor on $v$ and coincides with the conditional law given $B_1 = y$. From the point of view of intuition this is not surprising as the addition of the information $B_v = x$ means to add the information $B_v - B_1 = x - y$. As the increments of the Brownian motion are independent of the past it would have been reasonable to expect that the addition of this new information does not change the conditional law. This fact may be obtained directly using Exercise 4.1 b).

**4.13** Thanks to Exercise 3.11 the joint law of the two r.v.'s is Gaussian. In order to identify it, we just need to compute its mean and covariance matrix. The two r.v.'s are obviously centered. Let us compute the variance of the second one:

$$\begin{aligned} \mathrm{E}\Big[\Big(\int_0^1 B_s\,ds\Big)^2\Big] &= \mathrm{E}\Big[\int_0^1 B_s\,ds \cdot \int_0^1 B_t\,dt\Big] = \mathrm{E}\Big[\int_0^1 \int_0^1 B_s B_t\,ds\,dt\Big] \\ &= \int_0^1 \int_0^1 s \wedge t\,ds\,dt = \int_0^1 dt \int_0^t s\,ds + \int_0^1 dt \int_t^1 t\,ds = I_1 + I_2\,. \end{aligned}$$

We have easily

$$I_1 = \int_0^1 \frac{t^2}{2}\,dt = \frac{1}{6}$$

and similarly $I_2 = \frac{1}{6}$; therefore the variance is $\frac{1}{3}$. The variance of $B_1$ is equal to 1, of course, whereas for the covariance we find

$$\mathrm{E}\Big[B_1 \int_0^1 B_s\,ds\Big] = \int_0^1 \mathrm{E}(B_1 B_s)\,ds = \int_0^1 s\,ds = \frac{1}{2} \cdot$$

The covariance matrix of $B_1$ and $\int_0^1 B_s\,ds$ is therefore

$$\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix} .$$

The best $L^2$ estimate of the value of $B_1$ given $\int_0^1 B_t\,dt = x$ is the conditional mean, which is easily computed using formula (4.23) and is equal to $\frac{3}{2}x$.

**4.14**

a) We have $\mathrm{Cov}(\eta, Y_s) = \mathrm{Cov}(\eta, \eta s) + \mathrm{Cov}(\eta, \sigma B_s) = s\rho^2$ ($\eta$ and $B_s$ are independent). By the same argument

$$\mathrm{Cov}(Y_s, Y_t) = \mathrm{Cov}(\eta s, \eta t) + \mathrm{Cov}(\sigma B_s, \sigma B_t) = st\rho^2 + (s \wedge t)\sigma^2 .$$

b) As $\eta$ is independent of $B$, $(\eta, (B_t)_{t\geq 0})$ is a Gaussian family. Therefore for every $t_1, \dots, t_m$ the vector $(Y_{t_1}, \dots, Y_{t_m})$ is Gaussian, being a linear function of $(\eta, B_{t_1}, \dots, B_{t_m})$.

c) $\lambda$ must satisfy the relation $\mathrm{Cov}(\eta - \lambda Y_t, Y_s) = 0$ for every $s \leq t$. We have

$$\begin{aligned}\mathrm{Cov}(\eta - \lambda Y_t, Y_s) &= \mathrm{Cov}(\eta, Y_s) - \mathrm{Cov}(\lambda Y_t, Y_s) = s\rho^2 - \lambda(st\rho^2 + s\sigma^2)\\ &= s(\rho^2 - \lambda(t\rho^2 + \sigma^2))\end{aligned}$$

and therefore

$$\lambda = \frac{\rho^2}{\sigma^2 + t\rho^2} \cdot$$

With this choice of $\lambda$ the r.v. $Z = \eta - \lambda Y_t$ is not correlated with each of the r.v.'s $Y_s, s \leq t$. As these generate the $\sigma$-algebra $\mathscr{G}_t$ and $(\eta, Y_t, t \geq 0)$ is a Gaussian family, by Remark 1.2 $Z$ is independent of $\mathscr{G}_t$.

d) As $Y_t$ is $\mathscr{G}_t$-measurable whereas $Z = \eta - \lambda Y_t$ is independent of $\mathscr{G}_t$ and $\mathrm{E}[Y_t] = \mu t$,

$$\begin{aligned}\mathrm{E}[\eta|\mathscr{G}_t] &= \mathrm{E}[\lambda Y_t + Z|\mathscr{G}_t] = \lambda Y_t + \mathrm{E}[Z] = \lambda Y_t + \mathrm{E}[\eta - \lambda Y_t]\\ &= \lambda Y_t + \mu(1 - \lambda t) = \frac{\rho^2 Y_t}{\sigma^2 + t\rho^2} + \mu\Big(1 - \frac{t\rho^2}{\sigma^2 + t\rho^2}\Big) = \frac{\rho^2 Y_t + \sigma^2\mu}{\sigma^2 + t\rho^2} \cdot\end{aligned}$$

We have

$$\begin{aligned}\lim_{t\to+\infty} \mathrm{E}[\eta|\mathscr{G}_t] &= \lim_{t\to+\infty} \frac{\rho^2 Y_t + \sigma^2\mu}{\sigma^2 + t\rho^2} = \lim_{t\to+\infty} \frac{\rho^2 Y_t}{\sigma^2 + t\rho^2}\\ &= \lim_{t\to+\infty} \frac{\rho^2}{\sigma^2 + t\rho^2}\,(\eta t + \sigma B_t) = \eta \qquad \text{a.s.}\end{aligned}$$

as $\lim_{t\to+\infty} \frac{1}{t} B_t = 0$ a.s. by the Iterated Logarithm Law.

**4.15**

a) If $t_1, \dots, t_n \in \mathbb{R}^+$, then $(X_{t_1}, \dots, X_{t_n})$ is Gaussian, being a linear function of $(B_{t_1}, \dots, B_{t_n}, B_1)$. Moreover, if $t \leq 1$,

$$\mathrm{E}[X_t B_1] = \mathrm{E}[(B_t - tB_1)B_1] = t \wedge 1 - t = 0 .$$

The two r.v.'s $X_t$ and $B_1$, being jointly Gaussian and uncorrelated, are independent for every $t \le 1$. $X_t$ is centered and, if $s \le t$,

$$\mathrm{E}[X_t X_s] = \mathrm{E}[(B_t - tB_1)(B_s - sB_1)] = s - st - st + st = s(1-t)\ . \tag{S.13}$$

b) As $X_s$ and $X_t - \frac{1-t}{1-s}X_s$ are jointly Gaussian, we need only to show that they are not correlated. We have, recalling that $s \le t$,

$$\mathrm{E}\Big[\Big(X_t - \frac{1-t}{1-s}X_s\Big)X_s\Big] = s(1-t) - \frac{1-t}{1-s}s(1-s) = 0\ .$$

c) We have

$$\mathrm{E}[X_t|X_s] = \mathrm{E}\Big[X_t - \frac{1-t}{1-s}X_s + \frac{1-t}{1-s}X_s\,\Big|\,X_s\Big] = \frac{1-t}{1-s}X_s \tag{S.14}$$

as, thanks to b), $X_t - \frac{1-t}{1-s}X_s$ is independent of $X_s$ and centered whereas $\frac{1-t}{1-s}X_s$ is already $\sigma(X_s)$-measurable. The relation $\mathrm{E}[X_t|X_s] = \mathrm{E}[X_t|\mathscr{G}_s]$ follows if we show that $X_t - \frac{1-t}{1-s}X_s$ is independent of $\mathscr{G}_s$ and then repeat the same argument as in (S.14). In order to obtain this it is sufficient to show that $X_t - \frac{1-t}{1-s}X_s$ is independent of $X_u$ for every $u \le s$. This is true as, if $u \le s$,

$$\mathrm{E}\Big[\Big(X_t - \frac{1-t}{1-s}X_s\Big)X_u\Big] = u(1-t) - \frac{1-t}{1-s}u(1-s) = 0\ .$$

d) We have, for $s \le t \le 1$,

$$\mathrm{E}[X_t|\mathscr{F}_s] = \mathrm{E}[B_t - tB_1|\mathscr{F}_s] = B_s - tB_s = (1-t)B_s\ .$$

This result is different from

$$\mathrm{E}[X_t|\mathscr{G}_s] = \frac{1-t}{1-s}X_s$$

obtained in c), as it is easy to see that $\mathrm{Var}(\frac{1-t}{1-s}X_s) = (1-t)^2\frac{s}{1-s}$ whereas $\mathrm{Var}((1-t)B_s) = (1-t)^2 s$. Therefore the two $\sigma$-algebras $\mathscr{F}_s$ and $\mathscr{G}_s$ are different.

e) Let $0 \le t_1 < \dots < t_m \le 1$. The conditional law of the vector $(B_{t_1}, \dots, B_{t_m})$ given $B_1 = 0$ can be computed as explained in Sect. 4.4. It has actually already been computed in Exercise 4.12 b). It is Gaussian, centered, and with covariance matrix whose $(i,j)$-th entry is $t_i \wedge t_j - t_i t_j$, the same as the covariance matrix of the Brownian bridge, as it can be computed from (S.13).

**5.1** For every $t > s$, we have $\mathrm{E}(X_t|\mathscr{F}_s) \le X_s$ a.s. Therefore the r.v. $U = X_s - \mathrm{E}(X_t|\mathscr{F}_s)$ is positive a.s.; but it has zero mean, as

$$\mathrm{E}[U] = \mathrm{E}[X_s] - \mathrm{E}[\mathrm{E}(X_t|\mathscr{F}_s)] = \mathrm{E}(X_s) - \mathrm{E}(X_t) = 0\ .$$

We deduce that $U = 0$ a.s. (a positive r.v. having mean equal to 0 is $= 0$ a.s.). Therefore $\mathrm{E}(X_t|\mathscr{F}_s) = X_s$ a.s. for every $t > s$.

**5.2**

a) We know that $\mathrm{P}(X_{\tau_{a,b}} = b) = \frac{a}{a+b}$, hence

$$\lim_{a\to+\infty} \mathrm{P}(X_{\tau_{a,b}} = b) = 1 ,$$

which was to be expected: if the left endpoint of the interval $]-a, b[$ is far from the origin, it is more likely for the exit to take place at $b$.

b) The event $\{X_{\tau_{a,b}} = b\}$ is contained in $\{\tau_b < +\infty\}$, hence, for every $a > 0$,

$$\mathrm{P}(\tau_b < +\infty) \geq \mathrm{P}(X_{\tau_{a,b}} = b)$$

so that, thanks to a), $\mathrm{P}(\tau_b < +\infty) = 1$.

**5.3**

a) Thanks to the law of large numbers we have a.s.

$$\frac{1}{n}X_n = \frac{1}{n}(Y_1 + \cdots + Y_n) \underset{n\to\infty}{\to} \mathrm{E}[Y_1] = p - q < 0 .$$

Therefore a.s. there exists an $n_0 > 0$ such that $\frac{1}{n}X_n \leq \frac{1}{2}(p-q) < 0$ for $n \geq n_0$, which implies $X_n \leq \frac{n}{2}(p-q)$; hence $X_n \to -\infty$ a.s.

b1) Let $\mathscr{F}_n = \sigma(X_1, \dots, X_n) = \sigma(Y_1, \dots, Y_n)$. We have, $Y_{n+1}$ being independent of $\mathscr{F}_n$,

$$\begin{aligned}
\mathrm{E}[Z_{n+1}|\mathscr{F}_n] &= \mathrm{E}\Big[\Big(\frac{q}{p}\Big)^{X_{n+1}}\Big|\mathscr{F}_n\Big] = \mathrm{E}\Big[\Big(\frac{q}{p}\Big)^{X_n+Y_{n+1}}\Big|\mathscr{F}_n\Big] = \Big(\frac{q}{p}\Big)^{X_n}\mathrm{E}\Big[\Big(\frac{q}{p}\Big)^{Y_{n+1}}\Big|\mathscr{F}_n\Big] \\
&= \Big(\frac{q}{p}\Big)^{X_n}\mathrm{E}\Big[\Big(\frac{q}{p}\Big)^{Y_{n+1}}\Big] = \Big(\frac{q}{p}\Big)^{X_n}\Big(\frac{q}{p}\mathrm{P}(Y_{n+1} = 1) + \frac{p}{q}\mathrm{P}(Y_{n+1} = -1)\Big) \\
&= \Big(\frac{q}{p}\Big)^{X_n}\Big(\frac{q}{p}p + \frac{p}{q}q\Big) = \Big(\frac{q}{p}\Big)^{X_n}(p+q) = \Big(\frac{q}{p}\Big)^{X_n} = Z_n .
\end{aligned}$$

Note that this argument proves that the product of independent r.v.'s having mean equal to 1 always gives rise to a martingale with respect to their natural filtration. Here we are dealing with an instance of this case.

b2) As, thanks to a), $\lim_{n\to\infty} X_n = -\infty$ a.s. and $\frac{q}{p} > 1$, we have $\lim_{n\to\infty} Z_n = 0$.

c) As $n \wedge \tau$ is a bounded stopping time, by the stopping theorem $\mathrm{E}[Z_{n\wedge\tau}] = \mathrm{E}[Z_1] = 1$. By a) $\tau < +\infty$, therefore $\lim_{n\to\infty} Z_{n\wedge\tau} = Z_\tau$ a.s. As $\frac{q}{p} > 1$ and $-a \leq X_{n\wedge\tau} \leq b$, we have $(\frac{q}{p})^{-a} \leq Z_{n\wedge\tau} \leq (\frac{q}{p})^b$ and we can apply Lebesgue's theorem and obtain that $\mathrm{E}[Z_\tau] = \lim_{n\to\infty} \mathrm{E}[Z_{n\wedge\tau}] = 1$.

d) As $X_\tau$ can only take the values $-a, b$,

$$1 = \mathrm{E}[Z_\tau] = \mathrm{E}\Big[\Big(\frac{q}{p}\Big)^{X_\tau}\Big] = \Big(\frac{q}{p}\Big)^{b} \mathrm{P}(X_\tau = b) + \Big(\frac{q}{p}\Big)^{-a} \mathrm{P}(X_\tau = -a)$$
$$= \Big(\frac{q}{p}\Big)^{b} \mathrm{P}(X_\tau = b) + \Big(\frac{q}{p}\Big)^{-a} (1 - \mathrm{P}(X_\tau = b))\,.$$

Actually, as $\tau < +\infty$, $\mathrm{P}(X_\tau = -a) = 1 - \mathrm{P}(X_\tau = b)$. Hence

$$1 - \Big(\frac{q}{p}\Big)^{-a} = \mathrm{P}(X_\tau = b)\Big(\Big(\frac{q}{p}\Big)^{b} - \Big(\frac{q}{p}\Big)^{-a}\Big)\,,$$

i.e.

$$\mathrm{P}(X_\tau = b) = \frac{1 - (\frac{q}{p})^{-a}}{(\frac{q}{p})^{b} - (\frac{q}{p})^{-a}}\,\cdot$$

**5.4**

a) $Z_k$ is $\mathscr{F}_{k-1}$-measurable whereas $X_k$ is independent of $\mathscr{F}_k$, therefore $X_k$ and $Z_k$ are independent. Thus $Z_k^2 X_k^2$ is integrable, being the product of integrable independent r.v.'s and $Y_n$ is square integrable for every $n$ (beware of a possible confusion: "$(Y_n)_n$ square integrable" means $Y_n \in L^2$ for every $n$, "$(Y_n)_n$ bounded in $L^2$" means $\sup_{n>0} \mathrm{E}(Y_n^2) < +\infty$). Moreover,

$$\mathrm{E}[Y_{n+1}\,|\,\mathscr{F}_n] = \mathrm{E}[Y_n + Z_{n+1}X_{n+1}\,|\,\mathscr{F}_n] = Y_n + Z_{n+1}\mathrm{E}[X_{n+1}\,|\,\mathscr{F}_n]$$
$$= Y_n + Z_{n+1}\underbrace{\mathrm{E}[X_{n+1}]}_{=0} = Y_n\,.$$

b) As $Z_k$ and $X_k$ are independent, $\mathrm{E}[Z_k X_k] = \mathrm{E}[Z_k]\mathrm{E}[X_k] = 0$ and therefore $\mathrm{E}[Y_k] = 0$. Moreover, we have

$$\mathrm{E}[Y_k^2] = \mathrm{E}\Big[\Big(\sum_{k=1}^n Z_k X_k\Big)^2\Big] = \mathrm{E}\Big[\sum_{k,h=1}^n Z_k X_k Z_h X_h\Big] = \sum_{k,h=1}^n \mathrm{E}[Z_k X_k Z_h X_h]$$

but in the previous sum all the terms with $h \neq k$ vanish: let us assume $k > h$, then $Z_k X_h Z_h$ is $\mathscr{F}_{k-1}$-measurable whereas $X_k$ is independent of $\mathscr{F}_{k-1}$. Therefore

$$\mathrm{E}[Z_k X_k Z_h X_h] = \underbrace{\mathrm{E}[X_k]}_{=0}\,\mathrm{E}[Z_k Z_h X_h]\,,$$

hence, using again the independence of $Z_k$ and $X_k$,

$$\mathrm{E}[Y_k^2] = \sum_{k=1}^n \mathrm{E}[Z_k^2 X_k^2] = \sum_{k=1}^n \mathrm{E}[Z_k^2]\mathrm{E}[X_k^2] = \sigma^2 \sum_{k=1}^n \mathrm{E}[Z_k^2]\,. \tag{S.15}$$

The compensator $(A_n)_n$ of the submartingale $(M_n^2)_n$ is given by the relations $A_0 = 0$, $A_{n+1} = A_n + \mathrm{E}[M_{n+1}^2 - M_n^2 | \mathscr{F}_n]$. Now

$$\begin{aligned}\mathrm{E}[M_{n+1}^2 - M_n^2 | \mathscr{F}_n] &= \mathrm{E}[(M_n + X_{n+1})^2 - M_n^2 | \mathscr{F}_n]\\ = \mathrm{E}[M_n^2 + 2M_nX_{n+1} + X_{n+1}^2 - M_n^2 | \mathscr{F}_n] &= \mathrm{E}[2M_nX_{n+1} + X_{n+1}^2 | \mathscr{F}_n]\\ &= 2M_n\mathrm{E}[X_{n+1} | \mathscr{F}_n] + \mathrm{E}[X_{n+1}^2 | \mathscr{F}_n]\,.\end{aligned}$$

As $X_{n+1}$ is independent of $\mathscr{F}_n$,

$$\begin{aligned}\mathrm{E}[X_{n+1} | \mathscr{F}_n] &= \mathrm{E}[X_{n+1}] = 0 \qquad \text{a.s.}\\ \mathrm{E}[X_{n+1}^2 | \mathscr{F}_n] &= \mathrm{E}[X_{n+1}^2] = \sigma^2 \qquad \text{a.s.}\end{aligned} \tag{S.16}$$

and therefore $A_{n+1} = A_n + \sigma^2$, which with the condition $A_0 = 0$ gives $A_n = n\sigma^2$.

In order to compute the compensator of $(Y_n^2)_n$ just repeat the same argument. If we denote it by $(B_n)_n$, then

$$\begin{aligned}\mathrm{E}[Y_{n+1}^2 - Y_n^2 | \mathscr{F}_n] &= \mathrm{E}[(Y_n + Z_{n+1}X_{n+1})^2 - Y_n^2 | \mathscr{F}_n]\\ &= \mathrm{E}[2Y_nZ_{n+1}X_{n+1} + Z_{n+1}^2X_{n+1}^2 | \mathscr{F}_n]\\ = 2Y_nZ_{n+1}\mathrm{E}[X_{n+1} | \mathscr{F}_n] + \mathrm{E}[Z_{n+1}^2X_{n+1}^2 | \mathscr{F}_n] &= Z_{n+1}^2\mathrm{E}[X_{n+1}^2 | \mathscr{F}_n] = \sigma^2 Z_{n+1}^2\,.\end{aligned}$$

From the relations $B_0 = 0$ and $B_{n+1} = B_n + \mathrm{E}[Y_{n+1}^2 - Y_n^2 | \mathscr{F}_n]$ we find

$$B_n = \sigma^2 \sum_{k=1}^n Z_k^2\,.$$

c) By (S.15)

$$\mathrm{E}[Y_n^2] = \sigma^2 \sum_{k=1}^n \frac{1}{k^2}\,\cdot$$

As the series on the right-hand side is convergent, the martingale $(Y_n)_n$ is bounded in $L^2$. It is therefore convergent a.s. and in $L^2$ and therefore uniformly integrable.

**5.5**

a) We have

$$\mathrm{E}[Y_k] = -1 \cdot 2^{-k} + 0 \cdot (1 - 2^{1-k}) + 1 \cdot 2^{-k} = 0\,,$$

Therefore $(X_n)_n$, being the sum of independent centered r.v.'s, is a martingale (Example 5.1 a)).

b) The r.v.'s $Y_k$ are square integrable and therefore $X_n \in L^2$, as the sum of square integrable r.v.'s. The associated increasing process of the martingale $(X_n)_n$, i.e. the compensator of the submartingale $(X_n^2)_n$, is defined by $A_0 = 0$ and for $n \ge 1$

$$A_n = A_{n-1} + \mathrm{E}[(X_n - X_{n-1})^2 | \mathscr{F}_{n-1}] = A_{n-1} + \mathrm{E}[Y_n^2 | \mathscr{F}_{n-1}] = A_{n-1} + \mathrm{E}[Y_n^2] \; .$$

As $\mathrm{E}[Y_n^2] = 2 \cdot 2^{-n}$,

$$A_n = 1 + \frac{1}{2} + \cdots + \frac{1}{2^{n-1}} = 2(1 - 2^{-n}) \; .$$

(Note that the associated increasing process $(A_n)_n$ turns out to be deterministic, as is always the case for a martingale with independent increments).

c) Thanks to b) the associated increasing process $(A_n)_n$ is bounded. As

$$A_n = \mathrm{E}[X_n^2]$$

the martingale $(X_n)_n$ is bounded in $L^2$. Hence it converges a.s. and in $L^2$. It also converges in $L^1$, as $L^p$ convergence implies convergence in $L^{p'}$ for every $p' \le p$.

**5.6**

a) We have

$$\{\tau \le u\} = \begin{cases} \emptyset & \text{if } u < s \\ A & \text{if } s \le u < t \\ \Omega & \text{if } u \ge t \end{cases}$$

therefore, in any case, $\{\tau \le u\} \in \mathscr{F}_u$ for every $u$.

b) If $X$ is a martingale then the relation $\mathrm{E}(X_\tau) = \mathrm{E}(X_0)$ for every bounded stopping time $\tau$ is a consequence of the stopping theorem (Theorem 5.13) applied to the stopping times $\tau_2 = \tau$ and $\tau_1 = 0$.

Conversely, assume that $\mathrm{E}(X_\tau) = \mathrm{E}(X_0)$ for every bounded stopping time $\tau$ and let us prove the martingale property , i.e. that, if $t > s$, for every $A \in \mathscr{F}_s$

$$\mathrm{E}(X_t 1_A) = \mathrm{E}(X_s 1_A) \; . \tag{S.17}$$

The idea is to find two bounded stopping times $\tau_1, \tau_2$ such that from the relation $\mathrm{E}[X_{\tau_1}] = \mathrm{E}[X_{\tau_2}]$ (S.17) follows. Let us choose, for a fixed $A \in \mathscr{F}_s$, $\tau$ as in a) and $\tau_2 \equiv t$. Now $X_\tau = X_s 1_A + X_t 1_{A^c}$ and the relation $\mathrm{E}[X_\tau] = \mathrm{E}[X_t]$ can be written as

$$\mathrm{E}[X_s 1_A] + \mathrm{E}[X_t 1_{A^c}] = \mathrm{E}[X_\tau] = \mathrm{E}[X_t] = \mathrm{E}[X_t 1_A] + \mathrm{E}[X_t 1_{A^c}]$$

and by subtraction we obtain (S.17).

**5.7** Note first that $M_t$ is integrable for every $t > 0$, thanks to the integrability of $\mathrm{e}^{B_t}$. Then, as indicated at the end of Sect. 5.1, it suffices to prove that $g(x) = (\mathrm{e}^x - K)^+$ is a convex function. But this is immediate as $g$ is the composition of the functions $x \mapsto \mathrm{e}^x - K$, which is convex, and of $y \mapsto y^+$, which is convex *and* increasing.

**5.8**

a) If $s \le t$, as $\{M_s = 0\} \in \mathscr{F}_s$,

$$\mathrm{E}(1_{\{M_s=0\}}M_t) = \mathrm{E}(1_{\{M_s=0\}}M_s) = 0\ .$$

As $M_t \ge 0$, we must have $M_t = 0$ a.s. on $\{M_s = 0\}$ (this is also a consequence of Exercise 4.4 a)).

b1) $\tau$ is the infimum of the set of times at which $M$ vanishes; therefore there exists a sequence $(t_n)_n$ with $t_n \searrow \tau$ such that $M_{t_n} = 0$. Therefore, as $M$ is right-continuous, $M_\tau = 0$.

b2) Let $T > 0$ and $\tau = \inf\{t; M_t = 0\}$. Let us prove that $\mathrm{P}(\tau \le T) = 0$. By the stopping theorem and as $M_\tau = 0$ a.s.

$$\mathrm{E}(M_T) = \mathrm{E}(M_{T\wedge\tau}) = \mathrm{E}(M_T 1_{\{\tau>T\}} + M_\tau 1_{\{\tau\le T\}}) = \mathrm{E}(M_T 1_{\{\tau>T\}})\ .$$

As $M_T > 0$ a.s. this equality is possible only if $\mathrm{P}(\tau > T) = 1$.

**5.9**

a) This is an extension to an $m$-dimensional Brownian motion of what we have already seen in Example 5.2. Let $s < t$. As $B_s$ is $\mathscr{F}_s$-measurable whereas $B_t - B_s$ is independent of $\mathscr{F}_s$,

$$\begin{aligned}\mathrm{E}(X_t\,|\,\mathscr{F}_s) &= \mathrm{E}(\mathrm{e}^{\langle\lambda,B_s\rangle+\langle\lambda,B_t-B_s\rangle-\frac{1}{2}|\lambda|^2t}\,|\,\mathscr{F}_s)\\ &= \mathrm{e}^{\langle\lambda,B_s\rangle-\frac{1}{2}|\lambda|^2t}\,\mathrm{E}(\mathrm{e}^{\langle\lambda,B_t-B_s\rangle}\,|\,\mathscr{F}_s) = \mathrm{e}^{\langle\lambda,B_s\rangle-\frac{1}{2}|\lambda|^2t}\,\mathrm{E}(\mathrm{e}^{\langle\lambda,B_t-B_s\rangle})\\ &= \mathrm{e}^{\langle\lambda,B_s\rangle-\frac{1}{2}|\lambda|^2t}\,\mathrm{e}^{\frac{1}{2}|\lambda|^2(t-s)} = \mathrm{e}^{\langle\lambda,B_s\rangle-\frac{1}{2}|\lambda|^2s} = X_s\ .\end{aligned}$$

b) Two ways of reasoning are possible.

First let us assume that $m = 1$. For large $t$, $B_t(\omega) \le ((2+\varepsilon)t\log\log t)^{1/2}$ by the Iterated Logarithm Law. Therefore, a.s.,

$$\begin{aligned}&\lambda B_t - \frac{1}{2}|\lambda|^2 t\\ &= ((2+\varepsilon)t\log\log t)^{1/2}\Big(\frac{B_t}{((2+\varepsilon)t\log\log t)^{1/2}} - \frac{1}{2}\underbrace{\frac{|\lambda|^2\sqrt{t}}{((2+\varepsilon)\log\log t)^{1/2}}}_{\to+\infty}\Big) \underset{t\to+\infty}{\to} -\infty\end{aligned}$$

and therefore

$$X_t = e^{\lambda B_t - \frac{1}{2}|\lambda|^2 t} \underset{t\to+\infty}{\to} 0 \quad \text{a.s.}$$

If $m > 1$, then we know that $W_t = \frac{1}{|\lambda|}\langle \lambda, B_t\rangle$ is a Brownian motion. We can write

$$X_t = e^{|\lambda| W_t - \frac{1}{2}|\lambda|^2 t}$$

and then repeat the argument above with the Iterated Logarithm Law applied to $W$.

The second approach is the following: let $0 < \alpha < 1$. Then

$$\mathrm{E}[X_t^\alpha] = e^{-\frac{1}{2}\alpha|\lambda|^2 t}\mathrm{E}[e^{\langle \alpha\lambda, B_t\rangle}] = e^{-\frac{1}{2}\alpha|\lambda|^2 t} e^{\frac{1}{2}\alpha^2|\lambda|^2 t} = e^{-\frac{1}{2}|\lambda|^2 t\alpha(1-\alpha)} .$$

Hence $\lim_{t\to+\infty}\mathrm{E}[X_t^\alpha] = 0$. Let us denote by $X_\infty$ the a.s. limit of $X$ (whose existence is guaranteed, $X$ being a continuous positive martingale). Then by Fatou's lemma

$$\mathrm{E}[X_\infty^\alpha] \le \varliminf_{t\to+\infty} \mathrm{E}[X_t^\alpha] = 0 .$$

The positive r.v. $X_\infty^\alpha$, having expectation equal to 0, is therefore $= 0$ a.s.

This second approach uses the nice properties of martingales (the previous one with the Iterated Logarithm Law does not) and can be reproduced in other similar situations.

c) If the martingale $(X_t)_t$ was uniformly integrable, it would also converge to 0 in $L^1$ and we would have $\mathrm{E}(X_t) \to 0$ as $t \to +\infty$. But this it is not the case, as $\mathrm{E}(X_t) = \mathrm{E}(X_0) = 1$ for every $t \ge 0$.

**5.10**

a) The first point has already been proved in Remark 5.4 on p. 132. However, let us produce a direct proof. If $s \le t$, as $B_s$ is $\mathscr{F}_s$-measurable whereas $B_t - B_s$ is independent of $\mathscr{F}_s$,

$$\begin{aligned}\mathrm{E}(Y_t|\mathscr{F}_s) &= \mathrm{E}[(B_s + (B_t - B_s))^2|\mathscr{F}_s] - t\\ &= \mathrm{E}[B_s^2 + 2B_s(B_t - B_s) + (B_t - B_s)^2|\mathscr{F}_s] - t\\ &= B_s^2 + 2B_s\underbrace{\mathrm{E}(B_t - B_s|\mathscr{F}_s)}_{=0} + \underbrace{\mathrm{E}[(B_t - B_s)^2]}_{=t-s} - t = B_s^2 - s = Y_s .\end{aligned}$$

If $Y$ was uniformly integrable then it would converge a.s. and in $L^1$. This is not possible, as we know by the Iterated Logarithm Law that there exists a sequence of times $(t_n)_n$ such that $t_n \to +\infty$ and $B_{t_n} = 0$. Therefore $\varliminf_{t\to+\infty} Y_{t_n} = -\infty$ a.s.

b) The requested equality would be immediate if $\tau$ was bounded, which we do not know (actually it is not). But, for every $t > 0$, $t \wedge \tau$ is a bounded stopping time and by the stopping theorem $0 = \mathrm{E}[Y_{t\wedge\tau}] = \mathrm{E}[B^2_{t\wedge\tau} - (t\wedge\tau)]$. Therefore

$$\mathrm{E}(B^2_{t\wedge\tau}) = \mathrm{E}(t\wedge\tau)$$

and we can take the limit as $t \to +\infty$

$$\lim_{t\to+\infty} \mathrm{E}(B^2_{t\wedge\tau}) = \mathrm{E}(B^2_\tau)$$

by Lebesgue's theorem ($B^2_{t\wedge\tau} \le \max(a^2, b^2)$) and

$$\lim_{t\to+\infty} \mathrm{E}(t\wedge\tau) = \mathrm{E}(\tau)$$

by Beppo Levi's theorem ($t \mapsto t\wedge\tau$ is obviously increasing). In Example 5.3 we have seen that

$$\mathrm{P}(B_\tau = -a) = \frac{b}{a+b}\,, \qquad \mathrm{P}(B_\tau = b) = \frac{a}{a+b}\,\cdot$$

Therefore we find

$$\mathrm{E}(\tau) = \mathrm{E}(B^2_\tau) = \frac{a^2b + b^2a}{a+b} = ab\,.$$

c1) Immediate as

$$Y_t = B_1(t)^2 - t + \dots + B_m(t)^2 - t$$

so that, thanks to a), $(Y_t)_t$ turns out to be the sum of $m$ $(\mathscr{F}_t)_t$-martingales.

c2) Recall that we know already that $\tau < +\infty$ a.s. (Exercise 3.18). By the stopping theorem, for every $t > 0$, $\mathrm{E}[|B_{\tau\wedge t}|^2 - \tau\wedge t] = 0$, i.e.

$$\mathrm{E}[|B_{\tau\wedge t}|^2] = m\mathrm{E}[\tau\wedge t]\,.$$

By a repetition of the argument of b), i.e. using Lebesgue's theorem for the left-hand side and Beppo Levi's for the right-hand side, we find

$$\mathrm{E}[|B_\tau|^2] = m\mathrm{E}[\tau]$$

and, observing that $|B_\tau|^2 = 1$ a.s., we can conclude that

$$\mathrm{E}[\tau] = \frac{1}{m}\,\cdot$$

**5.11**

a) As $\mathrm{E}[M_t M_s\,|\,\mathscr{F}_s] = M_s \mathrm{E}[M_t\,|\,\mathscr{F}_s] = M_s^2$ we have

$$\begin{aligned}\mathrm{E}[(M_t - M_s)^2] &= \mathrm{E}[\mathrm{E}[(M_t - M_s)^2\,|\,\mathscr{F}_s]]\\ &= \mathrm{E}[\mathrm{E}[M_t^2 - 2M_t M_s + M_s^2\,|\,\mathscr{F}_s]] = \mathrm{E}[M_t^2 - M_s^2]\,.\end{aligned}$$

b) Let us assume $M_0 = 0$ for simplicity. This is possible because the two martingales $(M_t)_t$ and $(M_t - M_0)_t$ have the same associated increasing process. Note that the suggested associated increasing process vanishes at 0. We have then

$$\mathrm{E}[M_t^2\,|\,\mathscr{F}_s] = \mathrm{E}[(M_t - M_s + M_s)^2\,|\,\mathscr{F}_s] = \mathrm{E}[(M_t - M_s)^2 + 2(M_t - M_s)M_s + M_s^2\,|\,\mathscr{F}_s]\,.$$

But $\mathrm{E}[(M_t - M_s)^2\,|\,\mathscr{F}_s] = \mathrm{E}[(M_t - M_s)^2] = \mathrm{E}[M_t^2 - M_s^2]$ as $M$ has independent increments whereas $\mathrm{E}[(M_t - M_s)M_s\,|\,\mathscr{F}_s] = M_s \mathrm{E}[M_t - M_s\,|\,\mathscr{F}_s] = 0$. Therefore

$$\mathrm{E}[M_t^2\,|\,\mathscr{F}_s] = M_s^2 + \mathrm{E}[M_t^2 - M_s^2]\,,$$

from which it follows that $Z_t = M_t^2 - \mathrm{E}(M_t^2)$ is a martingale, i.e. that $\langle M\rangle_t = \mathrm{E}[M_t^2]$.

c) $M$ being a Gaussian process we know (Remark 1.2) that $M_t - M_s$ is independent of $\mathscr{G}_s = \sigma(M_u, u \le s)$ if and only if, for every $u \le s$,

$$\mathrm{E}[(M_t - M_s)M_u] = 0\,.$$

But this relation is immediate as, $M$ being a martingale,

$$\mathrm{E}[(M_t - M_s)M_u] = \mathrm{E}\big[\mathrm{E}[(M_t - M_s)M_u\,|\,\mathscr{G}_s]\big] = \mathrm{E}\big[\underbrace{\mathrm{E}[(M_t - M_s)\,|\,\mathscr{G}_s]}_{=0} M_u\big]\,.$$

d) As $\langle M\rangle$ is deterministic and because of the independence of the increments,

$$\begin{aligned}\mathrm{E}\big[\mathrm{e}^{\theta M_t - \frac{1}{2}\theta^2\langle M\rangle_t}\,\big|\,\mathscr{F}_s\big] &= \mathrm{e}^{\theta M_s - \frac{1}{2}\theta^2\langle M\rangle_t}\mathrm{E}\big[\mathrm{e}^{\theta(M_t - M_s)}\,\big|\,\mathscr{F}_s\big]\\ &= \mathrm{e}^{\theta M_s - \frac{1}{2}\theta^2\langle M\rangle_t}\mathrm{E}\big[\mathrm{e}^{\theta(M_t - M_s)}\big]\,.\end{aligned}$$

As $M_t - M_s$ is Gaussian, we have $\mathrm{E}\big[\mathrm{e}^{\theta(M_t - M_s)}\big] = \mathrm{e}^{\frac{1}{2}\theta^2 \mathrm{Var}(M_t - M_s)}$ (recall Exercise 1.6). As, thanks to a),

$$\mathrm{Var}(M_t - M_s) = \mathrm{E}[(M_t - M_s)^2] = \mathrm{E}[M_t^2 - M_s^2] = \langle M\rangle_t - \langle M\rangle_s\,,$$

we have

$$\mathrm{E}\big[\mathrm{e}^{\theta M_t - \frac{1}{2}\theta^2\langle M\rangle_t}\,\big|\,\mathscr{F}_s\big] = \mathrm{e}^{\theta M_s - \frac{1}{2}\theta^2\langle M\rangle_t}\mathrm{e}^{\frac{1}{2}\theta^2(\langle M\rangle_t - \langle M\rangle_s)} = Z_s\,.$$

**5.12**

a) We must prove that, for every $n$ and for every bounded Borel function $\phi : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{E}[Y_{n+1}\,\phi(Y_1,\dots,Y_n)] = \mathrm{E}[Y_n\,\phi(Y_1,\dots,Y_n)]\;. \tag{S.18}$$

But

$$\begin{aligned}\mathrm{E}[Y_{n+1}\,\phi(Y_1,\dots,Y_n)] &= \mathrm{E}[X_{n+1}\phi(X_1,\dots,X_n)]\;,\\ \mathrm{E}[Y_n\,\phi(Y_1,\dots,Y_n)] &= \mathrm{E}[X_n\phi(X_1,\dots,X_n)]\end{aligned}$$

as $(X_n)_n$ and $(Y_n)_n$ are equivalent and, $(X_n)_n$ being a martingale,

$$\mathrm{E}[X_{n+1}\phi(X_1,\dots,X_n)] = \mathrm{E}[X_n\,\phi(X_1,\dots,X_n)]\;,$$

from which (S.18) follows.

b) The proof follows the same idea as in a). Let us denote by $(\mathscr{G}_t)_t$ the natural filtration of the process $(Y_t)_t$. By Remark 4.2, in order to show that $\mathrm{E}[Y_t|\mathscr{G}_s] = Y_s$, $s \le t$, we must just prove that

$$\mathrm{E}[Y_t 1_C] = \mathrm{E}[Y_s 1_C]$$

for every $C$ in a class $\mathscr{C}$ of events that is stable with respect to finite intersections, containing $\Omega$ and generating $\mathscr{G}$. If we choose as $\mathscr{C}$ the family of events of the form $\{Y_{s_1} \in A_1,\dots,Y_{s_n} \in A_n\}$, for $n = 1,2,\dots,s_1,\dots,s_n \le s$ and $A_1,\dots,A_n \in \mathscr{B}(\mathbb{R})$, then we are led to show that

$$\mathrm{E}[Y_t\,1_{\{Y_{s_1}\in A_1\}}\dots 1_{\{Y_{s_n}\in A_n\}}] = \mathrm{E}[Y_s\,1_{\{Y_{s_1}\in A_1\}}\dots 1_{\{Y_{s_n}\in A_n\}}]\;.$$

But this follows from the fact that

$$\begin{aligned}\mathrm{E}[Y_t\,1_{\{Y_{s_1}\in A_1\}}\dots 1_{\{Y_{s_n}\in A_n\}}] &= \mathrm{E}[X_t 1_{\{X_{s_1}\in A_1\}}\dots 1_{\{X_{s_n}\in A_n\}}]\\ \mathrm{E}[Y_s\,1_{\{Y_{s_1}\in A_1\}}\dots 1_{\{Y_{s_n}\in A_n\}}] &= \mathrm{E}[X_s 1_{\{X_{s_1}\in A_1\}}\dots 1_{\{X_{s_n}\in A_n\}}]\end{aligned}$$

as $(X_t)_t$ and $(Y_t)_t$ are equivalent and

$$\mathrm{E}[X_t\,1_{\{X_{s_1}\in A_1\}}\dots 1_{\{X_{s_n}\in A_n\}}] = \mathrm{E}[X_s\,1_{\{X_{s_1}\in A_1\}}\dots 1_{\{X_{s_n}\in A_n\}}]$$

as $(X_t)_t$ is, by assumption, a martingale.

**5.13**

a) If $t > s$ and recalling Remark 4.5, then

$$\mathrm{E}[Y_t|\mathscr{F}_s] = \mathrm{E}\Big[tB_t - \int_0^t B_u\,du\,\Big|\,\mathscr{F}_s\Big] = \mathrm{E}\big[t(B_s + B_t - B_s)|\mathscr{F}_s\big] - \int_0^t \mathrm{E}\big[B_u|\mathscr{F}_s\big]\,du$$

$$= tB_s + \underbrace{\mathrm{E}[t(B_t - B_s|\mathscr{F}_s]}_{=0} - \int_0^s B_u\,du - \int_s^t B_s\,du$$

$$= tB_s - \int_0^s B_u\,du - (t-s)B_s = Y_s\,.$$

b) We have, for $u \le s < t$,

$$\mathrm{E}[(Y_t - Y_s)B_u] = \mathrm{E}[(tB_t - sB_s)B_u] - \mathrm{E}\Big[B_u \int_s^t B_v\,dv\Big]$$

$$= (t-s)u - \int_s^t \mathrm{E}[B_u B_v]\,dv = (t-s)u - \int_s^t u\,dv = 0\,.$$

**5.14**

a) Let us prove first that $X_\tau$ is $\mathscr{F}_\tau$-measurable. Let $A \in \mathscr{B}(\mathbb{R})$, then we have

$$\{X_\tau \in A, \tau \le k\} = \bigcup_{m=0}^{k} \{X_\tau \in A, \tau = m\} = \bigcup_{m=0}^{k} \{X_m \in A, \tau = m\}\,. \tag{S.19}$$

As $\{\tau = m\} = \{\tau \le m\} \setminus \{\tau \le m-1\} \in \mathscr{F}_m \subset \mathscr{F}_k$, from (S.19) we have $\{X_\tau \in A, \tau \le k\} \in \mathscr{F}_k$ for every $k$, i.e. $\{X_\tau \in A\} \in \mathscr{F}_\tau$.

We are left with the proof that $\mathrm{E}[X 1_A] = \mathrm{E}[X_\tau 1_A]$ for every $A \in \mathscr{F}_\tau$. If $A \in \mathscr{F}_\tau$, then $A \cap \{\tau = n\} \in \mathscr{F}_n$ and

$$\mathrm{E}(X 1_A) = \sum_{n=0}^{\infty} \mathrm{E}(X 1_{A\cap\{\tau=n\}}) = \sum_{n=0}^{\infty} \mathrm{E}[1_{A\cap\{\tau=n\}} \mathrm{E}(X|\mathscr{F}_n)] = \sum_{n=0}^{\infty} \mathrm{E}(X_n 1_{A\cap\{\tau=n\}})$$

$$= \sum_{n=0}^{\infty} \mathrm{E}(X_\tau 1_{A\cap\{\tau=n\}}) = \mathrm{E}(X_\tau 1_A)\,.$$

b) If $(\tau_n)_n$ is a sequence of stopping times decreasing to $\tau$ and such that $\tau_n$ takes a discrete set of values (see Lemma 3.3), then repeating the proof of a) we have $\mathrm{E}(X|\mathscr{F}_{\tau_n}) = X_{\tau_n}$. As $(X_t)_t$ is right-continuous, $X_{\tau_n} \to X_\tau$ as $n \to \infty$ a.s. Moreover, the family $(X_{\tau_n})_n$ is uniformly integrable by Proposition 5.4 and therefore $X_{\tau_n} \to X_\tau$ also in $L^1$. As $\mathrm{E}(X|\mathscr{F}_\tau) = \mathrm{E}[\mathrm{E}(X|\mathscr{F}_{\tau_n})|\mathscr{F}_\tau]$ we deduce that

$$\mathrm{E}(X|\mathscr{F}_\tau) = \mathrm{E}(X_{\tau_n}|\mathscr{F}_\tau) \underset{n\to\infty}{\to} \mathrm{E}(X_\tau|\mathscr{F}_\tau) = X_\tau\,.$$

The last equality in the relation above follows from the fact that $X_\tau$ is $\mathscr{F}_\tau$-measurable, as a consequence of Propositions 2.1 and 3.6.

**5.15** We have, recalling again Remark 4.5,

$$\begin{aligned}
\mathrm{E}[M_t\,|\,\mathscr{F}_s] &= \mathrm{e}^{\lambda t}\mathrm{E}(B_t\,|\,\mathscr{F}_s) - \lambda\int_0^t \mathrm{e}^{\lambda u}\mathrm{E}[B_u\,|\,\mathscr{F}_s]\,du\\
&= \mathrm{e}^{\lambda t}B_s - \lambda\int_0^s \mathrm{e}^{\lambda u}B_u\,du - \lambda\int_s^t \mathrm{e}^{\lambda u}B_s\,du\\
&= \mathrm{e}^{\lambda t}B_s - \lambda\int_0^s \mathrm{e}^{\lambda u}B_u\,du - (\mathrm{e}^{\lambda t}-\mathrm{e}^{\lambda s})B_s = M_s\;.
\end{aligned}$$

which proves the martingale property. Let us write down the increments of $M$, trying to express them in terms of the increments of the Brownian motion. We have

$$\begin{aligned}
M_t - M_s &= \mathrm{e}^{\lambda t}B_t - \mathrm{e}^{\lambda s}B_s - \lambda\int_s^t \mathrm{e}^{\lambda u}B_u\,du\\
&= \mathrm{e}^{\lambda t}(B_t-B_s) + (\mathrm{e}^{\lambda t}-\mathrm{e}^{\lambda s})B_s - \lambda\int_s^t \mathrm{e}^{\lambda u}(B_u-B_s)\,du - \lambda\int_s^t \mathrm{e}^{\lambda u}B_s\,du\\
&= \mathrm{e}^{\lambda t}(B_t-B_s) - \lambda\int_s^t \mathrm{e}^{\lambda u}(B_u-B_s)\,du\;.
\end{aligned}$$

As $M_t - M_s$ is a function of the increments of $B$ after time $s$, it follows that it is independent of $\mathscr{F}_s$ (recall Remark 3.2).

**5.16**

a) Recalling the expression of the Laplace transform of Gaussian laws (Exercise 1.6),

$$\mathrm{E}[\mathrm{e}^{\lambda B_t-\nu t}] = \mathrm{e}^{(\frac{\lambda^2}{2}-\nu)t}$$

so that the required limit is equal to $+\infty$, 1 or 0, according as $\frac{\lambda^2}{2} > \nu$, $\frac{\lambda^2}{2} = \nu$ or $\frac{\lambda^2}{2} < \nu$.

b) Let $t > s$. With the typical method of factoring out the increment we have

$$\mathrm{E}(\mathrm{e}^{\lambda B_t-\nu t}\,|\,\mathscr{F}_s) = \mathrm{e}^{\lambda B_s-\nu t}\mathrm{E}(\mathrm{e}^{\lambda(B_t-B_s)}\,|\,\mathscr{F}_s) = \mathrm{e}^{\lambda B_s+\frac{\lambda^2}{2}(t-s)-\nu t}$$

so that, if $\frac{\lambda^2}{2} = \nu$, $(X_t)_t$ is a martingale. Conversely, it will be a supermartingale if and only if

$$\frac{\lambda^2}{2}(t-s) - \nu t \le -\nu s,$$

i.e. if $\frac{\lambda^2}{2} \le \nu$. The same argument also allows us to prove the result in the submartingale case.

c) We have

$$X_t^\alpha = e^{\alpha\lambda B_t - \alpha\nu t} .$$

It is obvious that we can choose $\alpha > 0$ small enough so that $\frac{1}{2}\alpha^2\lambda^2 < \alpha\nu$, so that, for these values of $\alpha$, $X^\alpha$ turns out to be a supermartingale, thanks to b). Being positive, it converges a.s. Let us denote by $Z$ its limit: as by Fatou's lemma

$$\lim_{t\to\infty} E(X_t^\alpha) = \lim_{t\to\infty} E(e^{\alpha\lambda B_t - \alpha\nu t}) = \lim_{t\to\infty} e^{(\frac{1}{2}\alpha^2\lambda^2 - \alpha\nu)t} = 0 ,$$

we have

$$E(Z) \le \varliminf_{t\to+\infty} E(X_t^\alpha) = 0 ,$$

from which, as $Z \ge 0$, we deduce $Z = 0$ a.s. Therefore $\lim_{t\to\infty} X_t^\alpha = 0$ and also $\lim_{t\to\infty} X_t = 0$ a.s.

d) The integral

$$A_\infty = \int_0^{+\infty} e^{\lambda B_s - \nu s}\, ds$$

is well defined (possibly $= +\infty$), the integrand being positive. By Fubini's theorem

$$E(A_\infty) = E\Big[\int_0^{+\infty} e^{\lambda B_s - \nu s}\, ds\Big] = \int_0^{+\infty} E\big[e^{\lambda B_s - \nu s}\big]\, ds\Big] = \int_0^{+\infty} e^{(\frac{\lambda^2}{2} - \nu)s}\, ds .$$

Therefore, if $\frac{\lambda^2}{2} < \nu$, $E(A_\infty) = (\nu - \frac{\lambda^2}{2})^{-1}$. If $\frac{\lambda^2}{2} \ge \nu$ then $E[A_\infty] = +\infty$.

If $W_t = \frac{1}{\lambda} B_{\lambda^2 t}$, we know, thanks to the scaling property, that $(W_t)_t$ is also a Brownian motion. Therefore the r.v.'s

$$\int_0^{+\infty} e^{\lambda B_s - \nu s}\, ds$$

and

$$\int_0^{+\infty} e^{\lambda W_s - \nu s}\, ds = \int_0^{+\infty} e^{B_{\lambda^2 s} - \nu s}\, ds$$

have the same law. Now just make the change of variable $t = \lambda^2 s$.

• Note the apparent contradiction: we have $\lim_{t\to\infty} X_t = 0$ *for every* value of $\lambda \in \mathbb{R}$, $\nu > 0$, whereas, for $\frac{1}{2}\lambda^2 > \nu$, $\lim_{t\to\infty} E(X_t) = +\infty$.

**5.17**

a) Let us consider the two possibilities: if $\tau_x = +\infty$, then

$$\lim_{t\to+\infty} M_{t\wedge\tau_x} = \lim_{t\to+\infty} M_t = 0$$

whereas if $\tau_x < +\infty$, then $M_{t\wedge\tau_x} \to M_{\tau_x} = x$ as $t \to +\infty$ (the martingale is continuous). Putting together the two possible cases we obtain (5.34).

b) As the stopping time $t \wedge \tau_x$ is bounded, by the stopping theorem, Theorem 5.13, $\mathrm{E}[M_{t\wedge\tau_x}] = \mathrm{E}[M_0] = 1$. As the stopped martingale $(M_{t\wedge\tau_x})_t$ is bounded (actually $0 \le M_{t\wedge\tau_x} \le x$), by Lebesgue's theorem,

$$1 = \lim_{t\to+\infty} \mathrm{E}(M_{t\wedge\tau_x}) = \mathrm{E}[x1_{\{\tau_x<+\infty\}}] = x\mathrm{P}(\tau_x < +\infty) ,$$

i.e. $\mathrm{P}(\tau_x < +\infty) = \frac{1}{x}$.

c) Just observe that, for $x > 1$, $\{M^* \ge x\} = \{\tau_x < +\infty\}$, so that

$$\mathrm{P}(M^* \ge x) = \mathrm{P}(\tau_x < +\infty) = \frac{1}{x} \cdot \tag{S.20}$$

d) By Exercise 5.9 we know that $(M_t)_t$ is a continuous martingale such that $M_0 = 1$ and vanishing as $t \to +\infty$. Thanks to (S.20)

$$\mathrm{P}(X^* \le x) = \mathrm{P}(2\theta X^* \le 2\theta x) = \mathrm{P}\Big(\sup_{t\ge0}(2\theta B_t - 2\theta^2 t) \le 2\theta x\Big)$$

$$= \mathrm{P}\Big(\sup_{t\ge0} \mathrm{e}^{2\theta B_t - 2\theta^2 t} \le \mathrm{e}^{2\theta x}\Big) = \mathrm{P}(M^* \le \mathrm{e}^{2\theta x}) = 1 - \mathrm{P}(M^* > \mathrm{e}^{2\theta x}) = 1 - \mathrm{e}^{-2\theta x} .$$

We recognize the partition function of an exponential law with parameter $2\theta$, therefore $X^*$ is exponential with parameter $2\theta$.

- This is an example of a *ruin problem*. The process $X$ is a *ruin process* modeling, for instance, the net loss of an insurance company, coming from the difference between the claims and the premia. In the average $X_t$ tends to $-\infty$, as the premia are computed in order to make a profit, but the outflow produced by the claims is a random quantity which can (hopefully with small probability) become large.

  If $R$ denotes the reserves of the company at time 0, the event $\{X^* > R\}$ corresponds to the fact that the company is unable to face its obligations, hence of being bankrupt.

  Investigating this kind of situation is therefore of interest and Exercise 5.17 is a first example. The process $X_t = B_t - \theta t$ is not really a good ruin process and more realistic ones have been developed (see Asmussen 2000 if you want to go further). It is nevertheless a good first look.

**5.18**

a) If $s = \frac{t}{1-t}$, then $t = \frac{s}{s+1}$ and $1 - t = \frac{1}{s+1}$. (5.35) therefore holds if and only if the process $B_s = (1+s)X_{\frac{s}{s+1}}$ is a Brownian motion. As it is obviously a centered Gaussian process that vanishes for $s = 0$, we just have to prove that $\mathrm{E}(B_sB_t) = s \wedge t$. Note that $s \mapsto \frac{s}{s+1} = 1 - \frac{1}{s+1}$ is increasing. Therefore if $s \le t$, then also $\frac{s}{s+1} \le \frac{t}{t+1}$ and, recalling the form of the covariance function of the Brownian bridge,

$$\mathrm{E}(B_sB_t) = (1+s)(1+t)\mathrm{E}\big(X_{\frac{s}{s+1}}X_{\frac{t}{t+1}}\big) = (1+s)(1+t)\frac{s}{s+1}\Big(1 - \frac{t}{t+1}\Big) = s$$

and therefore $B$ is a Brownian motion.

b) We have, with the change of variable $s = \frac{t}{1-t}$,

$$\begin{aligned}\mathrm{P}\Big(\sup_{0\le t\le 1} X_t > a\Big) &= \mathrm{P}\Big(\sup_{0\le t<1} (1-t)B_{\frac{t}{1-t}} > a\Big) = \mathrm{P}\Big(\sup_{s>0} \frac{1}{s+1}B_s > a\Big)\\ &= \mathrm{P}\Big(\sup_{s>0} \frac{1}{s+1}(B_s - (s+1)a) > 0\Big) = \mathrm{P}\Big(\sup_{s>0} B_s - sa > a\Big)\,.\end{aligned}$$

Thanks to Exercise 5.17 the r.v. $\sup_{s>0} B_s - sa$ has an exponential law with parameter $2a$, therefore

$$\mathrm{P}\Big(\sup_{0\le t\le 1} X_t > a\Big) = \mathrm{e}^{-2a^2}$$

and the partition function of $\sup_{0\le t\le 1} X_t$ is $F(x) = 1 - \mathrm{e}^{-2x^2}$ for $x \ge 0$. Taking the derivative, the corresponding density is $f(x) = 4x\mathrm{e}^{-2x^2}$ for $x \ge 0$.

**5.19**

a) As computed in Example 5.3 (here $a = x$, $b = 1$) $\mathrm{P}(B_\tau = -x) = \frac{1}{1+x}$.

b) The important observation is that if $Z \ge x$, i.e.

$$\min_{t\le\tau_1} B_t \le -x\,,$$

then $B$ has gone below level $-x$ before passing at 1, so that $B_\tau = -x$. Therefore, by a),

$$\mathrm{P}(Z \ge x) = \mathrm{P}(B_\tau = -x) = \frac{1}{1+x}\,,$$

i.e. the partition function of $Z$ is $\mathrm{P}(Z \le x) = 1 - \frac{1}{1+x}$. Taking the derivative, the density of $Z$ is

$$f_Z(x) = \frac{1}{(1+x)^2}\,, \quad x > 0\,.$$

The expectation of $Z$ would be

$$\int_0^{+\infty} \frac{x}{(1+x)^2}\,dx$$

but the integral does not converge, as the integrand behaves like $\frac{1}{x}$ as $x \to +\infty$. Therefore $Z$ does not have finite expectation.

**5.20**

a) We have $M_t = \mathrm{e}^{-2\mu B_t - 2\mu^2 t}$ so that this is the martingale of Example 5.2 for $\theta = -2\mu$.

b1) By the Iterated Logarithm Law…

b2) By the stopping theorem

$$\mathrm{E}(M_{\tau\wedge t}) = \mathrm{E}(M_0) = 1\ . \tag{S.21}$$

As $-a \le X_{\tau\wedge t} \le b$, $t \mapsto M_{\tau\wedge t}$ is bounded and we can apply Lebesgue's theorem taking the limit as $t \to +\infty$ in (S.21), which gives

$$\begin{aligned} 1 = \mathrm{E}(M_\tau) = \mathrm{E}(\mathrm{e}^{-2\mu X_\tau}) &= \mathrm{e}^{2\mu a}\mathrm{P}(X_\tau = -a) + \mathrm{e}^{-2\mu b}\mathrm{P}(X_\tau = b) \\ &= \mathrm{e}^{2\mu a}(1 - \mathrm{P}(X_\tau = b)) + \mathrm{e}^{-2\mu b}\mathrm{P}(X_\tau = b) \end{aligned}$$

hence

$$\mathrm{P}(X_\tau = b) = \frac{1 - \mathrm{e}^{2\mu a}}{\mathrm{e}^{-2\mu b} - \mathrm{e}^{2\mu a}}$$

and the limit as $\mu \to +\infty$ of this probability is equal to 1.

**5.21**

a) We have

$$\begin{aligned} \mathrm{E}[B_t^3 - 3tB_t\,|\,\mathscr{F}_s] &= \mathrm{E}[(B_s + (B_t - B_s))^3 - 3t(B_s + (B_t - B_s))\,|\,\mathscr{F}_s] \\ &= \mathrm{E}[B_s^3 + 3B_s^2(B_t - B_s) + 3B_s(B_t - B_s)^2 + (B_t - B_s)^3\,|\,\mathscr{F}_s] - 3tB_s \\ &= B_s^3 + 3(t-s)B_s - 3tB_s = X_s\ . \end{aligned}$$

b) By the stopping theorem applied to the bounded stopping time $\tau \wedge t$,

$$0 = \mathrm{E}(X_{\tau\wedge t}) = \mathrm{E}(B^3_{\tau\wedge t}) - 3\mathrm{E}((\tau \wedge t)B_{\tau\wedge t})\ ,$$

i.e.

$$\mathrm{E}(B^3_{\tau\wedge t}) = 3\mathrm{E}[(\tau \wedge t)B_{\tau\wedge t}]\ .$$

Now $\mathrm{E}(B^3_{\tau\wedge t}) \to_{t\to+\infty} \mathrm{E}(B^3_\tau)$, as $|B^3_{\tau\wedge t}| \le \max(a,b)^3$ and we can apply Lebesgue's theorem. The same argument allows us to take the limit on the right-hand side, since $|(\tau \wedge t)B_{\tau\wedge t}| \le \tau \max(a,b)$ and we know (Exercise 5.31) that $\tau$ is integrable. We can therefore take the limit and obtain

$$\mathrm{E}(\tau B_\tau) = \frac{1}{3}\,\mathrm{E}(B_\tau^3) = \frac{1}{3}\Big(\frac{b^3a}{a+b} - \frac{a^3b}{a+b}\Big) = \frac{1}{3}\Big(\frac{ab}{a+b}\,(b^2-a^2)\Big) = \frac{1}{3}\,ab(b-a)\,.$$

We have actually seen in Exercise 5.31 that $\mathrm{P}(B_\tau = -a) = \frac{b}{a+b}$ and $\mathrm{P}(B_\tau = b) = \frac{a}{a+b}$. Therefore, recalling that $\mathrm{E}(B_\tau) = 0$, we find

$$\mathrm{Cov}(\tau, B_\tau) = \mathrm{E}(\tau B_\tau) = \frac{1}{3}\,ab(b-a)\,.$$

Note that the covariance is equal to zero if $a = b$, in agreement with the fact that, if $a = b$, then $\tau$ and $B_\tau$ are independent (Exercise 3.18).

**5.22** We must prove that, if $s \le t$,

$$\mathrm{E}[M_tN_t1_A] = \mathrm{E}[M_sN_s1_A] \tag{S.22}$$

for every $A \in \mathscr{H}_s$ or at least for every $A$ in a subclass $\mathscr{C}_s \subset \mathscr{H}_s$, which generates $\mathscr{H}_s$, contains $\Omega$ and is stable with respect to finite intersections (this is Remark 4.2). One can consider the class of the events of the form $A_1 \cap A_2$ with $A_1 \in \mathscr{M}_s$, $A_2 \in \mathscr{N}_s$. This class is stable with respect to finite intersections and contains both $\mathscr{M}_s$ (choosing $A_2 = \Omega$) and $\mathscr{N}_s$ (with $A_1 = \Omega$). As $M_t1_{A_1}$ and $N_t1_{A_2}$ are independent (the first is $\mathscr{M}_t$-measurable whereas the second one is $\mathscr{N}_t$-measurable) we have

$$\begin{aligned}\mathrm{E}[M_tN_t1_{A_1\cap A_2}] &= \mathrm{E}[M_t1_{A_1}N_t1_{A_2}] = \mathrm{E}[M_t1_{A_1}]\mathrm{E}[N_t1_{A_2}] = \mathrm{E}[M_s1_{A_1}]\mathrm{E}[N_s1_{A_2}]\\ &= \mathrm{E}[M_s1_{A_1}N_s1_{A_2}] = \mathrm{E}[M_sN_s1_{A_1\cap A_2}]\,,\end{aligned}$$

hence (S.22) is satisfied for the events of the class $\mathscr{C}_s$.

**5.23**

a) If $(M_t^2)_t$ is a martingale we have, for every $s \le t$,

$$\begin{aligned}0 = \mathrm{E}\big[M_t^2 - M_s^2\,|\,\mathscr{F}_s\big] &= \mathrm{E}\big[(M_t - M_s + M_s)^2 - M_s^2\,|\,\mathscr{F}_s\big]\\ \mathrm{E}\big[(M_t-M_s)^2 + 2M_s(M_t-M_s)\,|\,\mathscr{F}_s\big] &= \mathrm{E}\big[(M_t-M_s)^2\,|\,\mathscr{F}_s\big]\,.\end{aligned}$$

Now, taking the expectation, $\mathrm{E}[(M_t - M_s)^2] = 0$, i.e. $M_t = M_s$ a.s. Let us determine a negligible event $N$ such that, for every $\omega \notin N$, $B_t = B_s$ for every $t, s$. For $q, r \in \mathbb{Q}$ let $N_{q,r}$ be the negligible event such that for $\omega \notin N_{q,r}$ $M_q = M_r$. It is easy to verify that, as $M$ is supposed to be continuous, $N = \bigcup_{q,r\in\mathbb{Q}} N_{q,r}$ satisfies this requirement.

b1) The function $x \mapsto |x|^{p'}$ being convex, $(|M_t|^{p'})_t$ is a submartingale and, for $s \le t$,

$$\mathrm{E}\big[|M_t|^{p'}\,|\,\mathscr{F}_s\big] \ge |M_s|^{p'}\,. \tag{S.23}$$

If $p' < p$ then $x \mapsto |x|^{p/p'}$ is also convex so that

$$|M_s|^p = \mathrm{E}\big[|M_t|^p\,|\mathscr{F}_s\big] = \mathrm{E}\big[(|M_t|^{p'})^{p/p'}\,|\mathscr{F}_s\big] \ge \big(\mathrm{E}\big[|M_t|^{p'}\,|\mathscr{F}_s\big]\big)^{p/p'}\ ,$$

which gives $\mathrm{E}\big[|M_t|^{p'}\,|\mathscr{F}_s\big] \le |M_s|^{p'}$ and this inequality together with (S.23) allows us to conclude the proof.

b2) For $p = 2$ this is already proved. If for $p > 2$ $(|M_t|^p)_t$ is a martingale, then it is a also martingale for $p = 2$ by b1) and is therefore constant thanks to a).

**5.24**

a) With the usual method of factoring out the increment we have for $s \le t$,

$$\begin{aligned}\mathrm{E}\big[B_i(t)B_j(t)\,|\mathscr{F}_s\big] &= \mathrm{E}\big[\big(B_i(s) + (B_i(t) - B_i(s)\big)\big(B_j(s) + (B_j(t) - B_j(s)\big)\,|\mathscr{F}_s\big]\\ &= \mathrm{E}\big[B_i(s)B_j(s) + B_i(s)(B_j(t) - B_j(s)) + B_j(s)(B_i(t) - B_i(s))\\ &\qquad + (B_i(t) - B_i(s))(B_j(t) - B_j(s))\,|\mathscr{F}_s\big]\,.\end{aligned}$$

Now just observe that, the increments being independent of $\mathscr{F}_s$,

$$\begin{aligned}\mathrm{E}\big[B_i(s)(B_j(t) - B_j(s))\,|\mathscr{F}_s\big] &= B_i(s)\mathrm{E}\big[B_j(t) - B_j(s)\,|\mathscr{F}_s\big] = 0\\ \mathrm{E}\big[B_j(s)(B_i(t) - B_i(s))\,|\mathscr{F}_s\big] &= B_j(s)\mathrm{E}\big[B_i(t) - B_i(s)\,|\mathscr{F}_s\big] = 0\\ \mathrm{E}\big[(B_i(t) - B_i(s))(B_j(t) - B_j(s))\,|\mathscr{F}_s\big] &= \mathrm{E}\big[(B_i(t) - B_i(s))(B_j(t) - B_j(s))\big] = 0\end{aligned}$$

and therefore $\mathrm{E}[B_i(t)B_j(t)\,|\mathscr{F}_s] = B_i(s)B_j(s)$.

b) This is proved using the same idea as in a), only more complicated to express: one must factor out $B_{i_1}(t)\dots B_{i_d}(t)$ as

$$(B_{i_1}(s) + (B_{i_1}(t) - B_{i_1}(s))\dots(B_{i_d}(s) + (B_{i_d}(t) - B_{i_d}(s)) = B_{i_1}(s)\dots B_{i_d}(s) + \dots$$

where the rightmost $\dots$ denotes a r.v. which is the product of some $B_{i_k}(s)$ (which are already $\mathscr{F}_s$-measurable) and of some (at least one) terms of the kind $B_{i_k}(t) - B_{i_k}(s)$. These are centered r.v.'s which are, moreover, independent with respect to $\mathscr{F}_s$. The conditional expectation of their product with respect to $\mathscr{F}_s$ is therefore equal to 0, so that

$$\mathrm{E}[B_{i_1}(t)\dots B_{i_d}(t)\,|\mathscr{F}_s] = B_{i_1}(s)\dots B_{i_d}(s)\ ,$$

hence $t \mapsto B_{i_1}(t)\dots B_{i_d}(t)$ is a martingale.

c) $\det B_t$ is a sum of terms of the form $B_{1,j_1}B_{2,j_2}\dots B_{m,j_m}$, with $j_1,\dots,j_m$ being distinct indices. Thanks to b) every such term is a martingale. $\det B_t$ is therefore a sum of martingales, hence a martingale itself.

Moreover, $Y_t = \det(B_tB_t^*) = \det B_t \times \det B_t^* = (\det B_t)^2 = X_t^2$, so that $Y$ is the square of a martingale and therefore a submartingale.

Throughout this part of the solution we neglected to prove that both $X_t$ and $Y_t$ are integrable, but this immediate.

- Note that each of the processes $t \mapsto B_i(t)$ is a martingale with respect to the filtration $t \mapsto \sigma(B_i(u), u \le t)$. As these are independent (Remark 3.2 b)), from Exercise 5.22 it follows immediately that $t \mapsto B_i(t)B_j(t)$ is a martingale with respect to the filtration $\mathscr{G}_{i,j}(t) = \sigma(B_i(u), u \le t) \vee \sigma(B_j(u), u \le t)$.

**5.25**

a) If $t \le T$, thanks to the freezing Lemma 4.1 we have, as $B_T - B_t$ is independent of $\mathscr{F}_t$,

$$\mathrm{E}[1_{\{B_T>0\}} | \mathscr{F}_t] = \mathrm{E}[1_{\{B_t+B_T-B_t>0\}} | \mathscr{F}_t] = \Psi(B_t, t) ,$$

where, denoting by $\Phi$ the partition function of an $N(0,1)$ law,

$$\Psi(x,t) = \mathrm{E}[1_{\{x+B_T-B_t>0\}}] = \mathrm{P}(B_T - B_t > -x) = \Phi\Big(\frac{x}{\sqrt{T-t}}\Big) .$$

We use here the fact that $B_T - B_t \sim \sqrt{T-t}\, Z$ with $Z \sim N(0,1)$. Hence

$$\mathrm{E}[1_{\{B_T>0\}} | \mathscr{F}_t] = \Phi\Big(\frac{B_t}{\sqrt{T-t}}\Big) .$$

b) From a) $M_t = \mathrm{E}[1_{\{B_T>0\}} | \mathscr{F}_t]$ hence $M$ is a uniformly integrable martingale. As

$$\lim_{t\to T-} \frac{x}{\sqrt{T-t}} = \begin{cases} +\infty & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -\infty & \text{if } x < 0 \end{cases}$$

we have

$$\lim_{t\to T-} \Phi\Big(\frac{x}{\sqrt{T-t}}\Big) = \begin{cases} 1 & \text{if } x > 0 \\ \frac{1}{2} & \text{if } x = 0 \\ 0 & \text{if } x < 0 . \end{cases}$$

Hence (the event $\{B_T = 0\}$ has probability 0)

$$\lim_{t\to T-} \Phi\Big(\frac{B_t}{\sqrt{T-t}}\Big) = 1_{\{B_T>0\}} \qquad \text{a.s.}$$

**5.26**

a) The integral converges absolutely as $u \mapsto \frac{1}{\sqrt{u}}$ is integrable at the origin and the path $t \mapsto B_t(\omega)$ is bounded (there is no need here of the Iterated Logarithm Law…).

b) In order to prove that it is a Gaussian process, we use the usual technique of approximating the integral with Riemann sums, which are obviously Gaussian, using the property of stability of the Gaussian r.v.'s in the limit (Proposition 1.9). In this case it is not obvious that the Riemann sums converge to the integral, as the integrand diverges at 0. This argument, however, allows us to see that, for every $\varepsilon > 0$, the process $\widetilde{X}_t = 0$ for $t < \varepsilon$ and

$$\widetilde{X}_t = \int_\varepsilon^t \frac{B_u}{\sqrt{u}}\, du, \qquad t \geq \varepsilon$$

is Gaussian. One then takes the limit as $\varepsilon \searrow 0$ and uses again the properties of stability of Gaussianity with respect to limits in law. As the r.v.'s $X_t$ are centered, for the covariance we have, for $s \leq t$,

$$\begin{aligned}\mathrm{Cov}(X_s, X_t) = \mathrm{E}[X_s X_t] &= \int_0^t du \int_0^s \frac{\mathrm{E}[B_u B_v]}{\sqrt{u}\sqrt{v}}\, dv = \int_0^t du \int_0^s \frac{u \wedge v}{\sqrt{u}\sqrt{v}}\, dv \\ &= \int_0^s du \int_0^s \frac{u \wedge v}{\sqrt{u}\sqrt{v}}\, dv + \int_s^t du \int_0^s \frac{u \wedge v}{\sqrt{u}\sqrt{v}}\, dv\ .\end{aligned}$$

The last two integrals are computed separately: keeping carefully in mind which among $u$ and $v$ is smaller, we have for the second one

$$\begin{aligned}\int_s^t du \int_0^s \frac{u \wedge v}{\sqrt{u}\sqrt{v}}\, dv &= \int_s^t du \int_0^s \frac{v}{\sqrt{u}\sqrt{v}}\, dv = \int_s^t \frac{1}{\sqrt{u}}\, du \int_0^s \sqrt{v}\, dv \\ &= \frac{4}{3}\, s^{3/2}(\sqrt{t} - \sqrt{s})\ .\end{aligned}$$

Whereas for the first one

$$\begin{aligned}\int_0^s du \int_0^s \frac{u \wedge v}{\sqrt{u}\sqrt{v}}\, dv &= \int_0^s \frac{1}{\sqrt{u}}\, du \int_0^u \sqrt{v}\; dv + \int_0^s \sqrt{u}\, du \int_u^s \frac{1}{\sqrt{v}}\, dv \\ &= \frac{2}{3}\int_0^s u\, du + 2\int_0^s \sqrt{u}(\sqrt{s} - \sqrt{u})\, du = \frac{1}{3}\, s^2 + \frac{4}{3} s^2 - s^2 = \frac{2}{3}\, s^2\ .\end{aligned}$$

In conclusion

$$\mathrm{Cov}(X_s, X_t) = \frac{2}{3}\, s^2 + \frac{4}{3}\, s^{3/2}(\sqrt{t} - \sqrt{s})\ .$$

c) The most simple and elegant argument consists in observing that $(X_t)_t$ is a square integrable continuous process vanishing at the origin. If it were a martingale the paths would be either identically zero or with infinite variation a.s. Conversely the paths are $C^1$. Therefore $(X_t)_t$ cannot be a martingale.

Alternatively one might also compute the conditional expectation directly and check the martingale property. We have, for $s \le t$.

$$\mathrm{E}(X_t\,|\,\mathscr{F}_s) = \mathrm{E}\Big(\int_0^t \frac{B_u}{\sqrt{u}}\,du\,\Big|\,\mathscr{F}_s\Big) = \underbrace{\int_0^s \frac{B_u}{\sqrt{u}}\,du}_{=X_s} + \mathrm{E}\Big(\int_s^t \frac{B_u}{\sqrt{u}}\,du\,\Big|\,\mathscr{F}_s\Big)\,. \tag{S.24}$$

For the last conditional expectation we can write, as described in Remark 4.5,

$$\mathrm{E}\Big(\int_s^t \frac{B_u}{\sqrt{u}}\,du\,\Big|\,\mathscr{F}_s\Big) = \int_s^t \frac{1}{\sqrt{u}}\,\mathrm{E}[B_u\,|\,\mathscr{F}_s]\,du = B_s \int_s^t \frac{1}{\sqrt{u}}\,du = 2(\sqrt{t}-\sqrt{s})B_s\,,$$

so that $\mathrm{E}(X_t\,|\,\mathscr{F}_s) = X_s + 2(\sqrt{t}-\sqrt{s})B_s$ and $X$ is not a martingale.

**5.27** Let us denote by $\mathrm{E}^{\mathrm{P}}$, $\mathrm{E}^{\mathrm{Q}}$ the expectations with respect to P and Q, respectively.

a) Recall that, by definition, for $A \in \mathscr{F}_s$, $\mathrm{Q}(A) = \mathrm{E}^{\mathrm{P}}(Z_s 1_A)$. Let $s \le t$. We must prove that, for every $A \in \mathscr{F}_s$, $\mathrm{E}^{\mathrm{P}}(Z_t 1_A) = \mathrm{E}^{\mathrm{P}}(Z_s 1_A)$. But as $A \in \mathscr{F}_s \subset \mathscr{F}_t$, both these quantities are equal to $\mathrm{Q}(A)$.

b) We have $\mathrm{Q}(Z_t = 0) = \mathrm{E}^{\mathrm{P}}(Z_t 1_{\{Z_t=0\}}) = 0$ and therefore $Z_t > 0$ Q-a.s. Moreover, as $\{Z_t > 0\} \subset \{Z_s > 0\}$ a.s. if $s \le t$ (Exercise 5.8), then for every $A \in \mathscr{F}_s$,

$$\begin{aligned}\mathrm{E}^{\mathrm{Q}}(1_A Z_t^{-1}) = \mathrm{E}^{\mathrm{Q}}(1_{A\cap\{Z_t>0\}} Z_t^{-1}) &= \mathrm{P}(A\cap\{Z_t>0\}) \le \mathrm{P}(A\cap\{Z_s>0\})\\ &= \mathrm{E}^{\mathrm{Q}}(1_A Z_s^{-1})\end{aligned}$$

and therefore $(Z_t^{-1})_t$ is a Q-supermartingale.

c) Let us assume $\mathrm{P} \ll \mathrm{Q}$: this means that $\mathrm{P}(A) = 0$ whenever $\mathrm{Q}(A) = 0$. Therefore also $\mathrm{P}(Z_t = 0) = 0$ and

$$\mathrm{E}^{\mathrm{Q}}(Z_t^{-1}) = \mathrm{E}^{\mathrm{P}}(Z_t Z_t^{-1}) = 1\,.$$

The Q-supermartingale $(Z_t)_t$ therefore has constant expectation and is a Q-martingale by the criterion of Exercise 5.1. Alternatively, observe that

$$\frac{d\mathrm{P}_{|\mathscr{F}_t}}{d\mathrm{Q}_{|\mathscr{F}_t}} = Z_t^{-1}$$

and $(Z_t^{-1})_t$ is a martingale by a).

**5.28**

a) Note that the $\sigma$-algebra $\mathscr{G}_{n+1}$ is generated by the same r.v.'s $X_{sk/2^n}$ that generate $\mathscr{G}_n$ and some more in addition, therefore $\mathscr{G}_{n+1} \supset \mathscr{G}_n$.

Let, moreover, $\mathscr{G}' = \bigvee_{n\ge1}\mathscr{G}_n$. As $\mathscr{G}_n \subset \mathscr{G}$ for every $n$, clearly $\mathscr{G}' \subset \mathscr{G}$. Moreover, the r.v.'s $X_{sk/2^n}$, $k = 1,\dots,2^n$, $n = 1,2,\dots$, are all $\mathscr{G}'$-measurable. Let now $u \le s$. As the times of the form $sk/2^n$, $k = 1,\dots,2^n$, $n = 1,2,\dots$, are dense in $[0,s]$, there exists a sequence $(s_n)_n$ of times of this form such that $s_n \to u$ as $n \to \infty$. As the process $(X_t)_t$ is assumed to be continuous, $X_{s_n} \to X_u$ and $X_u$ turns out to be $\mathscr{G}'$-measurable for every $u \le s$. Therefore $\mathscr{G}' \supset \mathscr{G}$, hence $\mathscr{G}' = \mathscr{G}$.

b1) The sequence $(Z_n)_n$ is a $(\mathscr{G}_n)_n$-martingale: as $\mathscr{G}_{n+1} \supset \mathscr{G}_n$,

$$Z_n = \mathrm{E}[X_t|\mathscr{G}_n] = \mathrm{E}[\mathrm{E}(X_t|\mathscr{G}_{n+1})|\mathscr{G}_n] = \mathrm{E}[Z_{n+1}|\mathscr{G}_n]\ .$$

As $\mathrm{E}[Z_n^2] \le \mathrm{E}[X_t^2]$ by Jensen's inequality (Proposition 4.2 d) with $\Phi(x) = x^2$), the martingale $(Z_n)_n$ is bounded in $L^2$ and converges a.s. and in $L^2$ by Theorem 5.6. Its limit, by Proposition 5.5, is equal to $\mathrm{E}[X_t|\mathscr{G}]$ thanks to the relation $\mathscr{G} = \bigvee_{n\ge1}\mathscr{G}_n$ that was proved in a).

b2) If $X_t$ is only assumed to be integrable, then $(Z_n)_n$ is again a uniformly integrable martingale, thanks to Proposition 5.4. Hence it converges a.s. and in $L^1$ to $\mathrm{E}[X_t|\mathscr{G}]$. The difference is that in the situation b1) we can claim that $\mathrm{E}[X_t|\mathscr{G}_n]$ converges to a r.v. that is the best approximation in $L^2$ of $X_t$ by a $\mathscr{G}$-measurable r.v. (Remark 4.3), whereas now this optimality claim cannot be made.

**5.29**

a) This is the stopping theorem applied to the bounded stopping times $\tau \wedge t$ and to the martingales $(B_t)_t$ and $(B_t^2 - t)_t$.

b) We have, as $t \to +\infty$, $\tau \wedge t \nearrow \tau$ and $B_{\tau\wedge t} \to B_\tau$. By Beppo Levi's theorem $\mathrm{E}[\tau \wedge t] \nearrow \mathrm{E}[\tau] < +\infty$ ($\tau$ is integrable by assumption), so that

$$\mathrm{E}[B_{\tau\wedge t}^2] \le \mathrm{E}[\tau] < +\infty$$

for every $t \ge 0$. Therefore the martingale $(B_{\tau\wedge t})_t$ is bounded in $L^2$ and we can apply Doob's inequality, which gives

$$\mathrm{E}\Big[\sup_{t\ge0} B_{\tau\wedge t}^2\Big] \le 4\sup_{t\ge0}\mathrm{E}[B_{\tau\wedge t}^2] \le 4\mathrm{E}[\tau] < +\infty\ ,$$

i.e. (5.36).

c) Thanks to (5.36) the r.v. $B^* = \sup_{t\ge0} B_{\tau\wedge t}$ is square integrable and we have $|B_{\tau\wedge t}| \le B^*$ and $B_{\tau\wedge t}^2 \le {B^*}^2$. These relations allow us to apply Lebesgue's theorem and obtain

$$\begin{aligned}\mathrm{E}[B_\tau] &= \lim_{t\to+\infty}\mathrm{E}[B_{\tau\wedge t}] = 0\\ \mathrm{E}[B_\tau^2] &= \lim_{t\to+\infty}\mathrm{E}[B_{\tau\wedge t}^2] = \lim_{t\to+\infty}\mathrm{E}[\tau\wedge t] = \mathrm{E}[\tau]\ .\end{aligned}$$

**5.30**

a) By the stopping theorem applied to the bounded stopping time $t \wedge \tau_a$ we have $\mathrm{E}[M_{t\wedge\tau_a}] = \mathrm{E}[M_0] = 1$. If, moreover, $\lambda \geq 0$ the martingale $(M_{\tau_a\wedge t})_t$ is bounded, as $\lambda B_{\tau_a\wedge t} \leq \lambda a$, so that $M_{\tau_a\wedge t} \leq \mathrm{e}^{\lambda a}$. We can therefore apply Lebesgue's theorem and, recalling that $\tau_a < +\infty$ a.s.,

$$\mathrm{E}[M_{\tau_a}] = \lim_{t\to+\infty} \mathrm{E}(M_{\tau_a\wedge t}) = 1\ .$$

If $\lambda < 0$ then $M_{\tau_a\wedge t}$ is not bounded and Lebesgue's theorem can no longer be applied.

b) The previous equality can be rewritten as

$$1 = \mathrm{E}[M_{\tau_a}] = \mathrm{E}[\mathrm{e}^{\lambda B_{\tau_a}-\frac{1}{2}\lambda^2\tau_a}] = \mathrm{E}[\mathrm{e}^{\lambda a-\frac{1}{2}\lambda^2\tau_a}]\ ,$$

i.e.

$$\mathrm{E}[\mathrm{e}^{-\frac{1}{2}\lambda^2\tau_a}] = \mathrm{e}^{-\lambda a}\ . \tag{S.25}$$

Now if $\theta = -\frac{1}{2}\lambda^2$, i.e. $\lambda = \sqrt{-2\theta}$ (recall that (S.25) was proved for $\lambda \geq 0$ so that we have to discard the negative root), (S.25) can be rewritten, for $\theta \leq 0$, as

$$\mathrm{E}[\mathrm{e}^{\theta\tau_a}] = \mathrm{e}^{-a\sqrt{-2\theta}}\ .$$

For $\theta > 0$ the Laplace transform is necessarily equal to $+\infty$ as a consequence of the fact that $\mathrm{E}(\tau_a) = +\infty$ (see Exercise 3.20), thanks to the inequality $\tau_a \leq \frac{1}{\theta}\mathrm{e}^{\theta\tau_a}$.

c) If $X_1, \dots, X_n$ are i.i.d. r.v.'s, having the same law as $\tau_a$, then the Laplace transform of $n^{-2}(X_1 + \dots + X_n)$ is, for $\theta \leq 0$,

$$\Big(\exp\Big(-a\sqrt{\tfrac{-2\theta}{n^2}}\Big)\Big)^n = \mathrm{e}^{-a\sqrt{-2\theta}}\ .$$

The laws of the r.v.'s $n^{-2}(X_1 + \dots + X_n)$ and $X_1$ have the same Laplace transform and therefore they coincide, as seen in Sect. 5.7; hence the law of $\tau_a$ is stable with exponent $\frac{1}{2}$.

**5.31** We know from Example 5.2 that, for $\lambda \in \mathbb{R}$, $M_t = \mathrm{e}^{\lambda B_t - \frac{\lambda^2}{2}t}$ is a martingale. The stopping theorem gives

$$1 = \mathrm{E}[M_{t\wedge\tau}] = \mathrm{E}[\mathrm{e}^{\lambda B_{t\wedge\tau}-\frac{\lambda^2}{2}(t\wedge\tau)}]\ .$$

As $|B_{t\wedge\tau}| \le a$ we can apply Lebesgue's theorem and take the limit as $t \to +\infty$. We obtain

$$1 = \mathrm{E}[\mathrm{e}^{\lambda B_\tau - \frac{\lambda^2}{2}\tau}]\ .$$

As $B_\tau$ and $\tau$ are independent, as a particular case, for $m = 1$, of Exercise 3.18 b), we have

$$1 = \mathrm{E}[\mathrm{e}^{\lambda B_\tau}]\mathrm{E}[\mathrm{e}^{-\frac{\lambda^2}{2}\tau}]\ . \tag{S.26}$$

$B_\tau$ takes the values $a$ and $-a$ with probability $\frac{1}{2}$, hence

$$\mathrm{E}[\mathrm{e}^{\lambda B_\tau}] = \frac{1}{2}\,(\mathrm{e}^{\lambda a} + \mathrm{e}^{-\lambda a}) = \cosh(\lambda a)\ ,$$

so that by (S.26)

$$\mathrm{E}[\mathrm{e}^{-\frac{\lambda^2}{2}\tau}] = \frac{1}{\cosh(\lambda a)}$$

and we just have to put $\theta = \frac{\lambda^2}{2}$, i.e. $\lambda = \sqrt{2\theta}$.

**5.32**

a) In a way similar to Exercise 5.9 we can show that $M_t = \mathrm{e}^{i\lambda B_t + \frac{1}{2}\lambda^2 t}$ is a (complex) $(\mathscr{F}_t)_t$-martingale:

$$\begin{aligned}\mathrm{E}(M_t|\mathscr{F}_s) &= \mathrm{e}^{\frac{1}{2}\lambda^2 t}\mathrm{E}(\mathrm{e}^{i\lambda B_s}\mathrm{e}^{i\lambda(B_t-B_s)}|\mathscr{F}_s) = \mathrm{e}^{\frac{1}{2}\lambda^2 t}\mathrm{e}^{i\lambda B_s}\mathrm{E}(\mathrm{e}^{i\lambda(B_t-B_s)})\\ &= \mathrm{e}^{\frac{1}{2}\lambda^2 t}\mathrm{e}^{i\lambda B_s}\mathrm{e}^{-\frac{1}{2}\lambda^2(t-s)} = M_s\ .\end{aligned}$$

This implies that the real part of $M$ is itself a martingale and note now that $\Re M_t = \cos(\lambda B_t)\,\mathrm{e}^{\frac{1}{2}\lambda^2 t}$. The martingale relation $\mathrm{E}[X_t|\mathscr{F}_s] = X_s$ can also be checked directly using the addition formulas for the cosine function, giving rise to a more involved computation.

b) By the stopping theorem applied to the bounded stopping time $\tau \wedge t$,

$$1 = \mathrm{E}(X_0) = \mathrm{E}[\cos(\lambda B_{\tau\wedge t})\,\mathrm{e}^{\frac{1}{2}\lambda^2(\tau\wedge t)}]\ . \tag{S.27}$$

But $|B_{\tau\wedge t}| < a$ hence, with the conditions on $\lambda$, $|\lambda B_{\tau\wedge t}| < \frac{\pi}{2}$ and recalling the behavior of the cosine function, $\cos(\lambda B_{\tau\wedge t}) \ge \cos(\lambda a) > 0$. We deduce that

$$\mathrm{E}[\mathrm{e}^{\frac{1}{2}\lambda^2(\tau\wedge t)}] \le \cos(\lambda a)^{-1}$$

and, letting $t \to \infty$, by Beppo Levi's Theorem the r.v. $e^{\frac{1}{2}\lambda^2\tau}$ is integrable. Thanks to the upper bound

$$0 < \cos(\lambda B_{\tau\wedge t})\, e^{\frac{1}{2}\lambda^2(\tau\wedge t)} \le e^{\frac{1}{2}\lambda^2\tau}$$

we can apply Lebesgue's theorem as $t \to +\infty$ in (S.27). As $\tau < +\infty$ a.s. and $|B_\tau| = a$, we obtain

$$\mathrm{E}(e^{\frac{1}{2}\lambda^2\tau}) = \frac{1}{\cos(\lambda a)}$$

from which, replacing $\theta = \frac{1}{2}\lambda^2$, (5.38) follows.

c) The expectation of $\tau$ can be obtained as the derivative of its Laplace transform at the origin; therefore

$$\mathrm{E}(\tau) = \frac{d}{d\theta}\,\frac{1}{\cos(a\sqrt{2\theta})}\Big|_{\theta=0} = a^2.$$

This result has already been obtained in Exercise 5.31 b). Finally, for $p \ge 0$ and $\varepsilon > 0$, we have $x^p \le c(\varepsilon, p)\, e^{\varepsilon x}$ for $x \ge 0$ (just compute the maximum of $x \mapsto x^p e^{-\varepsilon x}$, which is $c(\varepsilon, p) = p^p \varepsilon^{-p} e^{-p}$). Therefore $\tau^p \le c(\varepsilon, p)\, e^{\varepsilon\tau}$. Now just choose an $\varepsilon$ with $0 < \varepsilon < \frac{\pi^2}{8a^2}$.

d) First of all note that $\theta = 0$ is not an eigenvalue, as the solutions of $\frac{1}{2}u'' = 0$ are linear-affine functions, which cannot vanish at two distinct points unless they vanish identically. For $\theta \ne 0$ the general integral of the equation

$$\frac{1}{2}u'' - \theta u = 0$$

is $u(x) = c_1 e^{x\sqrt{2\theta}} + c_2 e^{-x\sqrt{2\theta}}$. The boundary conditions impose on the constants the conditions

$$\begin{aligned} c_1 e^{a\sqrt{2\theta}} + c_2 e^{-a\sqrt{2\theta}} &= 0 \\ c_1 e^{-a\sqrt{2\theta}} + c_2 e^{a\sqrt{2\theta}} &= 0\,. \end{aligned}$$

This system, in the unknowns $c_1, c_2$, has solutions different from $c_1 = c_2 = 0$ if and only if the determinant of the matrix

$$\begin{pmatrix} e^{a\sqrt{2\theta}} & e^{-a\sqrt{2\theta}} \\ e^{-a\sqrt{2\theta}} & e^{a\sqrt{2\theta}} \end{pmatrix}$$

vanishes, i.e. if and only if $e^{2a\sqrt{2\theta}} - e^{-2a\sqrt{2\theta}} = 0$ which gives $2a\sqrt{2\theta} = ik\pi$ for $k \in \mathbb{Z}$. Therefore the eigenvalues are the numbers

$$-\frac{k^2\pi^2}{8a^2} \qquad k = 1, 2, \dots$$

They are all negative and the largest one is of course $-\frac{\pi^2}{8a^2}$.

- This exercise completes Exercise 5.30 where the Laplace transform of $\tau$ was computed for negative reals. A more elegant way of obtaining (5.38) is to observe that the relation

$$\mathrm{E}[e^{-\theta\tau}] = \frac{1}{\cosh(a\sqrt{2\theta})}, \qquad \theta > 0,$$

which is obtained in Exercise 5.30, defines the Laplace transform of $\tau$ as the holomorphic function

$$z \mapsto \frac{1}{\cosh(a\sqrt{-2z})} \tag{S.28}$$

on $\Im z < 0$. But $z \mapsto \cosh(a\sqrt{-2z})$ is a holomorphic function on the whole complex plane which can be written as $z \mapsto \cos(a\sqrt{2z})$ for $\Re z > 0$. Hence the function in (S.28) is holomorphic as far as $\Re z$ is smaller than the first positive zero of $\theta \mapsto \cos(a\sqrt{2\theta})$ i.e. $\frac{\pi^2}{8a^2}$. Note that $z \mapsto \cosh(a\sqrt{-2z})$ is holomorphic on $\mathbb{C}$, even in the presence of the square root, because the power series development of cosh only contains even powers, so that the square root "disappears".

**5.33**

a) By the Iterated Logarithm Law ...

b) We have

$$M_t = e^{\lambda(B_t+\mu t)-(\frac{\lambda^2}{2}+\lambda\mu)t} = e^{\lambda B_t - \frac{\lambda^2}{2}t},$$

which is a martingale (an old acquaintance...). As $t \wedge \tau$ is a bounded stopping time, by the stopping theorem $\mathrm{E}[M_{t\wedge\tau}] = \mathrm{E}[M_0] = 1$.

c) As $\frac{\lambda^2}{2} + \lambda\mu \geq 0$,

$$\lambda X_{t\wedge\tau} - \Big(\frac{\lambda^2}{2} + \lambda\mu\Big)(t \wedge \tau) \leq \lambda X_{t\wedge\tau} \leq \lambda a\,.$$

Hence $0 \le M_{t\wedge\tau} \le e^{\lambda a}$. Moreover, $X_{t\wedge\tau} \to_{t\to+\infty} X_\tau = a$, hence

$$M_{t\wedge\tau} \underset{t\to+\infty}{\to} e^{\lambda a-(\frac{\lambda^2}{2}+\lambda\mu)\,\tau}$$

and by Lebesgue's theorem

$$1 = \lim_{t\to+\infty} E[M_{t\wedge\tau}] = E[e^{\lambda a-(\frac{\lambda^2}{2}+\lambda\mu)\,\tau}]\,,$$

i.e. (5.40).

d) Let $\lambda > 0$ such that $\frac{\lambda^2}{2} + \lambda\mu = \theta$, i.e. $\lambda = \sqrt{\mu^2+2\theta} - \mu$. With this choice of $\lambda$ (5.40) becomes

$$E[e^{-\theta\tau}] = e^{-a(\sqrt{\mu^2+2\theta}-\mu)}\,. \tag{S.29}$$

The expectation of $\tau$ is obtained by taking the derivative at 0 of the Laplace transform (S.29):

$$E(\tau) = -\frac{d}{d\theta}\, e^{-a(\sqrt{\mu^2+2\theta}-\mu)}\Big|_{\theta=0} = \frac{a}{\mu}\,\cdot$$

- Remark: the passage time of a Brownian motion with a positive drift through a positive level has a finite expectation, a very different behavior compared to the zero drift situation (Exercise 3.20).

**6.1**

a) As in Sect. 4.4 and Exercise 4.10 b), we have, for $s \le t$,

$$C_{t,s} = K_{t,s}K_{s,s}^{-1}\,, \qquad Y_{t,s} = X_t - K_{t,s}K_{s,s}^{-1}X_s\,.$$

$Y_{t,s}$ is a centered Gaussian r.v. with covariance matrix $K_{t,t} - K_{t,s}K_{s,s}^{-1}K_{s,t}$. Moreover, by the freezing Lemma 4.1, for every bounded measurable function $f$

$$E[f(X_t)|X_s] = E[f(C_{t,s}X_s + Y_{t,s})|X_s] = \Phi_f(X_s)\,,$$

where $\Phi_f(x) = E[f(C_{t,s}x + Y_{t,s})]$. Therefore the conditional law of $X_t$ given $X_s = x$ is the law of the r.v. $C_{t,s}x + Y_{t,s}$, i.e. is Gaussian with mean $C_{t,s}x = K_{t,s}K_{s,s}^{-1}x$ and covariance matrix $K_{t,t} - K_{t,s}K_{s,s}^{-1}K_{s,t}$.

b) The Markov property with respect to the natural filtration requires that, for every bounded measurable function $f : \mathbb{R}^m \to \mathbb{R}$,

$$E[f(X_t)|\mathscr{G}_s] = \int f(y)p(s,t,X_s,dy)\,.$$

Let us first determine what the transition function $p$ should be: it is the law of $X_t$ given $X_s = x$ which, as seen in a), is the law of $C_{t,s}x + Y_{t,s}$. In a) we have also proved that

$$E[f(X_t)|X_s] = \int f(y)p(s,t,X_s,dy) .$$

We must therefore prove that

$$E[f(X_t)|\mathscr{G}_s] = E[f(X_t)|X_s] . \tag{S.30}$$

Let us prove that (6.32) implies the independence of $Y_{t,s}$ and the $\sigma$-algebra $\mathscr{G}_s = \sigma(X_u, u \le s)$. This will imply, again by the freezing Lemma 4.1,

$$E[f(X_t)|\mathscr{G}_s] = E[f(C_{t,s}X_s + Y_{t,s})|\mathscr{G}_s] = \Phi_f(X_s) .$$

The covariances between $Y_{t,s}$ and $X_u$ are given by the matrix (recall that all these r.v.'s are centered)

$$E(Y_{s,t}X_u^*) = E(X_tX_u^*) - E(K_{t,s}K_{s,s}^{-1}X_sX_u^*) = K_{t,u} - K_{s,t}K_{s,s}^{-1}K_{s,u} ,$$

which vanishes if and only if (6.32) holds. Hence if (6.32) holds then the Markov property is satisfied with respect to the natural filtration.

Conversely, let us assume that $(X_t)_t$ is a Markov process with respect to its natural filtration. For $s \le t$ we have, by the Markov property,

$$E(X_t|\mathscr{G}_s) = E(X_t|X_s) = K_{t,s}K_{s,s}^{-1}X_s$$

hence

$$K_{t,u} = E(X_tX_u^*) = E[E(X_t|\mathscr{G}_s)X_u^*] = E(K_{t,s}K_{s,s}^{-1}X_sX_u^*) = K_{t,s}K_{s,s}^{-1}K_{s,u} ,$$

i.e. (6.32).

**6.2**

a) The simplest approach is to observe that the paths of the process $(X_t)_t$ have finite variation (they are even differentiable), whereas if it was a square integrable continuous martingale its paths would have infinite variation (Theorem 5.15). Otherwise one can compute the conditional expectation

$$E(X_t|\mathscr{F}_s)$$

and check that it does not coincide with $X_s$. This can be done directly:

$$E(X_t|\mathscr{F}_s) = X_s + E\Big(\int_s^t B_u\,du\,\Big|\,\mathscr{F}_s\Big) = X_s + \int_s^t E[B_u|\mathscr{F}_s]\,du = X_s + (t-s)B_s .$$

b) If $s \le t$,

$$\begin{aligned}
K_{t,s} &= \mathrm{Cov}(X_t, X_s) = \mathrm{E}\Big(\int_0^t B_u\,du \int_0^s B_v\,dv\Big) = \int_0^t du \int_0^s \mathrm{E}[B_u B_v]\,dv\\
&= \int_0^t du \int_0^s u \wedge v\,dv = \int_s^t du \int_0^s v\,dv + \int_0^s du \int_0^s u \wedge v\,dv\\
&= \frac{s^2}{2}\,(t-s) + \int_0^s du \int_0^u v\,dv + \int_0^s du \int_u^s u\,dv\\
&= \frac{s^2}{2}\,(t-s) + \frac{s^3}{3}\,\cdot
\end{aligned}$$

In order for $(X_t)_t$ to be Markovian, using the criterion of Exercise 6.1 b), the following relation

$$K_{t,u} = K_{t,s} K_{s,s}^{-1} K_{s,u} \tag{S.31}$$

must hold for every $u \le s \le t$. But

$$K_{t,s} K_{s,s}^{-1} K_{s,u} = \Big(\frac{s^2}{2}\,(t-s) + \frac{s^3}{3}\Big)\,\frac{3}{s^3}\,\Big(\frac{u^2}{2}\,(s-u) + \frac{u^3}{3}\Big)\,.$$

If we choose $u = 1$, $s = 2$, $t = 3$, this quantity is equal to $\frac{7}{4}\,\frac{5}{6} = \frac{35}{24}$, whereas $K_{t,u} = \frac{u^2}{2}\,(t-u) + \frac{u^3}{3} = \frac{1}{2} + \frac{2}{3} = \frac{5}{6}$. $(X_t)_t$ is not a Markov process.

c) If $s \le t$,

$$\mathrm{Cov}(X_t, B_s) = \mathrm{E}\Big(\int_0^t B_u B_s\,du\Big) = \int_0^t u\wedge s\,du = \int_0^s u\,du + \int_s^t s\,du = \frac{s^2}{2} + s(t-s)\,.$$

Conversely, if $s \ge t$,

$$\mathrm{Cov}(X_t, B_s) = \mathrm{E}\Big(\int_0^t B_u B_s\,du\Big) = \int_0^t u \wedge s\,du = \int_0^t u\,du = \frac{t^2}{2}\,.$$

d) If $Y_t = (B_t, X_t)$, the covariance function of this process is, for $s \le t$,

$$K_{t,s} = \begin{pmatrix} \mathrm{Cov}(B_t, B_s) & \mathrm{Cov}(B_t, X_s)\\ \mathrm{Cov}(X_t, B_s) & \mathrm{Cov}(X_t, X_s)\end{pmatrix} = \begin{pmatrix} s & \frac{s^2}{2}\\ \frac{s^2}{2} + s(t-s) & \frac{s^3}{3} + \frac{s^2}{2}\,(t-s)\end{pmatrix}\,.$$

Therefore the right-hand side of (S.31) becomes here

$$\begin{pmatrix} s & \frac{s^2}{2}\\ \frac{s^2}{2} + s(t-s) & \frac{s^3}{3} + \frac{s^2}{2}\,(t-s)\end{pmatrix} \begin{pmatrix} s & \frac{s^2}{2}\\ \frac{s^2}{2} & \frac{s^3}{3}\end{pmatrix}^{-1} \begin{pmatrix} u & \frac{u^2}{2}\\ \frac{u^2}{2} + u(s-u) & \frac{u^3}{3} + \frac{u^2}{2}\,(s-u)\end{pmatrix}\,.$$

Since

$$K_{s,s}^{-1} = \begin{pmatrix} s & \frac{s^2}{2} \\ \frac{s^2}{2} & \frac{s^3}{3} \end{pmatrix}^{-1} = \frac{12}{s^4}\begin{pmatrix} \frac{s^3}{3} & -\frac{s^2}{2} \\ -\frac{s^2}{2} & s \end{pmatrix}$$

we have

$$K_{t,s}K_{s,s}^{-1} = \frac{12}{s^4}\begin{pmatrix} s & \frac{s^2}{2} \\ \frac{s^2}{2} + s(t-s) & \frac{s^3}{3} + \frac{s^2}{2}(t-s) \end{pmatrix}\begin{pmatrix} \frac{s^3}{3} & -\frac{s^2}{2} \\ -\frac{s^2}{2} & s \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ t-s & 1 \end{pmatrix}$$

so that

$$\begin{aligned} K_{t,s}K_{s,s}^{-1}K_{s,u} &= \begin{pmatrix} 1 & 0 \\ t-s & 1 \end{pmatrix}\begin{pmatrix} u & \frac{u^2}{2} \\ \frac{u^2}{2} + u(s-u) & \frac{u^3}{3} + \frac{u^2}{2}(s-u) \end{pmatrix} \\ &= \begin{pmatrix} u & \frac{u^2}{2} \\ \frac{u^2}{2} + u(t-u) & \frac{u^3}{3} + \frac{u^2}{2}(t-u) \end{pmatrix} = K_{t,u}\ . \end{aligned}$$

Condition (S.31) is therefore satisfied and $(B_t, X_t)_t$ is a Markov process.

- We have found an example of a process, $(X_t)_t$, which is not Markovian, but which is a function of a Markov process.

**6.3**

a1) Let us observe first that $X$ is adapted to the filtration $(\mathscr{F}_{g(t)})_t$. Moreover, $X$ is a Gaussian process: its finite-dimensional distributions turn out to be linear transformations of finite-dimensional distributions of a Brownian motion. Hence we expect its transition function to be Gaussian.

Using the Markov property enjoyed by the Brownian motion, for a measurable bounded function $f$ and denoting by $p$ the transition function of the Brownian motion, we have with the change of variable $z = h(t)y$

$$\begin{aligned} \mathrm{E}[f(X_t)|\mathscr{F}_{g(s)}] &= \mathrm{E}[f(h(t)B_{g(t)})|\mathscr{F}_{g(s)}] = \int f(h(t)y)\,p(g(s), g(t), B_{g(s)}, dy) \\ &= \frac{1}{\sqrt{2\pi(g(t)-g(s))}}\int_{-\infty}^{+\infty} f(h(t)y)\exp\Big(-\frac{(y-B_{g(s)})^2}{2(g(t)-g(s))}\Big)\,dy \\ &= \frac{1}{h(t)\sqrt{2\pi(g(t)-g(s))}}\int_{-\infty}^{+\infty} f(z)\exp\Big(-\frac{(\frac{z}{h(t)} - \frac{X_{g(s)}}{h(s)})^2}{2(g(t)-g(s))}\Big)\,dz \\ &= \frac{1}{h(t)\sqrt{2\pi(g(t)-g(s))}}\int_{-\infty}^{+\infty} f(z)\exp\Big(-\frac{(z - \frac{h(t)}{h(s)}X_{g(s)})^2}{2h(t)^2(g(t)-g(s))}\Big)\,dz \end{aligned}$$

from which we deduce that $X$ is a Markov process with respect to the filtration $(\mathscr{F}_{g(t)})_t$ and associated to the transition function

$$q(s,t,x,dy) \sim N\Big(\frac{h(t)}{h(s)}\,x, h(t)^2(g(t)-g(s))\Big)\,. \tag{S.32}$$

a2) From (S.32) we have in general $X_t \sim N(0, h(t)^2 g(t))$. Hence under the condition of a2) $X_t \sim N(0,t)$, as for the Brownian motion. However, for the transition function of $X$ we have

$$q(s,t,x,dy) \sim N\Big(\frac{\sqrt{tg(s)}}{\sqrt{sg(t)}}\,x, \frac{t}{g(t)}\,(g(t)-g(s))\Big)$$

whereas recall that the transition function of a Brownian motion is

$$p(s,t,x,sy) \sim N(x,t-s)$$

hence $X$ cannot be a Brownian motion, unless $g(t) = \sqrt{t}$ (and in this case $X_t = B_t$).

b) The clever reader has certainly sensed the apparition of the Iterated Logarithm Law. Actually, by a simple change of variable

$$\begin{aligned}\overline{\lim_{t\to+\infty}} \frac{X_t}{\sqrt{2g(t)h^2(t)\log\log g(t)}} &= \overline{\lim_{t\to+\infty}} \frac{h(t)B_{g(t)}}{\sqrt{2g(t)h^2(t)\log\log g(t)}}\\ &= \overline{\lim_{t\to+\infty}} \frac{B_{g(t)}}{\sqrt{2g(t)\log\log g(t)}} = 1\,.\end{aligned}$$

**6.4**

a) $X$ is clearly a Gaussian process and we know, by Exercise 6.1, that it is Markovian if it satisfies the relation

$$K_{t,u} = K_{t,s}K_{s,s}^{-1}K_{s,u}, \quad \text{for } u \le s \le t\,. \tag{S.33}$$

Now, if $s \le t$,

$$K_{t,s} = \mathrm{Cov}(X_s, X_t) = \mathrm{e}^{-\lambda(t+s)}\mathrm{E}[B_{\mathrm{e}^{2\lambda t}}B_{\mathrm{e}^{2\lambda s}}] = \mathrm{e}^{-\lambda(t+s)}\mathrm{e}^{2\lambda s} = \mathrm{e}^{-\lambda(t-s)}\,.$$

In particular, $K_{s,s} = \mathrm{Var}(X_s) \equiv 1$. Therefore

$$K_{t,s}K_{s,s}^{-1}K_{s,u} = \mathrm{e}^{-\lambda(t-s)}\mathrm{e}^{-\lambda(s-u)} = \mathrm{e}^{-\lambda(t-u)} = K_{t,u}$$

so that (S.33) is satisfied and $X$ is Markovian. In order to determine its transition function, just recall that $p(s,t,x,dy)$ is the conditional distribution of $X_t$ given

$X_s = x$. $X$ being a Gaussian process, we know that this is Gaussian with mean

$$\mathrm{E}(X_t) + \frac{K_{t,s}}{K_{s,s}}\,(x - \mathrm{E}(X_s)) = \mathrm{e}^{-\lambda(t-s)}x$$

and variance

$$K_{t,t} - \frac{K_{t,s}^2}{K_{s,s}} = 1 - \mathrm{e}^{-2\lambda(t-s)}\;.$$

Therefore $p(u, x, dy) \sim N(\mathrm{e}^{-\lambda u}x, 1 - \mathrm{e}^{-2\lambda u})$. As both mean and variance are functions of $t - s$ only, $X$ is time homogeneous.

b1) $X_t$ is Gaussian, centered and has variance $= 1$. In particular, the law of $X_t$ does not depend on $t$.

b2) This is immediate: the two random vectors are both Gaussian and centered. As we have seen in a1) that the covariance function of the process satisfies the relation $K_{t_1,t_2} = K_{t_1+h,t_2+h}$, they also have the same covariance matrix.

c) Under $\mathrm{P}^x$ the law of $Z_t$ is $p(t, x, dy) \sim N(\mathrm{e}^{-\lambda t}x, 1 - \mathrm{e}^{-2\lambda t})$. As $t \to +\infty$ the mean of this distribution converges to 0, whereas its variance converges to 1. Thanks to Exercise 1.14 a) this implies (6.33).

**6.5**

a) From Exercise 4.15 we know that $(X_t)_{t\le 1}$ is a centered Gaussian process and that, for $s \le t$,

$$K_{t,s} = \mathrm{E}(B_tB_s) + st\mathrm{E}(B_1^2) - s\mathrm{E}(B_tB_1) - t\mathrm{E}(B_sB_1) = s - st = s(1-t)\;.$$

$X_t$ therefore has variance $\sigma_t^2 = t(1-t)$. Going back to Exercise 6.1 a), the conditional law of $X_t$ given $X_s = x$ is Gaussian with mean

$$\frac{K_{t,s}}{K_{s,s}}\,x = \frac{1-t}{1-s}\,x \tag{S.34}$$

and variance

$$K_{t,t} - \frac{K_{t,s}^2}{K_{s,s}} = t(1-t) - \frac{s(1-t)^2}{1-s} = \frac{1-t}{1-s}\,(t-s)\;. \tag{S.35}$$

b) With the notations of Exercise 6.1 if $u \le s \le t$, $K_{t,u} = u(1-t)$, whereas

$$K_{t,s}K_{s,s}^{-1}K_{s,u} = s(1-t)\,\frac{1}{s(1-s)}\,u(1-s) = u(1-t) = K_{t,u}\;.$$

Therefore condition (6.32) is satisfied and $(X_t)_t$ is a Markov process. The transition function coincides with the conditional distribution of $X_t$ given $X_s = x$, i.e. $p(s, t, x, \cdot) \sim N(\frac{1-t}{1-s}\, x, \frac{1-t}{1-s}\, (t-s))$.

**6.6**

a) As the joint distributions of $X$ are also joint distributions of $B$, $X$ is also a Gaussian process. As for the covariance function, for $s \le t \le 1$,

$$K_{t,s} = \mathrm{E}(X_t X_s) = \mathrm{E}(B_{1-t} B_{1-s}) = \min(1-t, 1-s) = 1-t\,.$$

b) As, for $u \le s \le t$, $K_{t,s} K_{s,s}^{-1} K_{s,u} = (1-t)(1-s)^{-1}(1-s) = 1-t = K_{t,u}$, the Markovianity condition (6.32) is satisfied, hence $(X_t)_t$ is Markovian with respect to its natural filtration. Its transition function $p(s, t, x, \cdot)$ is Gaussian with mean and variance given respectively by (4.21) and (4.22), i.e. with mean

$$K_{t,s} K_{s,s}^{-1} x = \frac{1-t}{1-s}\, x$$

and variance

$$K_{t,t} - K_{t,s} K_{s,s}^{-1} K_{s,t} = (1-t) - \frac{(1-t)^2}{1-s} = \frac{1-t}{1-s}\,(t-s)\,.$$

- The transition function above is the same as that of the Brownian bridge (see Exercise 6.5). The initial distribution is different, as here it is the law of $B_1$, i.e. $N(0, 1)$.

**6.7**

a) In order to prove the existence of a continuous version we use Kolmogorov's Theorem 2.1. Let us assume that the process starts at time $u$ with initial distribution $\mu$. Recalling (6.6) which gives the joint law of $(X_s, X_t)$, we have for $u \le s \le t$,

$$\begin{aligned}\mathrm{E}[|X_t - X_s|^\beta] &= \int_{\mathbb{R}^m} \mu(dz) \int_{\mathbb{R}^m} p(u, s, z, dx) \int_{\mathbb{R}^m} |x-y|^\beta\, p(s, t, x, dy)\\ &\le c\,|t-s|^{m+\varepsilon} \int_{\mathbb{R}^m} \mu(dz) \int_{\mathbb{R}^m} p(u, s, z, dx) = c|t-s|^{m+\varepsilon}\,.\end{aligned}$$

Then by Theorem 2.1, $X$ has a continuous version (actually with Hölder continuous paths of exponent $\gamma$ for every $\gamma < \frac{\varepsilon}{\beta}$). In order to prove that the generator is local, let us verify condition (6.26): as $|y-x|^\beta > R^\beta$ for $y \notin B_R(x)$,

$$\frac{1}{h}\,p(s,s+h,x,B_R(x)^c) = \frac{1}{h}\int_{B_R(x)^c} p(s,s+h,x,dy)$$
$$\le \frac{R^{-\beta}}{h}\int_{\mathbb{R}^m} |y-x|^\beta p(s,s+h,x,dy) \le cR^{-\beta}|h|^{m+\varepsilon-1}\;,$$

which tends to 0 as $h \to 0$.

**6.8**

a) If $f = 1_A$ we have, recalling that $1_{A-x}(y) = 1_A(x+y)$,

$$\int 1_A(y)\,p(t,x,dy) = p(t,x,A) = p(t,0,A-x) = \int 1_{A-x}(y)\,p(t,0,dy)$$
$$\int 1_A(x+y)\,p(t,0,dy)$$

so that (6.34) is satisfied if $f$ is an indicator function. It is then also true for linear combinations of indicator functions and by the usual approximation methods (Proposition 1.11, for example) for every bounded Borel function $f$.

b) Recall that $p(h,x,\cdot)$ is the law of $X_h$ under $\mathrm{P}^x$, whereas $p(h,0,\cdot)$ is the law of $X_h$ under $\mathrm{P}^0$. Therefore, if we denote by $t_x f$ the "translated" function $y \mapsto f(x+y)$, thanks to a) we have

$$\mathrm{E}^x[f(X_h)] = \int f(y)\,p(h,x,dy) = \int f(x+y)\,p(h,0,dy) = \mathrm{E}^0[t_x f(X_h)]$$

for every bounded Borel function $f$. If $f \in C_K^2$ then

$$Lf(x) = \lim_{h\to 0+}\frac{1}{h}\big(T_h f(x) - f(x)\big) = \lim_{h\to 0+}\frac{1}{h}\big(\mathrm{E}^x[f(X_h)] - f(x)\big)$$
$$= \lim_{h\to 0+}\frac{1}{h}\big(\mathrm{E}^0[t_x f(X_h)] - t_x f(0)\big) = L(t_x f)(0)\;.$$

As we have

$$\frac{\partial f}{\partial y_i}(x) = \frac{\partial (t_x f)}{\partial y_i}(0), \qquad \frac{\partial^2 f}{\partial y_i^2}(x) = \frac{\partial^2 (t_x f)}{\partial y_i^2}(0)\;,$$

equating the two expressions

$$Lf(x) = \sum_{i,j=1}^{m} a_{ij}(x) \frac{\partial^2 f}{\partial x_i^2}(x) + \sum_{i=1}^{n} b_i(x) \frac{\partial f}{\partial x_i}(x)$$

$$L(t_x f)(0) = \sum_{i,j=1}^{m} a_{ij}(0) \frac{\partial^2 f}{\partial x_i^2}(x) + \sum_{i=1}^{n} b_i(0) \frac{\partial f}{\partial x_i}(x) ,$$

which must be equal for every choice of $f \in C_K^2$, we have $a_{ij}(x) = a_{ij}(0), b_i(x) = b_i(0)$, for every $i, j \le m$.

**6.9**

a) $p^h$ obviously satisfies condition i) on p. 151. Moreover, $p^h(t, x, \cdot)$ is a measure on $E$. As

$$p^h(t, x, E) = \frac{e^{-\alpha t}}{h(x)} \int_E h(y)\, p(t, x, dy) = \frac{e^{-\alpha t}}{h(x)} T_t h(x) = 1 ,$$

$p^h$ is also a probability. We still have to check the Chapman–Kolmogorov equation. Thanks to the Chapman–Kolmogorov equation for $p$,

$$\int_E p^h(s, y, A)\, p^h(t, x, dy) = \frac{e^{-\alpha t}}{h(x)} \int_E h(y) p(t, x, dy)\, \frac{e^{-\alpha s}}{h(y)} \int_A h(z)\, p(s, y, dz)$$
$$= \frac{e^{-\alpha(t+s)}}{h(x)} \int_A h(z)\, p(t + s, x, dz) = p^h(s + t, x, A) .$$

b) We have

$$T_t^h g(x) = \frac{e^{-\alpha t}}{h(x)} \int h(y) g(y) p(h, x, dy) = \frac{e^{-\alpha t}}{h(x)} T_t(hg)(x) .$$

Therefore, if $gh = f \in \mathscr{D}(L)$,

$$\begin{aligned} L^h g(x) &= \lim_{t\to 0+} \frac{1}{t}[T_t^h g(x) - g(x)] = \frac{1}{h(x)} \lim_{t\to 0+} \frac{1}{t}[e^{-\alpha t} T_t f(x) - f(x)] \\ &= \frac{1}{h(x)} \Big( \lim_{t\to 0+} \frac{1}{t}(T_t f(x) - f(x)) + \lim_{t\to 0+} \frac{1}{t}(e^{-\alpha t} - 1) T_t f(x) \Big) \\ &= \frac{1}{h(x)} (Lf(x) - \alpha f(x)) = \frac{1}{h(x)} L(gh)(x) - \alpha g(x) . \end{aligned} \tag{S.36}$$

c) If $h \in C^2$ and $g \in C^2_K$, then also $gh \in C^2_K$ and

$$L(gh) = h\,Lg + g\,Lh + \sum_{i,j=1}^m a_{ij} \frac{\partial g}{\partial x_i} \frac{\partial h}{\partial x_j} \cdot$$

From (6.35) we have $T_t h = e^{\alpha t} h$, hence $h \in \mathscr{D}(L)$ and $Lh = \alpha h$ so that, thanks to the relation between $L$ and $L^h$ given in (S.36),

$$L^h = \frac{1}{2} \sum_{i,j=1}^m a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^m \widetilde{b}_i \frac{\partial}{\partial x_i},$$

where

$$\widetilde{b}_i(x) = b(x) + \frac{1}{h(x)} \sum_{j=1}^m a_{ij}(x) \frac{\partial h}{\partial x_j} \cdot$$

d) Let us check that $h(x) = e^{\langle v,x \rangle}$ satisfies (6.35). We have

$$T_t h(x) = \int_{\mathbb{R}^m} h(y)\, p(t,x,dy) = \frac{1}{(2\pi t)^{m/2}} \int_{\mathbb{R}^m} e^{\langle v,y \rangle} e^{-\frac{1}{2t}|x-y|^2}\, dy\,.$$

We recognize in the rightmost integral the Laplace transform of an $N(x,tI)$ law computed at $v$. Hence (see Exercise 1.6)

$$T_t h(x) = e^{\langle v,x \rangle} e^{\frac{1}{2}t|v|^2}\,.$$

(6.35) is therefore verified with $\alpha = \frac{1}{2}|v|^2$. Thanks to b), therefore,

$$L^h = \frac{1}{2} \sum_{i=1}^m \frac{\partial^2}{\partial x_i^2} + \sum_{i=1}^m v_i \frac{\partial}{\partial x_i} \cdot$$

**6.10**

a) If $f$ is a bounded Borel function and $\mu$ is stationary, we have

$$\int_E f(x)\,\mu(dx) = \int_E T_t f(x)\,\mu(dx) = \int_E \mu(dx) \int_E f(y)\, p(t,x,dy)$$

and now just observe that $\int_E \mu(dx) p(t,x,\cdot)$ is the law of $X_t$, when the initial distribution is $\mu$.

b) This is a consequence of Fubini's theorem: if $f$ is a bounded Borel function with compact support, then

$$\int_{\mathbb{R}^m} T_t f(x)\,dx = \frac{1}{(2\pi t)^{m/2}} \int_{\mathbb{R}^m} dx \int_{\mathbb{R}^m} \mathrm{e}^{-\frac{1}{2t}|x-y|^2} f(y)\,dy$$
$$= \int_{\mathbb{R}^m} f(y)\,dy \underbrace{\int_{\mathbb{R}^m} \frac{1}{(2\pi t)^{m/2}}\, \mathrm{e}^{-\frac{1}{2t}|x-y|^2}\,dx}_{=1} = \int_{\mathbb{R}^m} f(y)\,dy\,.$$

c) Let us assume that (6.38) holds for every $x$; if an invariant probability $\mu$ existed, then we would have, for every $t > 0$ and every bounded Borel set $A$,

$$\mu(A) = \int_E p(t,x,A)\,d\mu(x)\,.$$

As $0 \le p(t,x,A) \le 1$ and $\lim_{t\to+\infty} p(t,x,A) = 0$, we can apply Lebesgue's theorem and obtain

$$\mu(A) = \lim_{t\to+\infty} \int_E p(t,x,A)\,d\mu(x) = 0$$

so that $\mu(A) = 0$ for every bounded Borel set $A$ in contradiction with the hypothesis that $\mu$ is a probability.

If $p$ is the transition function of a Brownian motion then, for every Borel set $A \subset \mathbb{R}^m$ having finite Lebesgue measure,

$$p(t,x,A) = \frac{1}{(2\pi t)^{m/2}} \int_A \mathrm{e}^{-\frac{1}{2t}|x-y|^2}\,dy \le \frac{1}{(2\pi t)^{m/2}}\, mis(A) \underset{t\to+\infty}{\to} 0\,.$$

Therefore a Brownian motion cannot have an invariant distribution.

d) Let $f : E \to \mathbb{R}$ be a bounded continuous function. Then (6.39) implies

$$\lim_{s\to+\infty} T_s f(x) = \lim_{s\to+\infty} \int_E f(y)\,p(s,x,dy) = \int_E f(y)\,d\mu(y)\,. \tag{S.37}$$

By the Feller property $T_t f$ is also bounded and continuous and, for every $x \in E$,

$$\lim_{s\to+\infty} T_s f(x) = \lim_{s\to+\infty} T_{s+t} f(x) = \lim_{s\to+\infty} T_s(T_t f)(x) = \int_E T_t f(y)\,\mu(dy)\,. \tag{S.38}$$

(S.37) and (S.38) together imply that the stationarity condition (6.37) is satisfied for every bounded continuous function. It is also satisfied for every bounded measurable function $f$ thanks to the usual measure theoretic arguments as in Theorem 1.5, thus completing the proof that $\mu$ is stationary.

## 6.11

a) Let $f : G \to \mathbb{R}$ be a bounded measurable function. Then we have for $s \le t$, thanks to the Markov property for the process $X$,

$$\begin{aligned}\mathrm{E}[f(Y_t)|\mathscr{F}_s] &= \mathrm{E}[f \circ \Phi(X_t)|\mathscr{F}_s] = \int_E f \circ \Phi(z)\, p(s,t,X_s,dz)\\ &= \int_E f \circ \Phi(z)\, p(s,t,\Phi^{-1}(Y_s),dz) = \int_G f(y)\, q(s,t,Y_s,dy)\;,\end{aligned}$$

where we denote by $q(s,t,y,\cdot)$ the image law of $p(s,t,\Phi^{-1}(y),\cdot)$ through the transformation $\Phi$. This proves simultaneously that $q$ is a transition function (thanks to Remark 6.2) and that $Y$ satisfies the Markov property.

b1) This is simply the integration rule with respect to an image probability (Proposition 1.1).

b2) If $f : G \to \mathbb{R}$ is a bounded measurable function and $s \le t$ we have, thanks to the Markov property of $X$,

$$\mathrm{E}[f(Y_s)|\mathscr{F}_s] = \mathrm{E}[f\circ\Phi(X_s)|\mathscr{F}_s] = \int_E f\circ\Phi(z)\, p(s,t,X_s,dz) = \int_E f(y) q(s,t,Y_s,dy),$$

which proves simultaneously the Markov property for $Y$ and the fact that $q$ is a transition function (Remark 6.2 again).

c) We must show that the invariance property (6.40) is satisfied when $X$ is a Brownian motion and $\Phi(x) = |x-z|$. This is rather intuitive by the property of rotational invariance of the transition function of the Brownian motion. Let us give a rigorous form to this intuition. Using (6.42), which is immediate as $p(s,x,dy)$ is Gaussian with mean $x$ and variance $s$, we must show that, if $|x-z| = |y-z|$ then

$$\mathrm{P}(\sqrt{s}Z \in A - x) = \mathrm{P}(\sqrt{s}Z \in A - y)$$

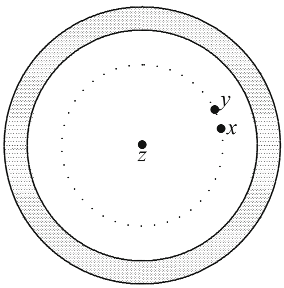


**Fig. S.5** Thanks to the property of rotational invariance, the probability of making a transition into the shaded area is the same starting from $x$ or $y$, or from whatever other point on the same sphere centered at $z$

for every $A$ of the form $A = \Phi^{-1}(\widetilde{A})$, $\widetilde{A} \in \mathscr{B}(\mathbb{R}^+)$. Such a subset of $\mathbb{R}^m$ is some kind of annulus around $z$ (see Fig. S.5). In any case the set $A-z$ is clearly rotationally invariant. Let $O$ be an orthogonal matrix such that $O(x-z) = y-z$ (the two vectors $x-z$ and $y-z$ have the same modulus). As the $N(0,I)$-distributed r.v.'s are rotationally invariant,

$$\begin{aligned}
&\mathrm{P}(\sqrt{s}Z \in A - x) = \mathrm{P}(\sqrt{s}Z \in A - z - (x-z)) \\
&= \mathrm{P}(\sqrt{s}Z \in O(A-z-(x-z))) = \mathrm{P}(\sqrt{s}Z \in A-z-(y-z)) = \mathrm{P}(\sqrt{s}Z \in A-y)\,.
\end{aligned}$$

**7.1**

a) We have $\mathrm{E}[Z] = 0$, as $Z$ is a stochastic integral with a bounded (hence belonging to $M^2$) integrand. We also have, by Fubini's theorem,

$$\mathrm{E}[Z^2] = \mathrm{E}\Big[\Big(\int_0^1 1_{\{B_t=0\}}\, dB_t\Big)^2\Big] = \mathrm{E}\Big[\int_0^1 1_{\{B_t=0\}}\, dt\Big] = \int_0^1 \mathrm{P}(B_t = 0)\, dt = 0\,.$$

$Z$, having both mean and variance equal to 0, is equal to 0 a.s.

b) $\mathrm{E}[Z] = 0$, as the integrand is in $M^2$. As for the variance we have

$$\mathrm{E}[Z^2] = \mathrm{E}\Big[\Big(\int_0^1 1_{\{B_t\ge 0\}}\, dB_t\Big)^2\Big] = \mathrm{E}\Big[\int_0^1 1_{\{B_t\ge 0\}}\, dt\Big] = \int_0^1 \mathrm{P}(B_t \ge 0)\, dt = \frac{1}{2}\,.$$

**7.2** Let $s \le t$. We have $B_s = \int_0^t 1_{[0,s[}(v)\, dB_v$ therefore, by Remark 7.1,

$$\mathrm{E}\Big(B_s \int_0^t B_u\, dB_u\Big) = \mathrm{E}\Big(\int_0^t 1_{[0,s[}(v)\, dB_v \int_0^t B_u\, dB_u\Big) = \int_0^t \mathrm{E}[1_{[0,s]}(u) B_u]\, du = 0\,.$$

If $t \le s$, the same argument leads to the same result.

**7.3** We have

$$\mathrm{E}(Y_t^2) = \int_0^t \mathrm{e}^{2s}\, ds = \frac{1}{2}\,(\mathrm{e}^{2t} - 1)$$

and

$$\mathrm{E}\Big[\int_0^t Y_s^2\, ds\Big] = \int_0^t \mathrm{E}[Y_s^2]\, ds = \int_0^t \frac{1}{2}\,(\mathrm{e}^{2s} - 1)\, ds = \frac{1}{4}\,(\mathrm{e}^{2t} - 1) - \frac{t}{2} < +\infty\,.$$

Hence $Y \in M^2$ so that $\mathrm{E}(Z_t) = 0$ and

$$\mathrm{E}[Z_t^2] = \mathrm{E}\Big[\int_0^t Y_s^2\, ds\Big] = \frac{1}{4}\,(\mathrm{e}^{2t} - 1) - \frac{t}{2}\,\cdot$$

Finally, using the argument of Example 7.1 c), if $s \le t$,

$$\mathrm{E}[Z_t Z_s] = \mathrm{E}\Big[\int_0^t Y_u\,dB_u \int_0^s Y_v\,dB_v\Big] = \mathrm{E}\Big[\Big(\int_0^s Y_u\,dB_u\Big)^2\Big] = \frac{1}{4}\,(\mathrm{e}^{2s}-1) - \frac{s}{2}\,\cdot$$

- In this and the following exercises we skip, when checking that a process is in $M^2$, the verification that it is progressively measurable. This fact is actually almost always obvious, thanks to Proposition 2.1 (an adapted right-continuous process is progressively measurable) or to the criterion of Exercise 2.3.

**7.4**

a) We have, by Theorems 7.1 and 7.6,

$$\mathrm{E}\Big[B_s^2\Big(\int_s^t B_u\,dB_u\Big)^2\Big] = \mathrm{E}\Big[\Big(\int_s^t B_s B_u\,dB_u\Big)^2\Big] = \mathrm{E}\Big[\int_s^t B_s^2 B_u^2\,du\Big]\,.$$

Recalling that $s \le u$ so that $B_u - B_s$ is independent of $B_s$,

$$\mathrm{E}(B_s^2 B_u^2) = \mathrm{E}[B_s^2(B_u - B_s + B_s)^2] = \underbrace{\mathrm{E}[B_s^2(B_u-B_s)^2]}_{=s(u-s)} + 2\,\underbrace{\mathrm{E}[B_s^3(B_u-B_s)]}_{=0} + \mathrm{E}(B_s^4)\,.$$

We can write $B_s = \sqrt{s}Z$ where $Z \sim N(0,1)$ and therefore $\mathrm{E}(B_s^4) = s^2\mathrm{E}(Z^4) = 3s^2$. In conclusion, $\mathrm{E}(B_s^2B_u^2) = s(u-s) + 3s^2$ and

$$\mathrm{E}\Big[B_s^2\Big(\int_s^t B_u\,dB_u\Big)^2\Big] = \int_s^t s(u-s) + 3s^2\,du = \frac{1}{2}s(t-s)^2 + 3s^2(t-s)\,.$$

b) One can write $Z = \int_0^t \widetilde{X}_u\,dB_u$, where $\widetilde{X}_u = X_u$ if $u \ge s$ and $\widetilde{X}_u = 0$ if $0 \le u < s$. It is immediate that $\widetilde{X} \in M^2([0,t])$. As $B_v = \int_0^t 1_{[0,v]}(u)\,dB_u$ and the product $\widetilde{X}_u 1_{[0,v]}(u)$ vanishes for every $0 \le u \le t$,

$$\mathrm{E}(ZB_v) = \mathrm{E}\Big(\int_0^t \widetilde{X}_u\,dB_u \int_0^t 1_{[0,v]}(u)\,dB_u\Big) = \mathrm{E}\Big(\int_0^t \widetilde{X}_u 1_{[0,v]}(u)\,du\Big) = 0 = \mathrm{E}(Z)\mathrm{E}(B_v)\,.$$

If $Z$ was independent of $\mathscr{F}_s$ it would be $\mathrm{E}[\phi(Z)W] = \mathrm{E}[\phi(Z)]\mathrm{E}[W]$ for every positive Borel function $\phi$ and every positive $\mathscr{F}_s$-measurable r.v. W. Let us prove that this is not true, in general, with a counterexample. Using the computation of a), let $\phi(Z) = Z^2$ with $Z = \int_s^t B_u\,dB_u$ and $W = B_s^2$. Then

$$\mathrm{E}(Z^2) = \int_s^t \mathrm{E}[B_u^2]\,du = \frac{1}{2}(t^2 - s^2)\,,$$

hence $\mathrm{E}[Z^2]\mathrm{E}[B_s^2] = \frac{s}{2}(t^2 - s^2)$, which is different from the value of $\mathrm{E}(B_s^2 Z^2)$ computed in a).

**7.5** We must check that $s \mapsto \mathrm{e}^{-B_s^2}$ is a process in $M^2$. For every $t \ge 0$ and recalling Remark 3.3,

$$\mathrm{E}\Big[\int_0^t \mathrm{e}^{-2B_s^2}\, ds\Big] = \int_0^t \mathrm{E}[\mathrm{e}^{-2B_s^2}]\, ds = \int_0^t \frac{1}{\sqrt{1+4s}}\, ds < +\infty\,.$$

In order to investigate whether $X$ is bounded in $L^2$, just observe that

$$\mathrm{E}[X_t^2] = \mathrm{E}\Big[\int_0^t \mathrm{e}^{-2B_s^2}\, ds\Big] = \int_0^t \frac{1}{\sqrt{1+4s}}\, ds \underset{t\to+\infty}{\to} +\infty$$

so that the answer is no.

**7.6**

a) Note that $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_2(t))_t, \mathrm{P})$ is a real Brownian motion and that $B_1$ is progressively measurable with respect to $(\mathscr{F}_t)_t$ and also $B_1 \in M^2_{B_2}([0,t])$ so that we can define

$$X_t = \int_0^t \mathrm{e}^{-B_1(u)^2}\, dB_2(u)\,.$$

Note also that $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_1(t))_t, \mathrm{P})$ is a real Brownian motion and that $X$ is progressively measurable with respect to $(\mathscr{F}_t)_t$. Moreover, for every $t \ge 0$,

$$\mathrm{E}[X_t^2] = \mathrm{E}\Big[\int_0^t \mathrm{e}^{-2B_1(u)^2}\, du\Big] = \int_0^t \frac{1}{\sqrt{1+4u}}\, du = \frac{1}{2}\big(\sqrt{1+4t}-1\big)\,. \tag{S.39}$$

As we can write

$$Z_t = \int_0^t \frac{X_s}{1+4s}\, dB_1(s) \tag{S.40}$$

and

$$\mathrm{E}\Big[\frac{X_s^2}{(1+4s)^2}\Big] = \frac{1}{2}\,\frac{\sqrt{1+4s}-1}{(1+4s)^2}\,,$$

the integrand $s \mapsto \frac{X_s}{1+4s}$ is in $M^2_{B_1}([0,t])$ for every $t > 0$ and $Z$ is a martingale.

Using the notation (S.40) and thanks to Theorem 7.3 and Remark 7.1

$$\langle Z\rangle_t = \int_0^t \frac{X_s^2}{(1+4s)^2}\, ds, \quad \langle Z, B_1\rangle_t = \int_0^t \frac{X_s}{1+4s}\, ds\,.$$

b) It is sufficient (and also necessary...) to prove that $Z$ is bounded in $L^2$. Now, recalling (S.39),

$$\mathrm{E}[Z_t^2]=\int_0^t \frac{\mathrm{E}[X_s^2]}{(1+4s)^2}\,ds=\frac{1}{2}\int_0^t\Big(\frac{1}{(1+4s)^{3/2}}-\frac{1}{(1+4s)^2}\Big)\,ds\,.$$

The rightmost integral being convergent as $t\to+\infty$, the $L^2$ norm of $Z$ is bounded, so that $Z$ converges as $t\to+\infty$ a.s. and in $L^2$. As $L^2$ convergence implies the convergence of the expectations we have $\mathrm{E}[Z_\infty]=0$. Also $L^2$ convergence implies the convergence of the second-order moment, so that

$$\begin{aligned}\mathrm{E}[Z_\infty^2]&=\lim_{t\to+\infty}\mathrm{E}[Z_t^2]=\frac{1}{2}\int_0^{+\infty}\Big(\frac{1}{(1+4s)^{3/2}}-\frac{1}{(1+4s)^2}\Big)\,ds\\&=-\frac{1}{4}\,\frac{1}{(1+4s)^{1/2}}\Big|_0^{+\infty}+\frac{1}{8}\,\frac{1}{1+4s}\Big|_0^{+\infty}=\frac{1}{8}\,\cdot\end{aligned}$$

**7.7**

a) If $f\in L^2([s,t])$ is a piecewise constant function the statement is immediate. Indeed, if $f=\sum_{i=1}^n \lambda_i 1_{[t_{i-1},t_i[}$ with $s=t_1<\dots<t_n=t$, then

$$\int_s^t f(u)\,dB_u=\sum_{i=1}^n \lambda_i(B_{t_i}-B_{t_{i-1}})$$

and all the increments $B_{t_i}-B_{t_{i-1}}$ are independent of $\mathscr{F}_s$. In general, if $(f_n)_n$ is a sequence of piecewise constant functions (that are dense in $L^2([s,t])$) converging to $f$ in $L^2([s,t])$, then, by the isometry property of the stochastic integral,

$$B_{f_n}\stackrel{def}{=}\int_s^t f_n(u)\,dB_u\ \underset{n\to\infty}{\overset{L^2}{\to}}\ \int_s^t f(u)\,dB_u\stackrel{def}{=}B_f$$

and, possibly taking a subsequence, we can assume that the convergence also takes place a.s. We must prove that, for every bounded $\mathscr{F}_s$-measurable r.v. $W$ and for every bounded Borel function $\psi$,

$$\mathrm{E}[W\psi(B_f)]=\mathrm{E}[W]\mathrm{E}[\psi(B_f)]\,. \tag{S.41}$$

But, if $\psi$ is bounded continuous, then $\psi(B_{f_n})\to\psi(B_f)$ a.s. as $n\to\infty$ and we can take the limit of both sides of the relation

$$\mathrm{E}[W\psi(B_{f_n})]=\mathrm{E}[W]\mathrm{E}[\psi(B_{f_n})]\,.$$

(S.41) is therefore proved for $\psi$ bounded continuous. We attain the general case with Theorem 1.5.

Alternatively, in a simpler but essentially similar way, we could have used the criterion of Exercise 4.5: for every $\lambda \in \mathbb{R}$ we have

$$\mathrm{E}(\mathrm{e}^{i\lambda B_f} \,|\, \mathscr{F}_s) = \lim_{n\to\infty} \mathrm{E}(\mathrm{e}^{i\lambda B_{f_n}} \,|\, \mathscr{F}_s) = \lim_{n\to\infty} \mathrm{E}(\mathrm{e}^{i\lambda B_{f_n}}) = \mathrm{E}(\mathrm{e}^{i\lambda B_f}) \;,$$

where the limits are justified by Lebesgue's theorem.

- Note that the previous proof would be much simpler if we only had to prove that $B_f$ is independent of $\mathscr{G}_s = \sigma(B_v, v \le s)$. In this case, by the criterion of Remark 1.2, it is sufficient to check that $B_f$ and $B_v$ are not correlated for every $v \le s$, which is immediate: if we denote by $\widetilde{f}$ the extension of $f$ to $[0,t]$ obtained by setting $f(v) = 0$ on $[0,s[$, then

$$\mathrm{E}(B_v B_f) = \mathrm{E}\Big(\int_0^t 1_{[0,v]}\, dB_u \int_0^t \widetilde{f}(u)\, dB_u\Big) = \int_0^t 1_{[0,v]}(u)\widetilde{f}(u)\, du = 0 \;.$$

b) The r.v. $B^\Phi_t$ is $\mathscr{F}_{\Phi^{-1}(t)}$-measurable. This suggests to try to see whether $B^\Phi$ is a Brownian motion with respect to $(\widetilde{\mathscr{F}}_t)_t$, with $\widetilde{\mathscr{F}}_t = \mathscr{F}_{\Phi^{-1}(t)}$. If $s \le t$, then

$$B^\Phi_t - B^\Phi_s = \int_{\Phi^{-1}(s)}^{\Phi^{-1}(t)} \sqrt{\Phi'(u)}\, dB_u$$

is independent of $\widetilde{\mathscr{F}}_s$ thanks to the result of a). We still have to prove c) of Definition 3.1. But $B^\Phi_t - B^\Phi_s$ is Gaussian by Proposition 7.1, is centered and has variance

$$\mathrm{E}[(B^\Phi_t - B^\Phi_s)^2] = \int_{\Phi^{-1}(s)}^{\Phi^{-1}(t)} \Phi'(u)\, du = t - s \;.$$

**7.8** Let us assume first that $f = 1_{[0,s]}$; therefore $\int_0^s f(u)\, dB_u = B_s$. Note that the r.v.'s $B_s B_u^2$ and $-B_s B_u^2$ have the same law ($-B$ is also a Brownian motion) so that $\mathrm{E}[B_s B_u^2] = 0$ and we have

$$\mathrm{E}\Big(B_s \int_0^t B_u^2\, du\Big) = \int_0^t \mathrm{E}(B_s B_u^2)\, du = 0 \;.$$

In conclusion $\int_0^t B_u^2\, du$ is orthogonal to all the r.v.'s $B_s, s > 0$. It is therefore orthogonal to $\int_0^s f(u)\, dB_u$, when $f$ is a piecewise constant function of $L^2([0,s])$, as in this case the stochastic integral is a linear combination of the r.v.'s $B_u, 0 \le u \le s$. The conclusion now follows by density.

**7.9**

a) We have

$$\sum_{i=0}^{n-1} B_{t_{i+1}}(B_{t_{i+1}}-B_{t_i})=\sum_{i=0}^{n-1} B_{t_i}(B_{t_{i+1}}-B_{t_i})+\sum_{i=0}^{n-1}(B_{t_{i+1}}-B_{t_i})^2\,.$$

The first term on the right-hand side converges, in probability, to the integral

$$\int_0^t B_s\,dB_s$$

as the amplitude of the partition tends to 0 thanks to Proposition 7.4. Conversely, by Proposition 3.4,

$$\sum_{i=0}^{n-1}(B_{t_{i+1}}-B_{t_i})^2\ \mathop{\to}_{|\pi|\to 0}\ t$$

in $L^2$, hence in probability. Therefore, in probability,

$$\sum_{i=0}^{n-1} B_{t_{i+1}}(B_{t_{i+1}}-B_{t_i})\ \mathop{\to}_{|\pi|\to 0}\ t+\int_0^t B_s\,dB_s\,.$$

b) We have again

$$\sum_{i=0}^{n-1} X_{t_{i+1}}(B_{t_{i+1}}-B_{t_i})=\sum_{i=0}^{n-1} X_{t_i}(B_{t_{i+1}}-B_{t_i})+\sum_{i=0}^{n-1}(X_{t_{i+1}}-X_{t_i})(B_{t_{i+1}}-B_{t_i})\,.$$

Now, if $X$ has paths with finite variation, we have a.s.

$$\begin{aligned}\Big|\sum_{i=0}^{n-1}(X_{t_{i+1}}-X_{t_i})(B_{t_{i+1}}-B_{t_i})\Big| &\le \sup_{i=0,\dots,n-1}|B_{t_{i+1}}-B_{t_i}|\sum_{i=0}^{n-1}|X_{t_{i+1}}-X_{t_i}|\\ &\le C\sup_{i=0,\dots,n-1}|B_{t_{i+1}}-B_{t_i}|\ \mathop{\to}_{|\pi|\to 0}\ 0\end{aligned}$$

thanks to the continuity of $B$, where $C=C(\omega)$ is the variation of $t\mapsto X_t$ in the interval $[0,t]$. Therefore the sums (7.38) and (7.39) have the same limit in probability.

**7.10**

a) We know already that $X_t = \int_0^t f(s)\,dB_s$ is a martingale, as $f \in M^2([0,T])$ for every $T$. As, for every $t \geq 0$,

$$\mathrm{E}[X_t^2] = \int_0^t f(s)^2\,ds \leq \|f\|_2^2\,,$$

$(X_t)_t$ is a martingale bounded in $L^2$ and converges a.s. and in $L^2$.

b1) There are various ways of showing that $(Y_t)_t$ is a martingale. For instance, as $t \mapsto \int_0^t g(s)\,dB_s$ is a Gaussian martingale with independent increments and its associated increasing process is $A_t = \int_0^t g(s)^2\,ds$, we have already seen this in Exercise 5.11 c2). A direct computation is also simple:

$$\begin{aligned}\mathrm{E}[Y_t|\mathscr{F}_s] = \mathrm{E}\big[\mathrm{e}^{\int_0^t g(u)\,dB_u - \frac{1}{2}\int_0^t g^2(u)\,du}\,\big|\,\mathscr{F}_s\big] &= \mathrm{e}^{-\frac{1}{2}\int_0^t g^2(u)\,du}\mathrm{E}\big[\mathrm{e}^{\int_0^t g(u)\,dB_u}\,\big|\,\mathscr{F}_s\big]\\ &= \mathrm{e}^{\int_0^s g(u)\,dB_u - \frac{1}{2}\int_0^t g^2(u)\,du}\mathrm{E}\big[\mathrm{e}^{\int_s^t g(u)\,dB_u}\,\big|\,\mathscr{F}_s\big]\\ &= \mathrm{e}^{\int_0^s g(u)\,dB_u - \frac{1}{2}\int_0^t g^2(u)\,du}\mathrm{E}\big[\mathrm{e}^{\int_s^t g(u)\,dB_u}\big]\\ &= \mathrm{e}^{\int_0^s g(u)\,dB_u - \frac{1}{2}\int_0^t g^2(u)\,du + \frac{1}{2}\int_s^t g^2(u)\,du} = Y_s\,.\end{aligned}$$

As the r.v. $\int_0^t g(s)\,dB_s$ is Gaussian, we have

$$\mathrm{E}(Y_t^2) = \mathrm{E}\big[\mathrm{e}^{2\int_0^t g(s)\,dB_s}\big]\,\mathrm{e}^{-\int_0^t g(s)^2\,ds} = \mathrm{e}^{2\int_0^t g(s)\,dB_s}\mathrm{e}^{-\int_0^t g(s)^2\,ds} = \mathrm{e}^{\int_0^t g(s)^2\,ds}\,. \tag{S.42}$$

b2) If $g \in L^2(\mathbb{R}^+)$, then by (S.42) $(Y_t)_t$ is a martingale bounded in $L^2$ and therefore uniformly integrable.

b3)

$$\mathrm{E}(Y_t^\alpha) = \mathrm{E}\big[\mathrm{e}^{\alpha\int_0^t g(s)\,dB_s}\big]\,\mathrm{e}^{-\frac{\alpha}{2}\int_0^t g(s)^2\,ds} = \mathrm{e}^{\frac{1}{2}\alpha(\alpha-1)\int_0^t g(s)^2\,ds} \underset{t\to+\infty}{\to} 0\,.$$

Therefore, defining $Y_\infty = \lim_{t\to+\infty} Y_t$ (the a.s. limit exists because $Y$ is a continuous positive martingale), by Fatou's lemma

$$\mathrm{E}(Y_\infty^\alpha) \leq \varliminf_{t\to+\infty} \mathrm{E}(Y_t^\alpha) = 0\,.$$

As $Y_\infty^\alpha \geq 0$, this implies $Y_\infty^\alpha = 0$ and therefore $Y_\infty = 0$. If $(Y_t)_t$ was a uniformly integrable martingale, then it would also converge in $L^1$ and this would imply $\mathrm{E}(Y_\infty) = 1$. Therefore $(Y_t)_t$ is not uniformly integrable

**7.11**

a1) Gaussianity is a consequence of Proposition 7.1. Moreover, for $s \le t$,

$$\mathrm{E}(Y_sY_t) = (1-t)(1-s)\mathrm{E}\Big[\int_0^s \frac{1}{1-u}\,dB_u \int_0^t \frac{1}{1-v}\,dB_v\Big]$$
$$= (1-t)(1-s)\int_0^s \frac{1}{(1-u)^2}\,du = (1-t)(1-s)\Big(1-\frac{1}{1-s}\Big) = s(1-t)\ ,$$

i.e. the same covariance function as a Brownian bridge.

a2) From a1) for $s=t$, $\mathrm{E}(Y_t^2) = t(1-t) \to 0$ for $t \to 1-$ hence $Y_t \to 0$ in $L^2$ as $t \to 1-$.

b) $(W_s)_s$ is obviously a Gaussian continuous process. Moreover, observe that $t \mapsto A(t)$ is increasing, therefore, for $s \le t$,

$$\mathrm{E}(W_sW_t) = \mathrm{E}\Big[\int_0^{A(s)} \frac{dB_u}{1-u} \int_0^{A(t)} \frac{dB_v}{1-v}\Big] = \int_0^{A(s)} \frac{1}{(1-u)^2}\,du$$
$$= \frac{1}{1-A(s)} - 1 = 1 + s - 1 = s = s \wedge t\ .$$

Therefore $(W_s)_s$ is a (natural) Brownian motion. The inverse of $s \mapsto A(s)$ is $t \mapsto \tau(t) = \frac{t}{1-t}$, i.e. $A(\tau(t)) = t$. We have therefore

$$Y_t = (1-t)\int_0^t \frac{dB_u}{1-u} = (1-t)W_{\tau(t)}$$

and

$$\lim_{t\to 1-} Y_t = \lim_{t\to 1-} (1-t)W_{\tau(t)} = \lim_{t\to 1-} (1-t)\cdot\sqrt{2\tau(t)\log\log\tau(t)}\cdot\frac{W_{\tau(t)}}{\sqrt{2\tau(t)\log\log\tau(t)}}\ \cdot$$

In order to conclude the proof, we simply recall that $\tau(t) \to +\infty$ as $t \to 1-$ and therefore, by the Iterated Logarithm Law,

$$\overline{\lim_{t\to 1-}}\ \frac{W_{\tau(t)}}{\sqrt{2\tau(t)\log\log\tau(t)}} = 1, \qquad \underline{\lim_{t\to 1-}}\ \frac{W_{\tau(t)}}{\sqrt{2\tau(t)\log\log\tau(t)}} = -1 \quad \text{a.s.}$$

whereas

$$\lim_{t\to 1-} (1-t)\sqrt{2\tau(t)\log\log\tau(t)} = \lim_{t\to 1-} \sqrt{2t(1-t)\log\log\tau(t)} = 0$$

hence $\lim_{t\to 1-} Y_t = 0$ a.s.

**7.12**

a) $Y_t$ and $Z_t$ are both Gaussian r.v.'s, as they are stochastic integrals with a deterministic square integrable integrand (Proposition 7.1). Both have mean equal to 0. As for the variance,

$$\mathrm{Var}(Y_t)=\mathrm{E}\Big[\Big(\int_0^t \mathrm{e}^{-\lambda(t-s)}\,dB_s\Big)^2\Big]=\int_0^t \mathrm{e}^{-2\lambda(t-s)}\,ds=\int_0^t \mathrm{e}^{-2\lambda u}\,du=\frac{1}{2\lambda}(1-\mathrm{e}^{-2\lambda t})$$

and in the same way

$$\mathrm{Var}(Z_t)=\mathrm{E}\Big[\Big(\int_0^t \mathrm{e}^{-\lambda s}\,dB_s\Big)^2\Big]=\int_0^t \mathrm{e}^{-2\lambda s}\,ds=\frac{1}{2\lambda}(1-\mathrm{e}^{-2\lambda t})\,.$$

Both $Y_t$ and $Z_t$ have a Gaussian law with mean 0 and variance $\frac{1}{2\lambda}(1-\mathrm{e}^{-2\lambda t})$.

b) $(Z_t)_t$ is a martingale (Proposition 7.3). For $(Y_t)_t$, if $s\le t$, we have instead

$$\mathrm{E}[Y_t\,|\,\mathscr{F}_s]=\mathrm{e}^{-\lambda t}\mathrm{E}\Big[\int_0^t \mathrm{e}^{\lambda u}\,dB_u\,\Big|\,\mathscr{F}_s\Big]=\mathrm{e}^{-\lambda t}\int_0^s \mathrm{e}^{\lambda u}\,dB_u=\mathrm{e}^{-\lambda(t-s)}Y_s\,.$$

Therefore $(Y_t)_t$ is not a martingale.

c) We have

$$\sup_{t\ge 0}\mathrm{E}[Z_t^2]=\lim_{t\ge 0}\mathrm{E}[Z_t^2]=\frac{1}{2\lambda}<+\infty\,.$$

$(Z_t)_t$ is therefore a martingale bounded in $L^2$ and converges a.s. and in $L^2$.

d1) Obviously $(Y_t)_t$ converges in law as $t\to+\infty$, as $Y_t$ and $Z_t$ have the same law for every $t$ and $(Z_t)_t$ converges a.s. and therefore in law.

d2) We have

$$\mathrm{E}[(Y_{t+h}-Y_t)^2]=\mathrm{E}[Y_{t+h}^2]+\mathrm{E}[Y_t^2]-2\mathrm{E}[Y_tY_{t+h}]\,.$$

The two first terms on the right-hand side have already been computed in a). As for the last, conversely,

$$\begin{aligned}\mathrm{E}[Y_tY_{t+h}]&=\mathrm{e}^{-\lambda(t+h)}\mathrm{e}^{-\lambda t}\mathrm{E}\Big(\int_0^t \mathrm{e}^{\lambda s}\,dB_s\int_0^{t+h}\mathrm{e}^{\lambda s}\,dB_s\Big)=\mathrm{e}^{-\lambda(2t+h)}\int_0^t \mathrm{e}^{2\lambda s}\,ds\\&=\frac{1}{2\lambda}(\mathrm{e}^{-\lambda h}-\mathrm{e}^{-\lambda(2t+h)})\,.\end{aligned}$$

Therefore, putting the pieces together,

$$\begin{aligned}\lim_{t\to+\infty}\mathrm{E}[(Y_{t+h}-Y_t)^2]&=\lim_{t\to+\infty}\frac{1}{2\lambda}\big(1-\mathrm{e}^{-2\lambda t}+1-\mathrm{e}^{-2\lambda(t+h)}-2\mathrm{e}^{-\lambda h}+2\mathrm{e}^{-\lambda(2t+h)}\big)\\&=\frac{1}{\lambda}(1-\mathrm{e}^{-\lambda h})\,.\end{aligned}$$

$(Y_t)_t$ is not therefore Cauchy in $L^2$ as $t \to +\infty$, and cannot converge in $L^2$. We shall see in Sect. 9.2 that $(Y_t)_t$ is an *Ornstein–Uhlenbeck process*.

**7.13**

a) $\widetilde{B}$ is a Gaussian process by Proposition 7.1. In order to take advantage of Proposition 3.1 let us show that $\mathrm{E}(\widetilde{B}_s\widetilde{B}_t) = s \wedge t$. Let us assume $s \le t$. As

$$\begin{aligned}\mathrm{E}[\widetilde{B}_t\widetilde{B}_s] &= \mathrm{E}\Big[\int_0^t \Big(3 - \frac{12u}{t} + \frac{10u^2}{t^2}\Big)\, dB_u \int_0^s \Big(3 - \frac{12u}{s} + \frac{10u^2}{s^2}\Big)\, dB_u\Big]\\ &= \int_0^s \Big(3 - \frac{12u}{t} + \frac{10u^2}{t^2}\Big)\Big(3 - \frac{12u}{s} + \frac{10u^2}{s^2}\Big)\, du\end{aligned}$$

$$= \int_0^s \Big(9 - \frac{36u}{s} + \frac{30u^2}{s^2} - \frac{36u}{t} + \frac{144u^2}{st} - \frac{120u^3}{s^2t} + \frac{30u^2}{t^2} - \frac{120u^3}{t^2s} + \frac{100u^4}{s^2t^2}\Big)\, du$$

we have

$$\mathrm{E}[\widetilde{B}_t\widetilde{B}_s] = 9s - 18s + 10s - 18\,\frac{s^2}{t} + 48\,\frac{s^2}{t} - 30\,\frac{s^2}{t} + 10\,\frac{s^3}{t^2} - 30\,\frac{s^3}{t^2} + 20\,\frac{s^3}{t^2} = s\,.$$

b) As $Y$ and $\widetilde{B}_t$ are jointly Gaussian, we must just check that they are uncorrelated. Let $t \le 1$: we have

$$\begin{aligned}\mathrm{E}[Y\widetilde{B}_t] &= \mathrm{E}\Big[\int_0^1 u\, dB_u \int_0^t \Big(3 - \frac{12u}{t} + \frac{10u^2}{t^2}\Big)\, dB_u\Big]\\ &= \mathrm{E}\Big[\int_0^t u\, dB_u \int_0^t \Big(3 - \frac{12u}{t} + \frac{10u^2}{t^2}\Big)\, dB_u\Big]\\ &= \int_0^t u\Big(3 - \frac{12u}{t} + \frac{10u^2}{t^2}\Big)\, du = \frac{3}{2}\,t^2 - 4t^2 + \frac{10}{4}\,t^2 = 0\,.\end{aligned}$$

The proof for $t > 1$ is straightforward and is left to the reader.

c) Let $\widetilde{\mathscr{G}}_\infty = \sigma(\widetilde{B}_t, t \ge 0)$ and $\mathscr{G}_\infty = \sigma(B_t, t \ge 0)$. Obviously $\widetilde{\mathscr{G}}_\infty \subset \mathscr{G}_\infty$ as every r.v. $\widetilde{B}_t$ is $\mathscr{G}_\infty$-measurable. On the other hand the r.v. $Y$ defined in b) is $\mathscr{G}_\infty$-measurable but not $\widetilde{\mathscr{G}}_\infty$-measurable. If it were we would have

$$\mathrm{E}[Y\,|\,\widetilde{\mathscr{G}}_\infty] = Y$$

whereas, conversely,

$$\mathrm{E}[Y|\widetilde{\mathscr{G}}_\infty] = \mathrm{E}(Y) = 0$$

as $Y$ is independent of $\widetilde{B}_t$ for every $t$ and therefore also of $\widetilde{\mathscr{G}}_\infty$ (Remark 1.2).

**7.14**

a) We have

$$\int_0^T 1_{\{s<\tau_n\}} X_s^2\,ds = \int_0^{\tau_n} X_s^2\,ds \le n\,,$$

hence the process $(X_s 1_{\{s<\tau_n\}})_s$ belongs $M^2([0,T])$ and

$$\begin{aligned}\mathrm{E}\Big[\int_0^{\tau_n} X_s^2\,ds\Big] &= \mathrm{E}\Big[\int_0^T 1_{\{s<\tau_n\}} X_s^2\,ds\Big] = \mathrm{E}\Big[\Big(\int_0^T 1_{\{s<\tau_n\}} X_s\,dB_s\Big)^2\Big]\\ &= \mathrm{E}\Big[\Big(\int_0^{\tau_n} X_s\,dB_s\Big)^2\Big] = \mathrm{E}(M_{t\wedge\tau_n}^2) \le \mathrm{E}\Big[\sup_{0\le t\le T} M_t^2\Big]\,.\end{aligned}$$

b) Assume (7.40): as $\tau_n \nearrow T$ a.s., by Beppo Levi's theorem,

$$\mathrm{E}\Big[\sup_{0\le t\le T} M_t^2\Big] \ge \lim_{n\to\infty} \mathrm{E}\Big[\int_0^{\tau_n} X_s^2\,ds\Big] = \mathrm{E}\Big[\int_0^T X_s^2\,ds\Big]\,,$$

hence $X \in M^2([0,T])$. Conversely, if $X \in M^2([0,T])$, $(M_t)_{0\le t\le T}$ is a square integrable martingale bounded in $L^2$ and (7.40) follows from Doob's inequality (the second one of the relations (7.23)).

**7.15** Let $(\tau_n)_n$ be a sequence reducing the local martingale $M$. Then, for $s \le t$,

$$\mathrm{E}[M_{t\wedge\tau_n}|\mathscr{F}_s] = M_{s\wedge\tau_n}\,. \tag{S.43}$$

As $\tau_n \nearrow +\infty$ as $n\to\infty$, we have $M_{t\wedge\tau_n} \to M_t$ and $M_{s\wedge\tau_n} \to M_s$ as $n\to\infty$. Moreover, as the sequence $(M_{t\wedge\tau_n})_n$ is assumed to be uniformly integrable, the convergence for $n\to\infty$ also takes place in $L^1$ and we can pass to the limit in (S.43) and obtain that $M_t$ is integrable for every $t\ge 0$ and also the martingale relation

$$\mathrm{E}[M_t|\mathscr{F}_s] = M_s\,.$$

**8.1**

a) The computation of the stochastic differential of $(M_t)_t$ is of course an application of Ito's formula, but this can be done in many ways. The first that comes to mind is to write $M_t = f(B_t, t)$ with

$$f(x,t) = (x+t)e^{-(x+\frac{1}{2}t)} .$$

But one can also write $M_t = X_t Y_t$, where

$$\begin{aligned}X_t &= B_t + t\\ Y_t &= e^{-(B_t+\frac{1}{2}t)} .\end{aligned}$$

By Ito's formula, considering that $Y_t = f(B_t + \frac{1}{2}t)$ with $f(x) = e^{-x}$, we have

$$dY_t = -e^{-(B_t+\frac{1}{2}t)}\Big(dB_t + \frac{1}{2}\,dt\Big) + \frac{1}{2}\,e^{-(B_t+\frac{1}{2}t)}\,dt = -Y_t\,dB_t ,$$

whereas obviously $dX_t = dB_t + dt$. Then, by Ito's formula for the product, as $d\langle X,Y\rangle_t = -Y_t\,dt$,

$$\begin{aligned}dM_t = d(X_tY_t) &= X_t\,dY_t + Y_t\,dX_t + d\langle X,Y\rangle_t\\ &= -(B_t+t)Y_t\,dB_t + Y_t\,dB_t + Y_t\,dt - Y_t\,dt\\ &= -(B_t+t+1)e^{-(B_t+\frac{1}{2}t)}\,dB_t .\end{aligned}$$

b) Thanks to a) $(M_t)_t$ is a local martingale. It is actually a martingale, as the process $t \mapsto -(B_t + t - 1)e^{-(B_t+\frac{1}{2}t)}$ is in $M^2$.

In order to show that it is a martingale we could also have argued without computing the stochastic differential with the usual method of factoring out the Brownian increments: if $s \le t$

$$\begin{aligned}\mathrm{E}[M_t\,|\,\mathscr{F}_s] &= \mathrm{E}\big[(B_s + (B_t - B_s) + t)e^{-(B_s+(B_t-B_s)+\frac{1}{2}t)}\,\big|\,\mathscr{F}_s\big]\\ &= \mathrm{E}\big[(B_s+t)e^{-(B_s+\frac{1}{2}t)}e^{-(B_t-B_s)} + (B_t-B_s)e^{-(B_s+\frac{1}{2}t)}e^{-(B_t-B_s)}\,\big|\,\mathscr{F}_s\big]\\ &= (B_s+t)e^{-(B_s+\frac{1}{2}t)}\mathrm{E}\big[e^{-(B_t-B_s)}\big] + e^{-(B_s+\frac{1}{2}t)}\mathrm{E}\big[(B_t-B_s)e^{-(B_t-B_s)}\big]\\ e^{-(B_s+\frac{1}{2}t)}\big((B_s+t)e^{\frac{1}{2}(t-s)} - (t-s)e^{\frac{1}{2}(t-s)}\big) &= (B_s+s)e^{-(B_s+\frac{1}{2}s)} = M_s .\end{aligned}$$

This old method appears to be a bit more complicated.

**8.2** It is convenient to write $X_t = ab + (b-a)B_t - B_t^2 + t$, so that, recalling that $dB_t^2 = 2B_t\,dB_t + dt$,

$$dX_t = (b-a)\,dB_t - 2B_t\,dB_t - dt + dt = (b-a-2B_t)\,dB_t ,$$

i.e.

$$X_t = ab + \int_0^t (b - a - 2B_t)\, dB_t \,.$$

The integrand obviously being in $M^2$, $X$ is a square integrable martingale.

**8.3**

a) Again Ito's formula can be applied in many ways. For instance, if $u(x,t) = e^{t/2}\sin x$,

$$dX_t = du(B_t,t) = \frac{\partial u}{\partial t}(B_t,t)\, dt + \frac{\partial u}{\partial x}(B_t,t)\, dB_t + \frac{1}{2}\,\frac{\partial^2 u}{\partial x^2}(B_t,t)\, dt$$

and as

$$\frac{\partial u}{\partial t}(x,t) = \frac{1}{2}\, e^{t/2}\sin x, \quad \frac{\partial u}{\partial x}(x,t) = e^{t/2}\cos x, \quad \frac{\partial^2 u}{\partial x^2}(x,t) = -e^{t/2}\sin x$$

we find

$$dX_t = \Big(\frac{1}{2}\, e^{t/2}\sin B_t - \frac{1}{2}\, e^{t/2}\sin B_t\Big)\, dt + e^{t/2}\cos B_t\, dB_t = e^{t/2}\cos B_t\, dB_t$$

and therefore $X$ is a local martingale and even a martingale, being bounded on bounded intervals. For $Y$ the same computation with $u(x,t) = e^{t/2}\cos x$ gives

$$dY_t = -e^{t/2}\sin B_t\, dB_t$$

so that $Y$ is also a martingale. As the expectation of a martingale is constant we have $E(X_t) = 0$, $E(Y_t) = 1$ for every $t \geq 0$.

More quickly the imaginative reader might have observed that

$$Y_t + iX_t = e^{iB_t + \frac{1}{2}t}$$

which is known to be an exponential complex martingale, hence both its real and imaginary parts are martingales.

b) The computation of the stochastic differentials of a) implies that both $X$ and $Y$ are Ito processes. It is immediate that

$$\langle X, Y\rangle_t = -\int_0^t e^s \cos B_s \sin B_s\, ds \,.$$

**8.4**

a) Just take the stochastic differential of the right-hand side: by Ito's formula

$$de^{\sigma B_t-\frac{1}{2}\sigma^2 t}=e^{\sigma B_t-\frac{1}{2}\sigma^2 t}(\sigma dB_t-\frac{1}{2}\sigma^2 dt)+\frac{1}{2}e^{\sigma B_t-\frac{1}{2}\sigma^2 t}\sigma^2\,dt=e^{\sigma B_t-\frac{1}{2}\sigma^2 t}\sigma\,dB_t\,.$$

Hence the left and right-hand sides in (8.52) have the same differential. As they both vanish at $t=0$ they coincide.

b) Recall that if $X\in M^2_{loc}$ and

$$Y_t=e^{\int_0^t X_u\,dB_u-\frac{1}{2}\int_0^t X_u^2\,du}$$

then $Y$ is a local martingale and satisfies the relation

$$dY_t=X_tY_t\,dB_t$$

and therefore

$$\int_0^t X_sY_s\,dB_s=Y_t-1\,.$$

In our case $X_t=B_t$, hence

$$\int_0^t e^{\int_0^s B_u\,dB_u-\frac{1}{2}\int_0^s B_u^2\,du}B_s\,dB_s=e^{\int_0^t B_u\,dB_u-\frac{1}{2}\int_0^t B_u^2\,du}-1\,.$$

**8.5**

a) If $f(x,t)=x^3-3tx$, by Ito's formula,

$$\begin{aligned}df(B_t,t)&=\frac{\partial f}{\partial x}(B_t,t)\,dB_t+\frac{\partial f}{\partial t}(B_t,t)\,dt+\frac{1}{2}\,\frac{\partial^2 f}{\partial x^2}(B_t,t)\,dt\\&=\frac{\partial f}{\partial x}(B_t,t)\,dB_t+\Big(\frac{1}{2}\frac{\partial^2 f}{\partial x^2}+\frac{\partial f}{\partial t}\Big)(B_t,t)\,dt\,.\end{aligned}$$

It is immediate that $\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x,t)+\frac{\partial f}{\partial t}(x,t)=3x-3x=0$; therefore, as $f(0,0)=0$,

$$f(B_t,t)=\int_0^t\frac{\partial f}{\partial x}(B_s,s)\,dB_s=\int_0^t(3B_s^2-3s)\,dB_s\,.$$

The integrand belongs to $M^2([0,+\infty[)$ (it is a polynomial in $t$ and $B_t$) and therefore $X_t=f(B_t,t)$ is a (square integrable) martingale.

b) The computation above shows that $X_t = P_n(B_t, t)$ is a martingale if and only if

$$\frac{1}{2}\frac{\partial^2 P_n}{\partial x^2} + \frac{\partial P_n}{\partial t} = 0 \,.$$

If $P_n$ is of the form (8.53), the coefficient of $x^{n-2m-2}t^m$ in the polynomial $\frac{1}{2}\frac{\partial^2 P_n}{\partial x^2} + \frac{\partial P_n}{\partial t}$ is equal to

$$\frac{1}{2}(n-2m)(n-2m-1)c_{n,m} + (m+1)c_{n,m+1} \,.$$

Requiring these quantities to vanish and setting $c_{n,0} = 1$, we can compute all the coefficients sequentially one after the other, thus determining $P_n$, up to a multiplicative constant. We find the polynomials

$$\begin{aligned} P_1(x,t) &= x \\ P_2(x,t) &= x^2 - t \\ P_3(x,t) &= x^3 - 3tx \\ P_4(x,t) &= x^4 - 6tx^2 + 3t^2 \,. \end{aligned}$$

The first two give rise to already known martingales, whereas the third one is the polynomial of a).

c) The stopping theorem, applied to the martingale $(P_3(B_t, t))_t$ and to the bounded stopping time $\tau \wedge n$, gives

$$\mathrm{E}(B^3_{\tau\wedge n}) = 3\,\mathrm{E}[(\tau \wedge n)B_{\tau\wedge n}] \,.$$

We can apply Lebesgue's theorem and take the limit as $n \to \infty$ (the r.v.'s $B_{\tau\wedge n}$ lie between $-a$ and $b$, whereas we already know that $\tau$ is integrable) so that $\mathrm{E}(B^3_\tau) = 3\,\mathrm{E}[\tau B_\tau]$. Therefore, going back to Exercise 5.31 where the law of $B_\tau$ was computed

$$\mathrm{E}(\tau B_\tau) = \frac{1}{3}\,\mathrm{E}(B^3_\tau) = \frac{1}{3}\,\frac{-a^3b + ab^3}{a+b} = ab(b-a) \,.$$

If $B_\tau$ and $\tau$ were independent we would have $\mathrm{E}(B_\tau \tau) = \mathrm{E}(B_\tau)\mathrm{E}(\tau) = 0$. Therefore $B_\tau$ and $\tau$ *are not* independent if $a \neq b$. If conversely $a = b$, $\tau$ and $B_\tau$ are uncorrelated and we know already that in this case they are independent (Exercise 3.18 b)).

**8.6**

a) Ito's formula gives

$$dM_t = \lambda \mathrm{e}^{\lambda t} B_t\,dt + \mathrm{e}^{\lambda t}\,dB_t - \lambda \mathrm{e}^{\lambda t} B_t\,dt = \mathrm{e}^{\lambda t}\,dB_t$$

so that $M$ is a square integrable martingale (also Gaussian...) with associated increasing process

$$\langle M\rangle_t = \int_0^t e^{2\lambda s}\, ds = \frac{1}{2\lambda}\,(e^{2\lambda t} - 1)\;.$$

Note that, as $M$ is Gaussian, this associated increasing process is deterministic as was to be expected from Exercise 5.11.

b) If $\lambda < 0$, as

$$E[M_t^2] = \langle M\rangle_t = \frac{1}{2\lambda}\,(e^{2\lambda t} - 1)\;,$$

$M$ is bounded in $L^2$ and therefore uniformly integrable and convergent a.s. and in $L^2$. As $M$ is Gaussian and Gaussianity is preserved with respect to convergence in law (Proposition 1.9), the limit is also Gaussian with mean 0 and variance

$$\lim_{t\to+\infty} \frac{1}{2\lambda}\,(e^{2\lambda t} - 1) = -\frac{1}{2\lambda}\,\cdot$$

c1) $Z$ is clearly an exponential martingale.

c2) If $\lambda < 0$, $\lim_{t\to+\infty} M_t$ exists in $L^2$ and is a Gaussian $N(0, -\frac{1}{2\lambda})$-distributed r.v. Hence $Z_\infty$ is the exponential of an $N(\frac{1}{4\lambda}, -\frac{1}{2\lambda})$ distributed r.v., i.e. a lognormal r.v. (Exercise 1.11).

c3) We have

$$E[Z_t^p] = E[e^{pM_t - \frac{p}{4\lambda}\,(e^{2\lambda t}-1)}] = e^{\frac{p^2-p}{4\lambda}\,(e^{2\lambda t}-1)}\;.$$

If $\lambda > 0$, for $p < 1$ we have $p^2 - p < 0$, hence

$$\lim_{t\to+\infty} E[Z_t^p] = 0\;.$$

Therefore, by Fatou's lemma,

$$E[Z_\infty^p] \le \varliminf_{t\to+\infty} E[Z_t^p] = 0\;.$$

The r.v. $Z_\infty^p$, being positive and with an expectation equal to 0, is equal to 0 a.s.

**8.7** By Ito's formula (see Example 8.6 for the complete computation)

$$dY_t = t\, dB_t$$

and therefore

$$d\Big(Y_t - \frac{1}{6}\,t^3\Big) = t\,dB_t - \frac{1}{2}\,t^2\,dt\,.$$

Hence, again by Ito's formula,

$$\begin{aligned}dZ_t &= \mathrm{e}^{Y_t-\frac{1}{6}t^3}\Big(t\,dB_t - \frac{1}{2}\,t^2\,dt\Big) + \frac{1}{2}\,\mathrm{e}^{Y_t-\frac{1}{6}t^3}\,d\langle Y\rangle_t\\ &= Z_t\Big(t\,dB_t - \frac{1}{2}\,t^2\,dt + \frac{1}{2}\,t^2\,dt\Big) = tZ_t\,dB_t\,.\end{aligned}$$

Therefore $Z$ is a local martingale (and a positive supermartingale). In order to prove that it is a martingale there are two possibilities. First we can prove that $Z \in M^2([0,T])$ for every $T \geq 0$: as $Y$ is a Gaussian process,

$$\mathrm{E}[Z_t^2] = \mathrm{E}[\mathrm{e}^{2Y_t-\frac{1}{3}t^3}] = \mathrm{e}^{-\frac{1}{3}t^3}\mathrm{e}^{2\mathrm{Var}(Y_t)} = \mathrm{e}^{-\frac{1}{3}t^3}\mathrm{e}^{\frac{2}{3}t^3} = \mathrm{e}^{\frac{1}{3}t^3}\,.$$

Therefore $Z \in M^2$ and $Z$, being the stochastic integral of a process in $M^2$, is a square integrable martingale. A second method is to prove that $\mathrm{E}[Z_t] = 1$ for every $t$ and to recall that a supermartingale with a constant expectation is a martingale (Exercise 5.1). This verification is left to the reader.

**8.8** By Corollary 8.1, there exists a Brownian motion $W$ such that, if

$$A_t = \int_0^t \frac{ds}{2+\sin s}\,,$$

then $X_t = W_{A_t}$. In order to investigate the behavior of $A_t$ as $t \to +\infty$ we must only pay attention to the fact that the primitive indicated in the hint is not defined for every value of $t$. Note that, as the integrand is a periodic function,

$$\int_0^{2\pi} \frac{ds}{2+\sin s} = \int_{-\pi}^{\pi} \frac{ds}{2+\sin s} = \frac{2}{\sqrt{3}}\arctan\Big(\frac{2}{\sqrt{3}}\tan\frac{t}{2} + \frac{1}{\sqrt{3}}\Big)\Big|_{-\pi}^{\pi} = \frac{2\pi}{\sqrt{3}}\,\cdot$$

Hence, denoting by [ ] the integer part function,

$$A_t = \int_0^t \frac{ds}{2+\sin s} = \frac{2\pi}{\sqrt{3}}\Big[\frac{t}{2\pi}\Big] + \int_{2\pi[\frac{t}{2\pi}]}^t \frac{ds}{2+\sin s}$$

so that $\lim_{t\to+\infty}\frac{1}{t}A_t = 3^{-1/2}$. We have therefore

$$\overline{\lim_{t\to+\infty}}\,\frac{X_t}{(2t\log\log t)^{1/2}} = \overline{\lim_{t\to+\infty}}\,\frac{W_{A_t}}{(2A_t\log\log A_t)^{1/2}}\,\frac{(2A_t\log\log A_t)^{1/2}}{(2t\log\log t)^{1/2}}\,\cdot$$

As

$$\lim_{t\to+\infty} \frac{(2A_t \log\log A_t)^{1/2}}{(2t\log\log t)^{1/2}} = 3^{-1/4}\,,$$

by the Iterated Logarithm Law we find

$$\overline{\lim_{t\to+\infty}}\, \frac{X_t}{(2t\log\log t)^{1/2}} = 3^{-1/4} \simeq 0.76\,.$$

**8.9** We have $d(\mathrm{e}^{-\lambda t}B_t) = -\lambda \mathrm{e}^{-\lambda t}B_t\,dt + \mathrm{e}^{-\lambda t}\,dB_t$ from which we derive

$$X_t = \int_0^t \mathrm{e}^{-\lambda s}B_s\,ds = -\frac{1}{\lambda}\,\mathrm{e}^{-\lambda t}B_t + \frac{1}{\lambda}\int_0^t \mathrm{e}^{-\lambda s}dB_s\,.$$

The stochastic integral on the right-hand side is a martingale bounded in $L^2$ and therefore it converges a.s. and in $L^2$. As $\lim_{t\to+\infty} \mathrm{e}^{-\lambda t}B_t = 0$ a.s. (recall the Iterated Logarithm Law) and in $L^2$, the limit in (8.54) also exists a.s. and in $L^2$, the limit being the r.v.

$$\frac{1}{\lambda}\int_0^{+\infty} \mathrm{e}^{-\lambda s}dB_s\,,$$

which is a centered Gaussian r.v. with variance

$$\frac{1}{\lambda^2}\int_0^{+\infty} \mathrm{e}^{-2\lambda s}\,ds = \frac{1}{2\lambda^3}\,\cdot$$

**8.10**

a) We know (Proposition 7.1) that $X_t^\varepsilon$ has a centered Gaussian distribution with variance

$$2\int_0^t \sin^2\frac{s}{\varepsilon}\,ds\,.$$

Integrating by parts we find

$$2\int_0^t \sin^2\frac{s}{\varepsilon}\,ds = -\,\varepsilon\sin\frac{t}{\varepsilon}\,\cos\frac{t}{\varepsilon} + t\,. \tag{S.44}$$

Hence, as $\varepsilon\to 0$, $X_t^\varepsilon$ converges in law to an $N(0,t)$-distributed r.v.

b) We know, by Theorem 8.2 and Corollary 8.1, that

$$X_t^\varepsilon = W^\varepsilon_{A_t^\varepsilon}$$

where $W^\varepsilon$ is a Brownian motion and

$$A_t^\varepsilon = 2\int_0^t \sin^2\frac{s}{\varepsilon}\,ds\,.$$

The Brownian motion $W^\varepsilon$ depends on $\varepsilon$ but of course it has the same law as $B$ for every $\varepsilon$. Hence $X^\varepsilon$ has the same distribution as $(B_{A_t^\varepsilon})_t$. As by (S.44) $A_t^\varepsilon \to t$ uniformly in $t$, thanks to the continuity of the paths we have

$$\sup_{0\le t\le T}\big|B_{A_t^\varepsilon}-B_t\big| \underset{\varepsilon\to 0}{\to} 0\,.$$

As a.s. convergence implies convergence in law, $(B_{A_t^\varepsilon})_t$, hence also $(X_t^\varepsilon)_t$, converges in law as $\varepsilon\to 0$ to the law of $B$, i.e. to the Wiener measure.

**8.11** By Corollary 8.1 there exists a Brownian motion $W$ such that, if

$$A_t = \int_0^t \frac{ds}{1+s} = \log(1+t)\,,$$

then $X_t = W_{A_t}$. Therefore, by the reflection principle applied to the Brownian motion $W$,

$$\mathrm{P}\Big(\sup_{0\le t\le 3} X_t \ge 1\Big) = \mathrm{P}\Big(\sup_{0\le t\le 3} W_{A_t}\ge 1\Big) = \mathrm{P}\Big(\sup_{0\le s\le \log 4} W_s \ge 1\Big)$$

$$= 2\mathrm{P}(W_{\log 4}\ge 1) = 2\mathrm{P}(\sqrt{\log 4}\,W_1\ge 1) = \frac{2}{\sqrt{2\pi}}\int_{(\log 4)^{-1/2}}^{+\infty} \mathrm{e}^{-x^2/2}\,dx \simeq 0.396\,.$$

**8.12**

a) We know that $X$ is a time changed Brownian motion, more precisely (Corollary 8.1)

$$X_t = W_{A_t}$$

where $W$ is a Brownian motion and

$$A_t = (\alpha+1)\int_0^t u^\alpha\,du = t^{\alpha+1}\,.$$

Therefore

$$\mathrm{P}\Big(\sup_{0\le s\le 2} X_s \ge 1\Big) = \mathrm{P}\Big(\sup_{0\le s\le 2} W_{A_s} \ge 1\Big) = \mathrm{P}\Big(\sup_{0\le t\le A_2} W_t \ge 1\Big) = 2\mathrm{P}(W_{A_2} \ge 1)$$
$$= 2\mathrm{P}(\sqrt{A_2}\, W_1 \ge 1) = 2\mathrm{P}(W_1 \ge 2^{-\frac{1}{2}(\alpha+1)})$$
$$= \frac{2}{\sqrt{2\pi}} \int_{2^{-\frac{1}{2}(\alpha+1)}}^{\infty} \mathrm{e}^{-t^2/2}\, dt\,.$$

b) The previous computation with $t$ instead of 2 gives

$$\mathrm{P}(\tau \le t) = \mathrm{P}\Big(\sup_{0\le s\le t} X_s \ge 1\Big) = \frac{2}{\sqrt{2\pi}} \int_{t^{-\frac{1}{2}(\alpha+1)}}^{\infty} \mathrm{e}^{-s^2/2}\, ds$$

and taking the derivative we find the density of $\tau$:

$$f_\tau(t) = \frac{2}{\sqrt{2\pi}}\, \frac{1}{2}\,(\alpha+1) t^{-\frac{1}{2}(\alpha+1)-1} \mathrm{e}^{-1/(2t^{\alpha+1})} = \frac{(\alpha+1)}{\sqrt{2\pi t^{\alpha+3}}}\, \mathrm{e}^{-1/(2t^{\alpha+1})}\,.$$

In order for $\tau$ to have finite mathematical expectation the function $t \mapsto tf_\tau(t)$ must be integrable. At zero the integrand tends to 0 fast enough because of the exponential whereas at infinity $tf_\tau(t) \sim t^{-\frac{1}{2}(\alpha+1)}$. It is therefore necessary that $-\frac{1}{2}(\alpha+1) < -1$, i.e. $\alpha > 1$.

**8.13** By Ito's formula $d(\phi_t B_t) = \phi_t' B_t\, dt + \phi_t\, dB_t$. Therefore, for $T > 0$,

$$\int_0^T \phi_t' B_t\, dt + \int_0^T \phi_t\, dB_t = \phi_T B_T - \underbrace{\phi_0 B_0}_{=0} \qquad \text{a.s.}$$

If $T$ is large enough, so that the support of $\phi$ is contained in $[0, T]$, we have $\phi_T = 0$ and therefore

$$\int_0^T \phi_t' B_t\, dt = -\int_0^T \phi_t\, dB_t \qquad \text{a.s.}$$

Moreover, $T$ can be replaced by $+\infty$, as both integrands vanish on $]T, +\infty[$.

**8.14**

a) Let us apply Ito's formula. If

$$\psi(x,t) = \frac{1}{\sqrt{1-t}} \exp\Big(-\frac{x}{2(1-t)}\Big)$$

then $Z_t = \psi(B_t^2, t)$ and we know that $dB_t^2 = 2B_t\, dB_t + dt$. Let us compute the derivatives of $\psi$:

$$\frac{\partial\psi}{\partial t}(x,t) = \frac{1}{2(1-t)^{3/2}} \exp\Big(-\frac{x}{2(1-t)}\Big) - \frac{x}{2(1-t)^{5/2}} \exp\Big(-\frac{x}{2(1-t)}\Big)$$
$$= \Big(\frac{1}{2(1-t)} - \frac{x}{2(1-t)^2}\Big)\psi(x,t)$$

and similarly

$$\frac{\partial\psi}{\partial x}(x,t) = -\frac{1}{2(1-t)}\,\psi(x,t)$$
$$\frac{\partial^2\psi}{\partial x^2}(x,t) = \frac{1}{4(1-t)^2}\,\psi(x,t)\;.$$

Hence

$$\begin{aligned} dZ_t &= \frac{\partial\psi}{\partial t}(B_t^2,t)\,dt + \frac{\partial\psi}{\partial x}(B_t^2,t)\,(2B_t\,dB_t + dt) + \frac{1}{2}\,\frac{\partial^2\psi}{\partial x^2}(B_t^2,t)\,4B_t^2\,dt \\ &= \Big[\Big(\frac{1}{2(1-t)} - \frac{B_t^2}{2(1-t)^2}\Big)\,dt - \frac{1}{2(1-t)}\big(2B_t\,dB_t + dt\big) \\ &\qquad + \frac{1}{8(1-t)^2}\,4B_t^2\,dt\Big]\psi(B_t^2,t) \\ &= -\frac{B_t}{1-t}\psi(B_t^2,t)\,dB_t = -\frac{1}{1-t}\,B_tZ_t\,dB_t\;. \end{aligned} \tag{S.45}$$

This proves that $Z$ is a local martingale. Being positive it is also a supermartingale. In order to prove that it is actually a martingale it does not seem immediate to verify that the stochastic integrand $t \mapsto -\frac{1}{1-t}\,B_tZ_t$ is in $M^2$, as it depends on $Z$ itself. It is however immediate that $Z$ is bounded ($\le (1-t)^{-1/2}$) in every interval $[0,T]$ with $T < 1$ and we know that a bounded local martingale is a martingale (Remark 7.7). It is also possible to check that, for every $t < 1$, $\mathrm{E}[Z_t] = \mathrm{E}[Z_0] = 1$, and we know that a supermartingale having constant expectation is a martingale (Exercise 5.1). This will be a byproduct of the computation of b).

b) We have

$$\mathrm{E}[Z_t^p] = \frac{1}{(1-t)^{p/2}}\mathrm{E}\Big[\exp\Big(-\frac{pB_t^2}{2(1-t)}\Big)\Big]\;.$$

Now, denoting by $W$ an $N(0,1)$-distributed r.v. (recall Remark 3.3)

$$\mathrm{E}\Big[\exp\Big(-\frac{pB_t^2}{2(1-t)}\Big)\Big] = \mathrm{E}\Big[\exp\Big(-\frac{ptW^2}{2(1-t)}\Big)\Big] = \frac{1}{\sqrt{1+\frac{pt}{1-t}}} = \frac{\sqrt{1-t}}{\sqrt{1+(p-1)t}}$$

so that

$$\mathrm{E}[Z_t^p] = \frac{1}{(1-t)^{(p-1)/2}\sqrt{1+(p-1)t}} \cdot$$

Therefore, for $p = 1$, $\mathrm{E}[Z_t] = 1$, thus a second proof that $Z$ is a martingale. Moreover, $Z$ belongs to $L^p$ but, if $p > 1$, it is not bounded in $L^p$: the $L^p$ norm diverges as $t \to 1-$.

c) $Z$ being a continuous positive martingale, the limit $\lim_{t\to 1-} Z_t := Z_1$ exists a.s. and is finite. From the computation of b) we have, for $p < 1$,

$$\lim_{t\to 1-} \mathrm{E}[Z_t^p] = 0$$

and, as $Z$ is positive, by Fatou's lemma

$$\mathrm{E}[Z_1^p] \le \underline{\lim}_{t\to 1-} \mathrm{E}[Z_t^p] = 0\,.$$

The r.v. $Z_1^p$ being positive and having expectation equal to 0 is necessarily equal to 0 a.s., i.e.

$$\lim_{t\to 1-} Z_t = 0 \qquad \text{a.s.}$$

d) From (S.45) we recognize the differential of an exponential martingale. Therefore a comparison between (S.45) and (8.11) suggests that

$$Z_t = \exp\Bigl(-\int_0^t \frac{B_s}{1-s}\,dB_s - \frac{1}{2}\int_0^t \frac{B_s^2}{(1-s)^2}\,ds\Bigr)\,,$$

i.e.

$$X_s = -\frac{B_s}{1-s} \cdot$$

This remark would be enough if we knew that (8.11) has a unique solution. Without this we can nevertheless verify directly that, for $t < 1$,

$$\frac{1}{\sqrt{1-t}}\exp\Bigl(-\frac{B_t^2}{2(1-t)}\Bigr) = \exp\Bigl(-\int_0^t \frac{B_s}{1-s}\,dB_s - \frac{1}{2}\int_0^t \frac{B_s^2}{(1-s)^2}\,ds\Bigr)\,. \tag{S.46}$$

Ito's formula for the product gives

$$d\frac{B_t^2}{2(1-t)} = \frac{2B_t\,dB_t + dt}{2(1-t)} + \frac{B_t^2}{2(1-t)^2}\,dt$$

hence integrating

$$\frac{B_t^2}{2(1-t)} = \int_0^t \frac{B_s}{1-s}\,dB_s + \int_0^t \frac{B_s^2}{2(1-s)^2}\,ds - \frac{1}{2}\log(1-t)$$

so that

$$\exp\Big(-\frac{B_t}{2(1-t)}\Big) = \sqrt{1-t} \times \exp\Big(-\int_0^t \frac{B_s}{1-s}\,dB_s - \frac{1}{2}\int_0^t \frac{B_s^2}{(1-s)^2}\,ds\Big)$$

and (S.46) follows.

**8.15**

a) It is immediate that $W_1$ is a centered Gaussian process. Its variance is equal to

$$\int_0^t (\sin s + \cos s)^2\,ds = \int_0^t (1 + 2\sin s \cos s)\,ds = t + (1 - \cos 2t) \neq t\,,$$

hence $W_1$ is not a Brownian motion.

b) We can write

$$W_2(t) = \int_0^t u(s)\,dB_s$$

with $u(s) = (\sin s, \cos s)$. As $u(s)$ is a vector having modulus equal to 1 for every $s$, $W_2$ is a Brownian motion thanks to Corollary 8.2.

c) As in b) we can write

$$W_3(t) = \int_0^t u(s)\,dB_s$$

where now $u(s) = (\sin B_2(s), \cos B_2(s))$. Again, as $u(s)$ has modulus equal to 1 for every $s$, $W_3$ is a Brownian motion. Note that, without taking advantage of Corollary 8.2, it is not immediate to prove that $W_3$ is a Gaussian process.

**8.16** Let us denote by $B$ the two-dimensional Brownian motion $(B_1, B_2)$.

a) We can write

$$X_t = \int_0^t Z_s\,dB_s$$

with $Z_s = (\sin B_3(s), \cos B_3(s))$. As $|Z_s| = 1$ for every $s$, $X$ is a Brownian motion by Corollary 8.2. The same argument applies for $Y$.

b1) We have, for $s \le t$,

$$\begin{aligned}
&\mathrm{E}[X_s Y_t]\\
&= \mathrm{E}\Big[\int_0^s \sin(B_3(u))\,dB_1(u) \int_0^t \cos(B_3(v))\,dB_1(v)\\
&\qquad + \int_0^s \cos(B_3(u))\,dB_2(u) \int_0^t \sin(B_3(v))\,dB_2(v)\Big]\\
&= \mathrm{E}\Big[\int_0^s \sin(B_3(u))\cos(B_3(u))\,du + \int_0^s \cos(B_3(u))\sin(B_3(u))\,du\Big]\\
&= 2\int_0^s \mathrm{E}[\sin(B_3(u))\cos(B_3(u))]\,du = 0\,,
\end{aligned}$$

where the last equality comes from the fact that $\sin(B_3(u))\cos(B_3(u))$ has the same distribution as $\sin(-B_3(u))\cos(-B_3(u)) = -\sin(B_3(u))\cos(B_3(u))$, and has therefore mathematical expectation equal to 0. Hence $X_s$ and $Y_t$ are uncorrelated.

b2) If $(X_t, Y_t)_t$ were a two-dimensional Brownian motion the product $t \mapsto X_t Y_t$ would be a martingale (see Exercises 5.22 or 5.24). This is not true because

$$\langle X, Y\rangle_t = 2\int_0^t \sin(B_3(u))\cos(B_3(u))\,du \not\equiv 0\,.$$

c) With the new definition we can write

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \int_0^t O_s\,dB_s$$

with

$$O_s = \begin{pmatrix} \sin(B_3(s)) & \cos(B_3(s)) \\ -\cos(B_3(s)) & \sin(B_3(s)) \end{pmatrix}\,.$$

It is immediate that $s \mapsto O_s$ is an orthogonal matrix-valued process. Hence the required statement follows from Proposition 8.8.

- Note the apparent contradiction: the processes $X$ and $Y$ in b) are each a Brownian motion and are uncorrelated. But they do not form a two-dimensional Brownian motion. Actually they are not jointly Gaussian.

**8.17**

a) Orthogonality with respect to $B_v$ for $v \le s$ imposes the condition

$$0 = \mathrm{E}\Big[\Big(B_t - \int_0^s \Phi(u)\,dB_u - \alpha B_1\Big)B_v\Big] = v - \int_0^v \Phi(u)\,du - \alpha v\,,$$

i.e.

$$v(1-\alpha)=\int_0^v \Phi(u)\,du, \quad \text{for every } v\le s \tag{S.47}$$

and therefore $\Phi\equiv 1-\alpha$ on $[0,s]$. Orthogonality with respect to $B_1$ conversely requires

$$0=\mathrm{E}\Big[\Big(B_t-\int_0^s \Phi(u)\,dB_u-\alpha B_1\Big)B_1\Big]=t-\int_0^s \Phi(u)\,du-\alpha\ ,$$

i.e., taking into account that $\Phi\equiv 1-\alpha$,

$$0=t-(1-\alpha)s-\alpha=t-s-\alpha(1-s)\ ,$$

i.e.

$$\alpha=\frac{t-s}{1-s},\qquad \Phi(u)\equiv\frac{1-t}{1-s}\,\cdot$$

b) Let $X=\frac{t-s}{1-s}\,B_1+\frac{1-t}{1-s}\,B_s$. In a) we have proved that the r.v. $B_t-X$, which is centered, is independent of $\widetilde{\mathscr G}_s$. Moreover, as $X$ is $\widetilde{\mathscr G}_s$-measurable,

$$\begin{aligned}\mathrm{E}[B_t\,|\widetilde{\mathscr G}_s]&=\mathrm{E}[(B_t-X)+X\,|\widetilde{\mathscr G}_s]\\ &=X+\mathrm{E}[B_t-X]=X=\frac{t-s}{1-s}\,B_1+\frac{1-t}{1-s}\,B_s\ .\end{aligned}\tag{S.48}$$

Hence $B$ is adapted to the filtration $(\widetilde{\mathscr G}_t)_t$ but is not a $(\widetilde{\mathscr G}_t)_t$-martingale, therefore it cannot be a Brownian motion with respect to this filtration.

c1) As the r.v.'s $\widetilde B_t-\widetilde B_s, B_1, B_v, v\le s$ form a Gaussian family, it suffices (Remark 1.2) to show that $\widetilde B_t-\widetilde B_s$ is orthogonal to $B_v$, $0\le v\le s$, and to $B_1$. We have

$$\mathrm{E}[(\widetilde B_t-\widetilde B_s)B_1]=\mathrm{E}[(B_t-B_s)B_1]-\int_s^t\frac{\mathrm{E}[(B_1-B_u)B_1]}{1-u}\,du=t-s-\int_s^t du=0$$

and, for $v\le s$,

$$\mathrm{E}[(\widetilde B_t-\widetilde B_s)B_v]=\mathrm{E}[(B_t-B_s)B_v]-\int_s^t\frac{\mathrm{E}[(B_1-B_u)B_v]}{1-u}\,du=0\ .$$

c2) Let us prove that $\mathrm{E}[\widetilde B_t\widetilde B_s]=s$, for $0\le s\le t$. This is elementary, albeit laborious. If $s<t\le 1$, note that $\mathrm{E}[\widetilde B_t\widetilde B_s]=\mathrm{E}[(\widetilde B_t-\widetilde B_s)\widetilde B_s]+\mathrm{E}[\widetilde B_s^2]=\mathrm{E}[\widetilde B_s^2]$,

thanks to c1). Hence we are reduced to the computation of

$$\begin{aligned}
&\mathrm{E}[\widetilde{B}_s^2]\\
&= \mathrm{E}(B_s^2) - 2\int_0^s \mathrm{E}\Big[B_s\frac{B_1-B_u}{1-u}\Big]\,du + \int_0^s dv\int_0^s \mathrm{E}\Big[\frac{B_1-B_v}{1-v}\,\frac{B_1-B_u}{1-u}\Big]\,du\\
&= s - 2\int_0^s \frac{s-u}{1-u}\,du + \int_0^s dv\int_0^s \frac{1-u-v+u\wedge v}{(1-v)(1-u)}\,du\\
&= s - 2I_1 + I_2\ .
\end{aligned}$$

With patience one can compute $I_2$ and find that it is equal to $2I_1$, which gives the result. The simplest way to check that $I_2 = 2I_1$ is to observe that the integrand in $I_2$ is a function of $(u, v)$ that is symmetric in $u, v$. Hence

$$\begin{aligned}
I_2 &= \int_0^s dv\int_0^s \frac{1-u-v+u\wedge v}{(1-v)(1-u)}\,dv = 2\int_0^s dv\int_v^s \frac{1-u-v+u\wedge v}{(1-v)(1-u)}\,du\\
&= 2\int_0^s dv\int_v^s \frac{1-u}{(1-v)(1-u)}\,du = 2\int_0^s \frac{s-v}{1-v}\,dv = 2I_1\ .
\end{aligned}$$

Hence $\mathrm{E}[\widetilde{B}_t\widetilde{B}_s] = s\wedge t$. If $s \le t$, the r.v. $\widetilde{B}_t - \widetilde{B}_s$ is centered Gaussian as $(\widetilde{B}_t)_t$ is clearly a centered Gaussian process, which together with c1), completes the proof that $\widetilde{B}$ is a $(\widetilde{\mathscr{G}}_t)_t$-Brownian motion.

c3) We have, of course,

$$dB_t = A_t\,dt + d\widetilde{B}_t$$

with

$$A_t = \frac{B_1 - B_t}{1-t}\ \cdot$$

Hence, since $(A_t)_t$ is $(\widetilde{\mathscr{G}}_t)_t$-adapted, $B$ is an Ito process with respect to the new Brownian motion $\widetilde{B}$.

**8.18**

a) If there existed a second process $(Y'_s)_s$ satisfying (8.56), we would have

$$\begin{aligned}
0 &= \mathrm{E}\Big[\Big(\int_0^T Y_s\,dB_s - \int_0^T Y'_s\,dB_s\Big)^2\Big] = \mathrm{E}\Big[\Big(\int_0^T (Y_s - Y'_s)\,dB_s\Big)^2\Big]\\
&= \mathrm{E}\Big[\int_0^T (Y_s - Y'_s)^2\,ds\Big]
\end{aligned}$$

and therefore $Y_s = Y'_s$ a.e. with probability 1 and the two processes would be indistinguishable.

b1) As usual, for $s \le T$,

$$X_s = \mathrm{E}[(B_s + (B_T - B_s))^3 \,|\, \mathscr{F}_s]$$
$$= B_s^3 + 3B_s^2 \mathrm{E}[B_T - B_s] + 3B_s \mathrm{E}[(B_T - B_s)^2] + \mathrm{E}[(B_T - B_s)^3] = B_s^3 + 3B_s(T - s)\,.$$

By Ito's formula,

$$dX_s = 3B_s^2\, dB_s + 3B_s\, ds + 3(T - s)\, dB_s - 3B_s\, ds = 3(B_s^2 + (T - s))\, dB_s\,.$$

Note that the part in $ds$ vanishes, which is not surprising as $(X_s)_s$ is clearly a martingale.

b2) Obviously

$$B_T^3 = X_T = \int_0^T \underbrace{3(B_s^2 + (T - s))}_{:=Y_s}\, dB_s\,.$$

c) We can repeat the arguments of b1) and b2): as $s \mapsto \mathrm{e}^{\sigma B_s - \frac{\sigma^2}{2}s}$ is a martingale,

$$X_s = \mathrm{E}[\mathrm{e}^{\sigma B_T} \,|\, \mathscr{F}_s] = \mathrm{e}^{\frac{\sigma^2}{2}T} \mathrm{E}[\mathrm{e}^{\sigma B_T - \frac{\sigma^2}{2}T} \,|\, \mathscr{F}_s] = \mathrm{e}^{\frac{\sigma^2}{2}T} \mathrm{e}^{\sigma B_s - \frac{\sigma^2}{2}s} = \mathrm{e}^{\sigma B_s + \frac{\sigma^2}{2}(T-s)}$$

and

$$dX_s = \mathrm{e}^{\frac{\sigma^2}{2}T} \sigma \mathrm{e}^{\sigma B_s - \frac{\sigma^2}{2}s}\, dB_s$$

and therefore

$$\mathrm{e}^{\sigma B_T} = X_T = X_0 + \sigma \int_0^T \mathrm{e}^{\sigma B_s + \frac{\sigma^2}{2}(T-s)}\, dB_s = \mathrm{e}^{\frac{\sigma^2}{2}T} + \int_0^T \underbrace{\sigma \mathrm{e}^{\sigma B_s + \frac{\sigma^2}{2}(T-s)}}_{:=Y_s}\, dB_s\,.$$

**8.19**

a1) We have

$$M_t = \mathrm{E}[Z|\mathscr{F}_t] = \mathrm{E}\Big[\int_0^T B_s\, ds \,\Big|\, \mathscr{F}_t\Big] = \int_0^t B_s\, ds + \mathrm{E}\Big[\int_t^T B_s\, ds \,\Big|\, \mathscr{F}_t\Big]$$
$$= \int_0^t B_s\, ds + (T - t)B_t\,.$$

a2) Ito's formula gives

$$dM_t = B_t\,dt + (T-t)\,dB_t - B_t\,dt = (T-t)\,dB_t$$

and therefore, as $M_0 = 0$,

$$\int_0^T B_s\,ds = M_T = \int_0^T \underbrace{(T-t)}_{:=X_t}\,dB_t\ .$$

b) Let again

$$M_t = \mathrm{E}\Big[\int_0^T B_s^2\,ds \,\Big|\, \mathscr{F}_t\Big] = \int_0^t B_s^2\,ds + \mathrm{E}\Big[\int_t^T B_s^2\,ds \,\Big|\, \mathscr{F}_t\Big]\ .$$

Now, recalling that $t \mapsto B_t^2 - t$ is a martingale,

$$\int_t^T \mathrm{E}[B_s^2|\mathscr{F}_t]\,ds = \int_t^T (\mathrm{E}[B_s^2 - s|\mathscr{F}_t] + s)\,ds = (T-t)(B_t^2 - t) + \frac{1}{2}\,(T^2 - t^2)$$

and putting things together

$$M_t = \int_0^t B_s^2\,ds + (T-t)(B_t^2 - t) + \frac{1}{2}\,(T^2 - t^2)\ .$$

As $dB_t^2 = 2B_t\,dB_t + dt$,

$$dM_t = B_t^2\,dt + (T-t)(2B_t\,dB_t + dt - dt) - (B_t^2 - t)\,dt - t\,dt = 2(T-t)B_t\,dB_t$$

and since $M_0 = \frac{T^2}{2}$ we have

$$\int_0^T B_s^2\,ds = M_T = \frac{T^2}{2} + 2\int_0^T (T-t)B_t\,dB_t\ .$$

**8.20** Ito's formula applied to the unknown process $Z$ gives

$$dZ_t = Z_t(\theta dM_t - dA_t) + \frac{\theta^2}{2}\,Z_t(B_i(t)^2 + B_j(t)^2)\,dt$$
$$= \theta Z_t(B_i(t)\,dB_j(t) + B_j(t)\,dB_i(t)) + Z_t\Big(\frac{1}{2}\,\theta^2(B_i(t)^2 + B_j(t)^2)\,dt - dA_t\Big)\ .$$

$Z$ is therefore a local martingale if and only if $dA_t = \frac{1}{2}\,\theta^2\,d\langle M\rangle_t = \frac{1}{2}\,\theta^2(B_i(t)^2 + B_j(t)^2)\,dt$.

- Of course this leaves open the question of whether such a $Z$ is a true martingale. The interested reader will be able to answer this easily later using Proposition 12.2.

**8.21**

a) We have

$$\mathrm{E}[Z^2 \mathrm{e}^{\alpha Z}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 \mathrm{e}^{\alpha x^2} \mathrm{e}^{-x^2/2}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 \mathrm{e}^{-\frac{1}{2}(1-2\alpha)x^2}\, dx$$
$$= (1-2\alpha)^{-1/2} \underbrace{\frac{(1-2\alpha)^{1/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 \mathrm{e}^{-\frac{1}{2}(1-2\alpha)x^2}\, dx} = (1-2\alpha)^{-3/2}$$

as the quantity over the brace is equal to the variance of an $N(0,(1-2\alpha)^{-1})$-distributed r.v. and is therefore equal to $(1-2\alpha)^{-1}$.

b) We have $\lim_{t\to 1-} H_t(\omega) = 0$ for $\omega \notin \{B_1 = 0\}$, which is a set of probability 0. Hence $t \mapsto H_t(\omega)$ is continuous for $t \in [0,1]$ if $\omega \notin \{B_1 = 0\}$, and $H \in M^2_{loc}([0,1])$. In order to check whether $H \in M^2([0,1])$ we have, denoting by $Z$ an $N(0,1)$-distributed r.v.,

$$\mathrm{E}\Big[\int_0^1 H_s^2\, ds\Big] = \int_0^1 \frac{1}{(1-s)^3} \mathrm{E}\Big[B_s^2 \exp\Big(-\frac{B_s^2}{1-s}\Big)\Big]\, ds$$
$$= \int_0^1 \frac{1}{(1-s)^3} \mathrm{E}\Big[sZ^2 \exp\Big(-\frac{sZ}{1-s}\Big)\Big]\, ds = \int_0^1 \frac{s}{(1-s)^3}\Big(1+\frac{2s}{1-s}\Big)^{-3/2} ds$$
$$= \int_0^1 \frac{s}{(1-s)^{3/2}(1+s)^{3/2}} = +\infty\,,$$

so that $H \notin M^2([0,1])$.

c) Let

$$f(x,t) = \frac{1}{\sqrt{1-t}} \exp\Big(-\frac{x^2}{2(1-t)}\Big)$$

and let us compute the stochastic differential, for $0 \le t < 1$, of $X_t = f(B_t, t)$. We have

$$\frac{\partial f}{\partial x}(x,t) = -\frac{x}{(1-t)^{3/2}} \exp\Big(-\frac{x^2}{2(1-t)}\Big)\,,$$

so that

$$\frac{\partial f}{\partial x}(B_t, t) = -H_t$$

and, with some patience,

$$\frac{\partial f}{\partial t}(x,t) = \Big(\frac{1}{2(1-t)^{3/2}} - \frac{x^2}{2(1-t)^{5/2}}\Big) \exp\Big(-\frac{x^2}{2(1-t)}\Big)$$
$$\frac{\partial^2 f}{\partial x^2}(x,t) = \Big(-\frac{1}{(1-t)^{3/2}} + \frac{x^2}{(1-t)^{5/2}}\Big) \exp\Big(-\frac{x^2}{2(1-t)}\Big)\,,$$

so that

$$\frac{\partial f}{\partial t}(x,t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x,t) = 0\,.$$

By Ito's formula,

$$dX_t = \Big(\frac{\partial f}{\partial t}(B_t,t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(B_t,t)\Big)\,dt + \frac{\partial f}{\partial x}(B_t,t)\,dB_t = \frac{\partial f}{\partial x}(B_t,t)\,dB_t = -H_t\,dB_t\,,$$

from which, as $X_0 = 1$, (8.57) follows. It is immediate that if $\omega \notin \{B_1 = 0\}$, then $\lim_{t\to 1-} X_t = 0$.

d) The integral

$$\int_0^1 H_s\,dB_s$$

is well defined as $H \in M^2_{loc}([0,1])$. Moreover, by continuity and thanks to (8.57),

$$\int_0^1 H_s\,dB_s = \lim_{t\to 1-}\int_0^t H_s\,dB_s = 1 - \lim_{t\to 1-} X_t = 1\,.$$

**8.22**

a) This was proved in Exercise 5.10 c).
b) In Example 8.9 it is proved that, if $X_n(t) = |B_t^{(n)}|^2$, then there exists a real Brownian motion $W$ such that

$$dX_n(t) = n\,dt + 2\sqrt{X_n(t)}\,dW_t$$

from which (8.58) follows.
c) We have, thanks to (8.58),

$$R_n(t\wedge\tau_n) - t\wedge\tau_n = \frac{2}{\sqrt{n}}\int_0^{t\wedge\tau_n}\sqrt{R_n(s)}\,dW_s\,.$$

As $\sqrt{R_n(s)} \le 1$ for $s \le \tau_n$, $Z_t = R_n(t \wedge \tau_n) - t \wedge \tau_n$ is a square integrable martingale. By Doob's inequality,

$$\begin{aligned}\mathrm{E}\Big[\sup_{0\le t\le\tau_n} |R_n(t) - t|^2\Big] &= \mathrm{E}\Big(\sup_{t\ge 0} Z_t^2\Big) \le 4 \sup_{t\ge 0} \mathrm{E}(Z_t^2)\\ &= \frac{16}{n}\mathrm{E}\Big[\int_0^{\tau_n} R_n(s)\,ds\Big] \le \frac{16}{n}\mathrm{E}(\tau_n) = \frac{16}{n}\,\cdot\end{aligned}$$

In particular, $\mathrm{E}(|R_n(\tau_n)-\tau_n)|^2) \le \frac{16}{n}$ and, recalling that $R_n(\tau_n) = 1$, by Markov's inequality,

$$\mathrm{P}(|1 - \tau_n)| \ge \varepsilon) \le \frac{16}{n\varepsilon^2}\,,$$

which proves that $\tau_n \to_{n\to\infty} 1$ in probability.

d) $\pi_{n,d}(\sigma_n)$ is the law of $X_n = (B_1(\tau_n), \dots, B_d(\tau_n))$, as indicated in the hint. As $\tau_n \to_{n\to\infty} 1$ in probability, from every subsequence of $(\tau_n)_n$ we can extract a further subsequence converging to 1 a.s. Hence from every subsequence of $(X_n)_n$ we can extract a subsequence converging to $X = (B_1(1), \dots, B_d(1))$ a.s. Therefore $X_n \to N(0, I)$ in law as $n \to \infty$.

**8.23**

a) Let $f(z) = -\log|z - x|$ so that $X_t = f(B_t)$. We cannot apply Ito's formula to $f$, which is not even defined at $x$, but we can apply it to a $C^2$ function that coincides with $f$ outside the ball of radius $\frac{1}{n}$ centered at $x$. Let us compute the derivatives of $f$. As remarked in Sect. 8.5, we have for $g(z) = |z-x|$, $z \ne x$,

$$\frac{\partial g}{\partial z_i}(z) = \frac{\partial}{\partial z_i}\sqrt{\sum_{j=1}^m (z_j - x_j)^2} = \frac{z_i - x_i}{|z-x|}$$

and

$$\begin{aligned}\frac{\partial f}{\partial z_i}(z) &= -\frac{z_i - x_i}{|z-x|^2}\\ \frac{\partial^2 f}{\partial z_i^2}(z) &= -\frac{1}{|z-x|^2} + 2\frac{(z_i - x_i)^2}{|z-x|^4}\end{aligned}$$

from which $\triangle f(z) = 0$ for $z \ne x$. Hence

$$\begin{aligned}X_{t\wedge\tau_n} &= -\log|x| + \int_0^{t\wedge\tau_n} \frac{1}{2}\,\triangle f(B_s)\,ds + \int_0^{t\wedge\tau_n} f'(B_s)\,dB_s\\ &= -\log|x| - \int_0^{t\wedge\tau_n} \frac{B_1(s) - x_1}{|B(s) - x|^2}\,dB_1(s) - \int_0^{t\wedge\tau_n} \frac{B_2(s) - x_2}{|B(s) - x|^2}\,dB_2(s)\,.\end{aligned} \tag{S.49}$$

As $|B_s - x| \geq \frac{1}{n}$ for $s \leq \tau_n$, we have

$$\frac{|B_i(s) - x_i|}{|B(s) - x|^2} \leq \frac{1}{|B(s) - x|} \leq n$$

so that the integrands in (S.49) are processes in $M^2$ and $(X_{t\wedge\tau_n})_n$ is a square integrable martingale.

b) Of course, $\tau_{n,M} < +\infty$ by the Iterated Logarithm Law. Moreover, by the stopping theorem we have for every $t > 0$

$$-\log|x| = \mathrm{E}[X_{t\wedge\tau_{n,M}}]$$

and, as $-\log M \leq X_{t\wedge\tau_{n,M}} \leq -\log\frac{1}{n}$, we can apply Lebesgue's theorem, taking the limit as $t \to +\infty$. Therefore

$$-\log|x| = \mathrm{E}[X_{\tau_{n,M}}] = -\log\tfrac{1}{n}\cdot\mathrm{P}(|B_{\tau_{n,M}} - x| = \tfrac{1}{n}) - \log M\,\big(1 - \mathrm{P}(|B_{\tau_{n,M}} - x| = \tfrac{1}{n})\big)$$

from which we obtain

$$\mathrm{P}(|B_{\tau_{n,M}} - x| = \tfrac{1}{n}) = \frac{\log M - \log|x|}{\log M - \log\frac{1}{n}}\,\cdot \tag{S.50}$$

c) We have $\{\tau_n < +\infty\} \supset \{\tau_n \leq \tau_{n,M}\} = \{|B_{\tau_{n,M}} - x| = \frac{1}{n}\}$. Therefore by (S.50)

$$\mathrm{P}(\tau_n < +\infty) \geq \frac{\log M - \log|x|}{\log M - \log\frac{1}{n}}\,\cdot$$

This inequality holds for every $M$ and taking the limit as $M \to +\infty$ we obtain $\mathrm{P}(\tau_n < +\infty) = 1$. Hence $B$ visits a.s. every neighborhood of $x$. The point $x$ being arbitrary, $B$ visits every open set a.s. and is therefore recurrent.

d1) The probability $\mathrm{P}(\zeta_k < \sigma_k)$ is obtained from (S.50) by replacing $M$ with $k$ and $n$ by $k^k$. Hence

$$\mathrm{P}(\zeta_k < \sigma_k) = \frac{\log k - \log|x|}{\log k - \log\frac{1}{k^k}} = \frac{\log k - \log|x|}{(k+1)\log k}$$

and as $k \to \infty$ this probability converges to 0.

d2) Note that $\zeta_k < \tau$: before reaching $x$, $B$ must enter the ball of radius $\frac{1}{k^k}$ centered at $x$. Hence we have $\{\tau < \sigma_k\} \subset \{\zeta_k < \sigma_k\}$ and $\mathrm{P}(\tau < \sigma_k) \leq \mathrm{P}(\zeta_k < \sigma_k)$. As $\sigma_k \to +\infty$ for $k \to \infty$,

$$\mathrm{P}(\tau < +\infty) = \lim_{k\to\infty}\mathrm{P}(\tau < \sigma_k) \leq \lim_{k\to\infty}\mathrm{P}(\zeta_k < \sigma_k) = 0\,.$$

**8.24**

a) We would like to apply Ito's formula to $z \mapsto |z+x|$, which is not possible immediately, as this is not a $C^2$ function. In order to circumvent this difficulty let $g_n : \mathbb{R}^m \to \mathbb{R}$ be such that $g_n(z) = |z+x|$ for $|z+x| \geq \frac{1}{n}$ and extended for $|z+x| < \frac{1}{n}$ in such a way as to be $C^2(\mathbb{R}^m)$. For $|z+x| \geq \frac{1}{n}$ therefore

$$\frac{\partial g_n}{\partial z_i}(z) = \frac{z_i + x_i}{|z+x|}, \qquad \frac{\partial g_n^2}{\partial z_i \partial z_j}(z) = \frac{\delta_{ij}}{|z+x|} - \frac{(z_i+x_i)(z_j+x_j)}{|z+x|^3}$$

hence

$$\triangle g_n(x) = \sum_{i=1}^m \frac{\partial g_n^2}{\partial z_i^2}(x) = \frac{m-1}{|z+x|} \cdot$$

If $\tau_n = \inf\{t; |B_t + x| < \frac{1}{n}\}$, Ito's formula applied to the function $g_n$ gives

$$\begin{aligned}|B_{t\wedge\tau_n} + x| &= |x| + \int_0^{t\wedge\tau_n} \frac{1}{2} \triangle g_n(B_s)\, ds + \int_0^{t\wedge\tau_n} g_n'(B_s)\, dB_s \\ &= |x| + \int_0^{t\wedge\tau_n} \frac{m-1}{2\,|B_s+x|}\, ds + \int_0^{t\wedge\tau_n} \frac{B_s+x}{|B_s+x|}\, dB_s\,.\end{aligned}$$

By Corollary 8.2 the process

$$W_t = \int_0^t \frac{B_s+x}{|B_s+x|}\, dB_s$$

is a Brownian motion. Note, as seen in Sect. 8.5 (or in Exercise 8.23 for dimension $d = 2$), that $|B_t + x| > 0$ for every $t > 0$ a.s. and $\tau_n \to +\infty$ as $n \to \infty$. Therefore, taking the limit as $n \to \infty$ and setting $\xi = |x|$,

$$X_t = \xi + \int_0^t \frac{m-1}{2X_s}\, ds + W_t\,,$$

which gives the required stochastic differential.

b) If $f \in C_K^2(]0, +\infty[)$, then by Ito's formula

$$df(X_t) = f'(X_t)\, dX_t + \frac{1}{2} f''(X_t)\, dt = Lf(X_t)\, dt + f'(X_t)\, dW_t\,,$$

where

$$L = \frac{1}{2}\frac{d^2}{dy^2} + \frac{m-1}{2y}\frac{d}{dy} \cdot$$

As $f$ has compact support $\subset ]0,+\infty[$, its derivative is bounded and therefore the expectation of the stochastic integral $\int_0^t f'(X_s)\,dW_s$ vanishes. In conclusion

$$\frac{1}{t}\big(\mathrm{E}[f(X_t)]-f(\xi)\big)=\frac{1}{t}\int_0^t \mathrm{E}[Lf(X_s)]\,ds \underset{t\to+0}{\to} Lf(\xi)\;.$$

**9.1**

a) Starting from the explicit solution, formula (9.3), the law of $\xi_t$ is Gaussian with mean $\mathrm{e}^{-\lambda t}x$ and variance

$$\Gamma_t=\sigma^2\int_0^t \mathrm{e}^{-2\lambda(t-s)}\,ds=\frac{\sigma^2}{2\lambda}\,(1-\mathrm{e}^{-2\lambda t})\,\cdot$$

As $\lambda>0$,

$$\lim_{t\to+\infty}\mathrm{e}^{-\lambda t}x=0,\qquad \lim_{t\to+\infty}\Gamma_t=\frac{\sigma^2}{2\lambda}\,\cdot$$

This implies (Exercise 1.14) that, for every $x$, the law of $\xi_t$ converges weakly, as $t\to+\infty$, to a Gaussian law $\mu$ with mean 0 and variance $\frac{\sigma^2}{2\lambda}$.

b) This follows from Exercise 6.10 d) but let us verify this point directly. Let $\eta$ be a Gaussian r.v. with distribution $\mu$, i.e. $\mu\sim N(0,\frac{1}{2\lambda})$, and independent of $B$. Then a repetition of the arguments of Example 9.1 gives that a solution of (9.45) with the initial condition $\xi_0=\eta$ is

$$\xi_t=\underbrace{\mathrm{e}^{-\lambda t}\eta}_{=Y_1}+\underbrace{\mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{\lambda s}\sigma\,dB_s}_{=Y_2}\;.$$

$Y_1$ and $Y_2$ are independent and Gaussian, centered and with variances

$$\frac{\sigma^2}{2\lambda}\,\mathrm{e}^{-2\lambda t}\quad\text{and}\quad\frac{\sigma^2}{2\lambda}\,(1-\mathrm{e}^{-2\lambda t})\;,$$

respectively. Therefore $\xi_t$ is Gaussian, centered and with variance

$$\frac{\sigma^2}{2\lambda}\,\mathrm{e}^{-2\lambda t}+\frac{\sigma^2}{2\lambda}\,(1-\mathrm{e}^{-2\lambda t})=\frac{\sigma^2}{2\lambda}\;,$$

i.e. $\xi_t\sim\mu$.

**9.2**

a) Let us follow the idea of the variation of constants of Example 9.1. The solution of the ordinary differential equation

$$\begin{aligned} x'_t &= b(t)x_t \\ x_0 &= x \end{aligned}$$

is $x_t = \mathrm{e}^{\Lambda(t)}x$, where $\Lambda(t) = \int_0^t b(s)\,ds$. Let us look for a solution of (9.46) of the form $x_t = \mathrm{e}^{\Lambda(t)}C(t)$. One sees easily that $C$ must be the solution of

$$\mathrm{e}^{\Lambda(t)}\,dC(t) = \sigma(t)\,dB_t\;,$$

i.e.

$$C(t) = \int_0^t \mathrm{e}^{-\Lambda(s)}\sigma(s)\,dB_s\;.$$

The solution of (9.46) is therefore

$$\xi_t = \mathrm{e}^{\Lambda(t)}x + \mathrm{e}^{\Lambda(t)}\int_0^t \mathrm{e}^{-\Lambda(s)}\sigma(s)\,dB_s\;. \tag{S.51}$$

As the stochastic integral of a deterministic function is, as a function of the integration endpoint, a Gaussian process (Proposition 7.1), $\xi$ is Gaussian. Obviously $\mathrm{E}(\xi_t) = \mathrm{e}^{\Lambda(t)}x$. Let us compute its covariance function. Let $s \le t$ and let

$$Y_t = \mathrm{e}^{\Lambda(t)}\int_0^t \mathrm{e}^{-\Lambda(s)}\sigma(s)\,dB_s\;.$$

Then, by Proposition 8.5,

$$\begin{aligned} \mathrm{Cov}(\xi_t,\xi_s) &= \mathrm{E}(Y_tY_s^*) \\ &= \mathrm{E}\Big[\mathrm{e}^{\Lambda(t)}\int_0^t \mathrm{e}^{-\Lambda(u)}\sigma(u)\,dB_u\Big(\mathrm{e}^{\Lambda(s)}\int_0^s \mathrm{e}^{-\Lambda(v)}\sigma(v)\,dB_v\Big)^*\Big] \\ &= \mathrm{e}^{\Lambda(t)}\mathrm{E}\Big[\int_0^t \mathrm{e}^{-\Lambda(u)}\sigma(u)\,dB_u\Big(\int_0^s \mathrm{e}^{-\Lambda(v)}\sigma(v)\,dB_v\Big)^*\Big]\mathrm{e}^{\Lambda(s)^*} \\ &= \mathrm{e}^{\Lambda(t)}\int_0^s \mathrm{e}^{-\Lambda(u)}\sigma(u)\sigma^*(u)\,\mathrm{e}^{-\Lambda(u)^*}\,du\;\mathrm{e}^{\Lambda(s)^*}\;. \end{aligned} \tag{S.52}$$

In particular, if $m = 1$ the covariance is

$$\mathrm{e}^{\Lambda(t)}\mathrm{e}^{\Lambda(s)}\int_0^s \mathrm{e}^{-2\Lambda(u)}\sigma^2(u)\,du\;.$$

b) We have, for this equation,

$$\Lambda(t) = \int_0^t -\frac{1}{1-s}\,ds = \log(1-t)\ .$$

Therefore $e^{\Lambda(t)} = 1 - t$ and the solution of (9.47) is

$$\xi_t = (1-t)x + (1-t)\int_0^t \frac{dB_s}{1-s}\ \cdot$$

Hence $E[\xi_t] = (1-t)x$. $\xi$ is a Gaussian process with covariance function

$$K(t,s) = (1-t)(1-s)\int_0^s \frac{1}{(1-u)^2}\,du = (1-t)(1-s)\Big(\frac{1}{1-s}-1\Big) = s(1-t)\ ,$$

for $s \le t$. If $x = 0$, then $E(\xi_t) = 0$ for every $t$ and the process has the same mean and covariance functions as a Brownian bridge.

**9.3**

a1) Recall, from Example 9.1, that the solution is

$$\xi_t = e^{-\lambda t}x + e^{-\lambda t}\int_0^t e^{\lambda s}\sigma\,dB_s\ .$$

By Corollary 8.1 there exists a Brownian motion $W$ such that, if

$$A_t = \int_0^t e^{2\lambda u}\sigma^2\,du = \frac{\sigma^2}{2\lambda}\,(e^{2\lambda t}-1)\ ,$$

then

$$\int_0^t e^{\lambda s}\sigma\,dB_s = W_{A_t}\ .$$

We have

$$\overline{\lim_{t\to+\infty}}\,\frac{\xi_t}{\sqrt{\log t}} = \overline{\lim_{t\to+\infty}}\,\frac{e^{-\lambda t}x + e^{-\lambda t}W_{A_t}}{\sqrt{\log t}} = \overline{\lim_{t\to+\infty}}\,\frac{e^{-\lambda t}W_{A_t}}{\sqrt{\log t}}\ \cdot$$

Now, trying to go back to the Iterated Logarithm Law,

$$\overline{\lim_{t\to+\infty}}\,\frac{e^{-\lambda t}W_{A_t}}{\sqrt{\log t}} = \overline{\lim_{t\to+\infty}}\,\frac{e^{-\lambda t}(2A_t\log\log A_t)^{1/2}}{\sqrt{\log t}}\,\frac{W_{A_t}}{(2A_t\log\log A_t)^{1/2}}$$

and

$$\overline{\lim_{t\to+\infty}} \frac{W_{A_t}}{(2A_t \log\log A_t)^{1/2}} = 1$$

whereas

$$\lim_{t\to+\infty} \frac{e^{-\lambda t}(2A_t \log\log A_t)^{1/2}}{\sqrt{\log t}} = \lim_{t\to+\infty} (2e^{-2\lambda t}A(t))^{1/2} \lim_{t\to+\infty} \Big(\frac{\log\log A(t)}{\log t}\Big)^{1/2}$$

and we can conclude the proof as

$$\lim_{t\to+\infty} 2e^{-2\lambda t}A_t = \frac{\sigma^2}{\lambda}$$

and $\log\log A_t \sim \log t$ as $t \to +\infty$. The the case of $\underline{\lim}$ is treated in the same way.

a2) Thanks to a1), for a sequence $t_n \nearrow +\infty$ we have

$$\xi_{t_n} \ge \frac{\sigma}{2\sqrt{\lambda}} \sqrt{\log t_n}\ .$$

Therefore

$$\overline{\lim_{t\to+\infty}}\ \xi_t = +\infty\ .$$

Similarly we obtain $\underline{\lim}_{t\to+\infty} \xi_t = -\infty$. Therefore the Ornstein–Uhlenbeck process, in this case, is recurrent in the sense explained p. 243.

b) We can write

$$\xi_t = e^{-\lambda t}\Big(x + \int_0^t e^{\lambda u}\sigma\, dB_u\Big)\ .$$

The stochastic integral above is a martingale bounded in $L^2$, as its variance is equal to $\int_0^t e^{2\lambda u}\sigma^2\, du < -\frac{\sigma^2}{2\lambda}$, as $\lambda < 0$. Therefore there exists a r.v. $X$ such that

$$x + \int_0^t e^{\lambda u}\sigma\, dB_u \underset{t\to+\infty}{\to} X \quad \text{a.s.}$$

The limit $X$, moreover, is Gaussian with mean $x$ and variance $\frac{\sigma^2}{-2\lambda}$. Therefore $\lim_{t\to+\infty} \xi_t = +\infty$ on the set $A = \{X > 0\}$ and $\lim_{t\to+\infty} \xi_t = -\infty$ on $A^c = \{X < 0\}$. As $X$ is $N(x, \frac{\sigma^2}{-2\lambda})$-distributed, we have $X = x + \frac{\sigma}{\sqrt{-2\lambda}} Z$, where

$Z \sim N(0, 1)$. Hence

$$P(A) = P(X > 0) = P\Big(x + \frac{\sigma}{\sqrt{-2\lambda}} Z > 0\Big) = P\Big(Z > -\frac{x\sqrt{-2\lambda}}{\sigma}\Big)$$
$$= P\Big(Z < \frac{x\sqrt{-2\lambda}}{\sigma}\Big) = \Phi\Big(\frac{x\sqrt{-2\lambda}}{\sigma}\Big)\,.$$

**9.4**

a) This is a particular case of the general situation of Exercise 9.2 a). The solution of the "homogeneous equation"

$$\begin{aligned} d\xi_t &= -\frac{1}{2}\,\frac{\xi_t}{1-t}\,dt \\ \xi_0 &= x \end{aligned}$$

is immediately found to be equal to

$$\xi_t = \sqrt{1-t}\,x\,,$$

and the "particular solution" is given by $t \mapsto \sqrt{1-t}\,C_t$, where $C$ must satisfy

$$\sqrt{1-t}\,dC_t = \sqrt{1-t}\,dB_t\,,$$

i.e. $C_t = B_t$ and the solution of (9.48) is found to be

$$\xi_t = \sqrt{1-t}(x + B_t)\,.$$

$\xi$ is obviously a Gaussian process (this is a particular case of Exercise 9.2 a)).

b) We have

$$\mathrm{Var}(\xi_t) = (1-t)t\,,$$

which is the same as the variance of a Brownian bridge. As for the covariance function we have, for $s \le t$,

$$\mathrm{Cov}(\xi_t, \xi_s) = \mathrm{E}\big[\sqrt{1-t}\,B_t\,\sqrt{1-s}\,B_s\big] = s\sqrt{1-s}\sqrt{1-t}\,,$$

which is different from the covariance function of a Brownian bridge. Note that, if the starting point $x$ is the origin, then, for every $t, 0 \le t \le 1$, the distribution of $\xi_t$ coincides with the distribution of a Brownian bridge at time $t$, but $\xi$ is not a Brownian bridge.

**9.5**

a) Similarly to the idea of Example 9.2, if we could apply Ito's formula to the function log we would obtain the stochastic differential

$$d\log\xi_t = \frac{1}{\xi_t}\,d\xi_t - \frac{1}{2\xi_t^2}\,\sigma^2(t)\,\xi_t^2\,dt = \Big(b(t) - \frac{\sigma^2(t)}{2}\Big)\,dt + \sigma(t)\,dB_t \tag{S.53}$$

and we would have

$$\log\xi_t = \int_0^t \big(b(s) - \frac{\sigma^2(s)}{2}\big)\,ds + \int_0^t \sigma(s)\,dB_s\,,$$

which gives for (9.49) the solution

$$\xi_t = xe^{\int_0^t (b(s) - \frac{\sigma^2(s)}{2})\,ds + \int_0^t \sigma(s)\,dB_s}\,. \tag{S.54}$$

It is now immediate to check by Ito's formula correctly applied to the exponential function that the process $\xi$ of (S.54) is a solution of (9.49).

b) In case (1) $\sigma$ is square integrable, hence

$$\int_0^t \sigma(s)\,dB_s \underset{t\to+\infty}{\to} Z\,,$$

where $Z$ is a centered Gaussian r.v. with variance $\int_0^{+\infty}\sigma^2(t)\,dt = 1$. The convergence takes place in $L^2$ and also a.s. (it is a martingale bounded in $L^2$). On the other hand $b(s) - \frac{\sigma^2(s)}{2} = \frac{1}{2(1+s)}$ and

$$\int_0^{+\infty} \Big(b(s) - \frac{\sigma^2(s)}{2}\Big)\,ds = +\infty$$

and therefore $\xi_t \to +\infty$ a.s.

In case (2)

$$\int_0^{+\infty} \sigma^2(s)\,ds = 1$$

$$\int_0^{+\infty} \Big(b(s) - \frac{\sigma^2(s)}{2}\Big)\,ds = -\int_0^{+\infty} \frac{1}{6(1+s)^2}\,ds = -\frac{1}{6}\,\cdot$$

The martingale $t \mapsto \int_0^t \sigma(s)\,dB_s$ converges a.s. and in $L^2$, being bounded in $L^2$, therefore, as $t \to +\infty$, $\xi_t \to xe^Z$ where $Z \sim N(-\frac{1}{6}, 1)$, the convergence taking place also a.s.

Finally, in situation (3) we have $b(s) - \frac{\sigma^2(s)}{2} = 0$ but

$$\int_0^{+\infty} \sigma^2(s)\, ds = \int_0^{+\infty} \frac{1}{1+s}\, ds = +\infty\,.$$

Therefore, if

$$A_t = \int_0^t \sigma^2(s)\, ds = \int_0^t \frac{1}{1+s}\, ds = \log(1+t)\,,$$

then

$$\int_0^t \sigma(s)\, dB_s = W_{A_t}\,,$$

where $W$ is a new Brownian motion. As $\lim_{t\to+\infty} A_t = +\infty$, by the Iterated Logarithm Law,

$$\overline{\lim_{t\to+\infty}} \int_0^t \sigma(s)\, dB_s = +\infty, \qquad \underline{\lim_{t\to+\infty}} \int_0^t \sigma(s)\, dB_s = -\infty$$

which implies

$$\overline{\lim_{t\to+\infty}}\, \xi_t = +\infty, \qquad \underline{\lim_{t\to+\infty}}\, \xi_t = 0\,.$$

**9.6**

a) Using the same idea as in Example 9.2, applying Ito's formula formally we have

$$d\log\xi_i(t) = \frac{1}{\xi_i(t)}\, d\xi_i(t) - \frac{1}{2\xi_i(t)^2}\, \xi_i(t)^2 d\Big\langle \sum_{j=1}^d \sigma_{ij} B_j(t), \sum_{k=1}^d \sigma_{ik} B_k(t)\Big\rangle_t$$

$$= b_i\, dt + \sum_{j=1}^d \sigma_{ij}\, dB_j(t) - \frac{1}{2}\sum_{j=1}^d \sigma_{ij}^2\, dt = \Big(b_i - \frac{a_{ii}}{2}\Big)\, dt + \sum_{j=1}^d \sigma_{ij}\, dB_j(t)\,,$$

where we note $a = \sigma\sigma^*$. This would yield the solution

$$\xi_i(t) = x_i \exp\Big(\big(b_i - \frac{a_{ii}}{2}\big)t + \sum_{j=1}^d \sigma_{ij} B_j(t)\Big)\,. \tag{S.55}$$

Actually, Ito's formula cannot be applied in the way we did, log not being defined on the whole of $\mathbb{R}$. We can, however, apply Ito's formula to the exponential function and check that a process $\xi$ whose components are given by (S.55) actually is a solution of (9.50). From (S.55) we have also that if $x_i > 0$ then $\xi_i(t) > 0$ for every $t$ a.s.

Recalling the expression of the Laplace transform of Gaussian r.v.'s we have

$$\mathrm{E}[\xi_i(t)] = x_i \mathrm{e}^{(b_i - \frac{a_{ii}}{2})t} \mathrm{E}\big[\mathrm{e}^{\sum_{j=1}^d \sigma_{ij} B_j(t)}\big] = x_i \mathrm{e}^{(b_i - \frac{a_{ii}}{2})t + \frac{1}{2}\sum_{j=1}^d \sigma_{ij}^2} = x_i \mathrm{e}^{b_i t} .$$

b) We have

$$\xi_i(t)^2 \xi_j(t)^2 = \exp\Big(2\big(b_i - \frac{a_{ii}}{2}\big)t + 2\sum_{h=1}^d \sigma_{ih} B_h(t) + 2\big(b_j - \frac{a_{jj}}{2}\big)t + 2\sum_{k=1}^d \sigma_{ik} B_k(t)\Big) ,$$

which is an integrable r.v. and, moreover, is such that $t \mapsto \mathrm{E}[\xi_i(t)^2 \xi_j(t)^2]$ is continuous (again recall the expression of the Laplace transform of Gaussian r.v.'s), hence $t \mapsto \xi_i(t)\xi_j(t)$ is in $M^2$.

Moreover, by Ito's formula,

$$\begin{aligned} d\xi_i(t)\xi_j(t) &= \xi_i(t)\, d\xi_j(t) + \xi_j(t)\, d\xi_i(t) + d\langle \xi_i, \xi_j \rangle_t \\ &= b_j \xi_i(t)\xi_j(t)\, dt + \xi_i(t)\xi_j(t) \sum_{h=1}^d \sigma_{jh}\, dB_h(t) + b_i \xi_j(t)\xi_i(t)\, dt \\ &\qquad + \xi_j(t)\xi_i(t) \sum_{k=1}^d \sigma_{ik}\, dB_k(t) + \xi_i(t)\xi_j(t) a_{ij}\, dt . \end{aligned}$$

Writing the previous formula in integrated form and taking the expectation, we see that the stochastic integrals have expectation equal to 0, as the integrands are in $M^2$. We find therefore

$$\mathrm{E}[\xi_i(t)\xi_j(t)] = x_i x_j + \int_0^t \mathrm{E}[\xi_i(s)\xi_j(s)](b_i + b_j + a_{ij})\, ds .$$

If we set $v(t) = \mathrm{E}[\xi_i(t)\xi_j(t)]$, then $v$ satisfies the ordinary equation

$$v'(t) = (b_i + b_j + a_{ij})\, v(t) ,$$

hence

$$\mathrm{E}[\xi_i(t)\xi_j(t)] = v(t) = x_i x_j \mathrm{e}^{(b_i + b_j + a_{ij})t} .$$

Therefore

$$\begin{aligned} \mathrm{Cov}(\xi_i(t), \xi_j(t)) &= \mathrm{E}[\xi_i(t)\xi_j(t)] - \mathrm{E}[\xi_i(t)]\mathrm{E}[\xi_j(t)] = x_i x_j \big(\mathrm{e}^{(b_i + b_j + a_{ij})t} - \mathrm{e}^{(b_i + b_j)t}\big) \\ &= x_i x_j \mathrm{e}^{(b_i + b_j)t} \big(\mathrm{e}^{a_{ij} t} - 1\big) . \end{aligned}$$

- This is one of the possible extensions of geometric Brownian motion to a multidimensional setting.

**9.7**

a) The clever reader has certainly observed that the processes $\xi_1, \xi_2$ are both geometric Brownian motions for which an explicit solution is known, which allows us to come correctly to the right answer. Let us, however, work otherwise. Observing that the process $\langle \xi_1, \xi_2 \rangle$ vanishes,

$$d\xi_1(t)\xi_2(t) = \xi_1(t)\, d\xi_2(t) + \xi_2(t)\, d\xi_1(t)$$
$$= \xi_1(t)\big(r_2\xi_2(t)\, dt + \sigma_2\xi_2(t)\, dB_2(t)\big) + \xi_2(t)\big(r_1\xi_1(t)\, dt + \sigma_1\xi_1(t)\, dB_1(t)\big)\ ,$$

hence

$$dX_t = (r_1 + r_2)X_t\, dt + X_t\big(\sigma_1\, dB_1(t) + \sigma_2\, dB_2(t)\big)\ .$$

If

$$W_t = \frac{1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\,\big(\sigma_1 B_1(t) + \sigma_2 B_2(t)\big)$$

then $W$ is a Brownian motion and the above relation for $dX_t$ becomes

$$dX_t = (r_1 + r_2)X_t\, dt + \sqrt{\sigma_1^2 + \sigma_2^2}\, X_t\, dW_t\ .$$

$X$ is therefore also a geometric Brownian motion and

$$X_t = x_0 \mathrm{e}^{(r_1+r_2-\frac{1}{2}(\sigma_1^2+\sigma_1^2))t+\sqrt{\sigma_1^2+\sigma_2^2}\, W_t}\ . \tag{S.56}$$

In order to investigate the case $\sqrt{\xi_1(t)\xi_2(t)}$ we can take either the square root in (S.56) or compute the stochastic differential of $Z_t = \sqrt{X_t}$. The latter strategy gives

$$dZ_t = d\sqrt{X_t} = \frac{1}{2\sqrt{X_t}}\, dX_t - \frac{1}{8X_t^{3/2}}\,(\sigma_1^2 + \sigma_2^2)X_t^2\, dt$$
$$= \frac{1}{2\sqrt{X_t}}\Big((r_1 + r_2)X_t\, dt + \sqrt{\sigma_1^2 + \sigma_2^2}\, X_t\, dW_t\Big) - \frac{1}{8}\,(\sigma_1^2 + \sigma_2^2)\sqrt{X_t}\, dt$$
$$= \Big(\frac{1}{2}(r_1 + r_2) - \frac{1}{8}\,(\sigma_1^2 + \sigma_2^2)\Big)Z_t\, dt + \frac{1}{2}\,\sqrt{\sigma_1^2 + \sigma_2^2}\, Z_t\, dW_t\ ,$$

which is again a geometric Brownian motion with some new parameters $r$ and $\sigma$. This argument is not completely correct as $x \mapsto \sqrt{x}$ is not a $C^2$ function, but the clever reader has certainly learned how to go round this kind of difficulty.

b) We can repeat the previous arguments: now $d\langle \xi_1, \xi_2 \rangle_t = \rho\sigma_1\sigma_2\xi_1(t)\xi_2(t)\,dt$ so that

$$
\begin{aligned}
d\xi_1(t)\xi_2(t) &= \xi_1(t)\,d\xi_2(t) + \xi_2(t)\,d\xi_1(t) + d\langle \xi_1, \xi_2 \rangle_t \\
&= \xi_1(t)\big(r_2\xi_2(t)\,dt + \sigma_2\sqrt{1-\rho^2}\,\xi_2(t)\,dB_2(t) + \sigma_2\rho\xi_2(t)\,dB_1(t)\big) \\
&\qquad + \xi_2(t)\big(r_1\xi_1(t)\,dt + \sigma_1\xi_1(t)\,dB_1(t)\big) + \rho\sigma_1\sigma_2\xi_1(t)\xi_2(t)\big)\,dt
\end{aligned}
$$

so that, if $X_t = \xi_1(t)\xi_2(t)$,

$$
dX_t = (r_1 + r_2 + \rho\sigma_1\sigma_2)X_t\,dt + X_t\big((\sigma_1 + \rho\sigma_2)\,dB_1(t) + \sigma_2\sqrt{1-\rho^2}\,dB_2(t)\big)\;.
$$

Now we have a new Brownian motion by letting

$$
W_t = \frac{1}{\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}}\,\big((\sigma_1 + \rho\sigma_2)\,B_1(t) + \sigma_2\sqrt{1-\rho^2}\,B_2(t)\big)
$$

and we find that $X$ is again a geometric Brownian motion with stochastic differential

$$
dX_t = (r_1 + r_2 + \rho\sigma_1\sigma_2)X_t\,dt + \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}\,X_t\,dW_t\;.
$$

The same arguments as above give

$$
\begin{aligned}
dZ_t &= \Big(\frac{1}{2}(r_1 + r_2 + \rho\sigma_1\sigma_2) - \frac{1}{8}\,(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)\Big)Z_t\,dt \\
&\qquad + \frac{1}{2}\,\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}\,Z_t\,dW_t \\
&= \Big(\frac{1}{2}(r_1 + r_2) - \frac{1}{8}\,(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)\Big)Z_t\,dt + \frac{1}{2}\,\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}\,Z_t\,dW_t\;,
\end{aligned}
$$

which is again a geometric Brownian motion.

**9.8**

a) Two possibilities: by Ito's formula, for every real number $\alpha$

$$
\begin{aligned}
d\xi_t^\alpha &= \alpha\xi_t^{\alpha-1}\,d\xi_t + \frac{1}{2}\,\alpha(\alpha-1)\xi_t^{\alpha-2}\sigma^2\xi_t^2\,dt \\
&= \Big(\alpha b + \frac{\sigma^2}{2}\,\alpha(\alpha-1)\Big)\xi_t^\alpha\,dt + \sigma\xi_t^\alpha\,dB_t\;.
\end{aligned}
$$

If the coefficient of $dt$ in the stochastic differential above vanishes then $\xi^\alpha$ will be a local martingale and actually a martingale (why?). The condition is therefore $\alpha = 0$ (obviously) or

$$b + \frac{\sigma^2}{2}(\alpha - 1) = 0 ,$$

i.e.

$$\alpha = 1 - \frac{2b}{\sigma^2} \cdot \tag{S.57}$$

Second possibility: we know that $\xi$ has the explicit form

$$\xi_t = \mathrm{e}^{(b-\frac{\sigma^2}{2})t+\sigma B_t}$$

and therefore

$$\xi_t^\alpha = \mathrm{e}^{\alpha(b-\frac{\sigma^2}{2})t+\alpha\sigma B_t} ,$$

which turns out to be an exponential martingale if

$$\alpha(b - \tfrac{\sigma^2}{2}) = -\alpha^2 \frac{\sigma^2}{2}$$

and we obtain again (S.57).

Note, however, that the use of Ito's formula above requires an explanation, as the function $x \mapsto x^\alpha$ is not defined on the whole of $\mathbb{R}$.

b) By the stopping theorem we have, for every $t \ge 0$ and the value of $\alpha$ determined in a),

$$\mathrm{E}[\xi_{\tau\wedge t}^\alpha] = 1 .$$

As $t \mapsto \xi_{\tau\wedge t}^\alpha$ remains bounded, Lebesgue's theorem gives

$$1 = \mathrm{E}[\xi_\tau^\alpha] = 2^\alpha \mathrm{P}(\xi_\tau = 2) + 2^{-\alpha}\big(1 - \mathrm{P}(\xi_\tau = 2)\big) ,$$

from which we find

$$\mathrm{P}(\xi_\tau = 2) = \frac{1 - 2^{-\alpha}}{2^\alpha - 2^{-\alpha}} \cdot$$

**9.9**

a) It is immediate that $(\eta_t)_t$ is a geometric Brownian motion and

$$\eta_t = y\mathrm{e}^{(\nu-\frac{1}{2})t+B_2(t)}$$

and therefore

$$\xi_t = x + y \int_0^t e^{(\nu-\frac{1}{2})t+B_2(t)}\, dB_1(t)\ . \tag{S.58}$$

This is a martingale, as the integrand in (S.58) is in $M_2([0,T])$ for every $T > 0$ since

$$E\Big(\int_0^T e^{(2\nu-1)t+2B_2(t)}\, dt\Big) = \int_0^T e^{(2\nu-1)t+2t}\, dt < +\infty\ .$$

Alternatively, in a more immediate way, the pair $(\xi, \eta)$ is the solution of an SDE with coefficients with sublinear growth, and we know that the solutions are automatically in $M_2([0,T])$ for every $T > 0$ (Theorem 9.1).

b) Obviously $E(\xi_t) = x$, as $(\xi_t)_t$ is a martingale and $\xi_0 = x$. Moreover, if $\nu \neq -\frac{1}{2}$,

$$\begin{aligned}\mathrm{Var}(\xi_t) = E[(\xi_t - x)^2] &= E\Big[y^2 \int_0^t e^{(2\nu-1)t+2B_2(t)}\, dt\Big] = y^2 \int_0^t E[e^{(2\nu-1)t+2B_2(t)}]\, dt\\ &= y^2 \int_0^t e^{(2\nu-1)t+2t}\, dt = y^2 \int_0^t e^{(2\nu+1)t}\, dt = \frac{y^2}{2\nu+1}\,(e^{(2\nu+1)t} - 1)\end{aligned}$$

whereas, if $\nu = -\frac{1}{2}$, $\mathrm{Var}(\xi_t) = y^2 t$.

c) From b) if $\nu < -\frac{1}{2}$ then $E(\xi_t^2) = E(\xi_t)^2 + \mathrm{Var}(\xi_t) = x^2 + \mathrm{Var}(\xi_t)$ is bounded in $t$. Therefore $(\xi_t)_t$ is a martingale bounded in $L^2$ and therefore convergent a.s. and in $L^2$.

**9.10**

a) We note that (9.51) is an SDE with sublinear coefficients. Therefore its solution belongs to $M^2$ and taking the expectation we find

$$E[\xi_t] = x + \int_0^t (a + bE[\xi_s])\, ds\ .$$

Thus the function $v(t) = E[\xi_t]$ is differentiable and satisfies

$$v'(t) = a + bv(t)\ . \tag{S.59}$$

The general integral of the homogeneous equation is $v(t) = e^{bt}v_0$. Let us look for a particular solution of the form $t \mapsto e^{bt}c(t)$. The equation becomes

$$be^{bt}c(t) + e^{bt}c'(t) = a + be^{bt}c(t)\ ,$$

i.e. $c$ must satisfy

$$e^{bt}c'(t) = a\ ,$$

hence

$$c(t) = a \int_0^t \mathrm{e}^{-bs}\, ds = \frac{a}{b}\,(1 - \mathrm{e}^{-bt})$$

so that $\mathrm{e}^{bt} c(t) = \frac{a}{b}\,(\mathrm{e}^{bt} - 1)$ and, as $v(0) = \mathrm{E}[\xi_0] = x$, we find

$$\mathrm{E}[\xi_t] = v(t) = \mathrm{e}^{bt} x + \frac{a}{b}\,(\mathrm{e}^{bt} - 1)\ . \tag{S.60}$$

b1) Immediate from (S.60).

b2) (S.60) can be written as

$$\mathrm{E}[\xi_t] = \Big(x + \frac{a}{b}\Big)\mathrm{e}^{bt} - \frac{a}{b}\ ,$$

therefore if $x + \frac{a}{b} = 0$, the expectation is constant and

$$\mathrm{E}[\xi_t] \equiv -\frac{a}{b} = x\ .$$

b3) Again by (S.60) $\lim_{t\to+\infty} \mathrm{E}[\xi_t] = \pm\infty$ according to the sign of $x + \frac{a}{b}$.

**9.11**

a1) The process $C$ must satisfy

$$\begin{aligned} d(\xi_0(t)C_t) &= (a + b\xi_0(t)C_t)\, dt + (\lambda + \sigma\xi_0(t)C_t)\, dB_t \\ &= C_t\xi_0(t)\big(b\, dt + \sigma\, dB_t\big) + a\, dt + \lambda\, dB_t\ . \end{aligned}$$

Let us assume that $dC_t = \zeta_t\, dt + \rho_t\, dB_t$ for some processes $\zeta, \rho$ to be determined. We have $d\xi_0(t) = b\xi_0(t)\, dt + \sigma\xi_0(t)\, dt$, hence, by Ito's formula,

$$\begin{aligned} d(\xi_0(t)C_t) &= C_t\, d\xi_0(t) + \xi_0(t)\, dC_t + d\langle C, \xi_0\rangle_t \\ &= C_t\xi_0(t)\big(b\, dt + \sigma\, dB_t\big) + \xi_0(t)\, dC_t + \sigma\rho_t\xi_0(t)\, dt\ . \end{aligned}$$

Therefore $C$ must satisfy

$$\xi_0(t)\, dC_t + \sigma\rho_t\xi_0(t)\, dt = \xi_0(t)(\zeta_t\, dt + \rho_t\, dB_t) + \sigma\rho_t\xi_0(t)\, dt = a\, dt + \lambda\, dB_t\ .$$

From this relation we obtain

$$\begin{aligned} \rho_t &= \lambda\xi_0(t)^{-1} \\ \zeta_t &= (a - \sigma\lambda)\xi_0(t)^{-1} \end{aligned}$$

and, recalling the expression of $\xi_0$,

$$C_t = (a-\sigma\lambda)\int_0^t e^{-(b-\frac{\sigma^2}{2})s-\sigma B_s}\,ds + \lambda\int_0^t e^{-(b-\frac{\sigma^2}{2})s-\sigma B_s}\,dB_s\,,$$

which gives for (9.52) the general solution

$$\begin{aligned}\xi_t &= \xi_0(t)x + \xi_0(t)C_t\\ &= e^{(b-\frac{\sigma^2}{2})t+\sigma B_t}\Big(x + (a-\sigma\lambda)\int_0^t e^{-(b-\frac{\sigma^2}{2})s-\sigma B_s}\,ds\\ &\qquad\qquad + \lambda\int_0^t e^{-(b-\frac{\sigma^2}{2})s-\sigma B_s}\,dB_s\Big)\,.\end{aligned} \tag{S.61}$$

a2) If $\sigma = 0$ (S.61) becomes

$$\xi_t = e^{bt}\Big(x + a\int_0^t e^{-bs}\,ds + \lambda\int_0^t e^{-bs}\,dB_s\Big)\,.$$

The stochastic integral in this expression has a deterministic integrand and is therefore Gaussian. The other terms appearing are also deterministic and therefore $\xi$ is Gaussian itself. We have

$$\begin{aligned}\mathrm{E}[\xi_t] &= e^{bt}x + \frac{a}{b}\,(e^{bt}-1)\\ \mathrm{Var}(\xi_t) &= \lambda^2\int_0^t e^{2b(t-s)}\,ds = \frac{\lambda^2}{2b}\,(e^{2bt}-1)\,.\end{aligned} \tag{S.62}$$

a3) $\xi$ being Gaussian it suffices to investigate the limits as $t \to +\infty$ of its mean and its variance. (S.62) for $b < 0$ give

$$\lim_{t\to+\infty}\mathrm{E}[\xi_t] = \frac{a}{-b}, \qquad \lim_{t\to+\infty}\mathrm{Var}(\xi_t) = \frac{\lambda^2}{-2b}\,\cdot$$

Therefore, $\xi_t$ converges in law as $t \to +\infty$ to an $N(\frac{a}{-b}, \frac{\lambda^2}{-2b})$ distribution. Observe that the mean of this distribution is the point at which the drift vanishes.

b) If Ito's formula could be applied to the function log (but it cannot, as it is not even defined on the whole real line) we would obtain

$$d(\log Y_t) = \frac{1}{Y_t}\,dY_t - \frac{1}{2Y_t^2}\,d\langle Y\rangle_t = (b+\theta\log Y_t)\,dt + \sigma\,dB_t - \frac{\sigma^2}{2}\,dt\,. \tag{S.63}$$

Therefore $\xi_t = \log Y_t$ would be a solution of

$$\begin{aligned}d\xi_t &= \big(b - \tfrac{\sigma^2}{2} + \theta\xi_t\big)\,dt + \sigma\,dB_t\\ \xi_0 &= x = \log y\,,\end{aligned} \tag{S.64}$$

which is a particular case of (9.52). Once these heuristics have been performed, it is immediate to check that if $\xi$ is a solution of (S.64) then $Y_t = e^{\xi_t}$ is a solution of (S.63). We have therefore proved the existence of a solution of the SDE (9.53), but this equation does not satisfy the existence and uniqueness results of Chap. 9 (the drift does not have a sublinear growth) so that uniqueness is still to be proved.

Let $Y$ be a solution of (9.53) and $\tau_\varepsilon$ its exit time from the interval $]\varepsilon, \frac{1}{\varepsilon}[$ or, which is the same, the exit time of $\log Y_t$ from $]-\log\frac{1}{\varepsilon}, \log\frac{1}{\varepsilon}[$. Let $f_\varepsilon$ be a $C^2$ function on $\mathbb{R}$ that coincides with log on $]\varepsilon, \frac{1}{\varepsilon}[$. Then we can apply Ito's formula to $f_\varepsilon(Y_t)$, which gives that

$$\log Y_{\tau_\varepsilon\wedge t} = \log y + \int_0^{\tau_\varepsilon\wedge t} \big(b - \tfrac{\sigma^2}{2} + \theta\log(Y_{\tau_\varepsilon\wedge s})\big)\, ds + \sigma\, B_{\tau_\varepsilon\wedge t}\,.$$

Therefore by localization (Theorem 9.3) $\log Y_t$ coincides with the solution of (S.64) until the exit from the interval $]-\log\frac{1}{\varepsilon}, \log\frac{1}{\varepsilon}[$. But Eq. (S.64) is a nice equation with Lipschitz continuous coefficients and hence has a unique solution. As we know that $\tau_\varepsilon \to +\infty$ as $\varepsilon \to 0$ (Remark 9.3), the uniqueness of the solution of (S.64) implies the uniqueness of the solution of (9.53).

As $Y_t = e^{\xi_t}$, necessarily $Y_t > 0$ for every $t$ a.s. Thanks to a), $Y_t$ has a lognormal law (see Exercise 1.11 for its definition) with parameters $\mu_t$ and $\sigma_t^2$ that can be obtained from (S.62) i.e.

$$\mu_t = e^{\theta t}x + \frac{b - \frac{1}{2}\sigma^2}{\theta}\,(e^{\theta t} - 1)$$

$$\sigma_t^2 = \frac{\sigma^2}{2\theta}\,(e^{2\theta t} - 1)\,.$$

If $\theta < 0$,

$$\lim_{t\to+\infty} \mu_t = \frac{1}{-\theta}\Big(b - \frac{\sigma^2}{2}\Big), \qquad \lim_{t\to+\infty} \sigma_t^2 = \frac{\sigma^2}{-2\theta}$$

and $(Y_t)_t$ converges in law as $t \to +\infty$ to a lognormal distribution with these parameters.

**9.12**

a) Let us apply Ito's formula to the function $u(x,t) = e^{-2\lambda t}(x^2 + \frac{\sigma^2}{2\lambda})$, so that $dZ_t = du(\xi_t, t)$. Recall the formula (see Remark 9.1)

$$du(\xi_t, t) = \Big(\frac{\partial u}{\partial t} + Lu\Big)(\xi_t, t)\, dt + \frac{\partial u}{\partial x}(\xi_t, t)\,\sigma\, dB_t\,,$$

where

$$Lu(x,t) = \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2}(x,t) + \lambda x \frac{\partial u}{\partial x}(x,t)$$

is the generator of the Ornstein–Uhlenbeck process. As

$$\frac{\partial u}{\partial t} = -2\lambda\, u(x,t), \qquad \frac{\partial u}{\partial x} = 2x\mathrm{e}^{-2\lambda t}, \qquad \frac{\partial^2 u}{\partial x^2} = 2\mathrm{e}^{-2\lambda t}$$

we have

$$Lu = 2\lambda\, u(x,t)$$

and the term in $dt$ of the stochastic differential vanishes. Therefore

$$dZ_t = du(\xi_t, t) = \frac{\partial u}{\partial x}(\xi_t, t)\,\sigma\, dB_t = \mathrm{e}^{-2\lambda t} 2\sigma \xi_t\, dB_t$$

and $(Z_t)_t$ is a martingale, being the stochastic integral of a process of $M^2$.

b) We have

$$\mathrm{E}[Z_t] = Z_0 = x^2 + \frac{\sigma^2}{2\lambda}\,,$$

hence

$$\mathrm{E}[Y_t] = \mathrm{e}^{2\lambda t}\mathrm{E}[Z_t] = \mathrm{e}^{2\lambda t}\Big(x^2 + \frac{\sigma^2}{2\lambda}\Big)\,,$$

which entails that $\lim_{t\to+\infty} \mathrm{E}[Y_t]$ is equal to $+\infty$ or to 0 according as $\lambda > 0$ or $\lambda < 0$.

**9.13**

a) The coefficients are Lipschitz continuous, therefore we have strong existence and uniqueness.

b1) We have $Y_t = f(\xi_t)$ with $f(z) = \log\big(\sqrt{1+z^2} + z\big)$. As $\sqrt{1+z^2} + z > 0$ for every $z \in \mathbb{R}$, $f : \mathbb{R} \to \mathbb{R}$ is differentiable infinitely many times and

$$f'(z) = \frac{1}{\sqrt{1+z^2}+z} \times \Big(\frac{z}{\sqrt{1+z^2}} + 1\Big) = \frac{1}{\sqrt{1+z^2}}$$

$$f''(z) = -\frac{z}{(1+z^2)^{3/2}}\,\cdot$$

Ito's formula gives

$$\begin{aligned} dY_t &= \frac{1}{\sqrt{1+\xi_t^2}}\,d\xi_t - \frac{\xi_t}{2(1+\xi_t^2)^{3/2}}d\langle\xi\rangle_t \\ &= \frac{1}{\sqrt{1+\xi_t^2}}\Big(\sqrt{1+\xi_t^2}+\frac{1}{2}\,\xi_t\Big)\,dt + dB_t - \frac{\xi_t}{2\sqrt{1+\xi_t^2}}\,dt \\ &= dt + dB_t\,. \end{aligned}$$

Therefore, as $Y_0 = \log\big(\sqrt{1+x^2}+x\big)$,

$$Y_t = \log\Big(\sqrt{1+x^2}+x\Big) + t + B_t\,.$$

b2) We have $\xi_t = f^{-1}(Y_t)$, assuming that $f$ is invertible. Actually if $y = f(z)$, then

$$e^y = \sqrt{1+z^2}+z\,,$$

i.e. $e^y - z = \sqrt{1+z^2}$ and taking the square $e^{2y} - 2ze^y + z^2 = 1 + z^2$, i.e. $e^{2y} - 1 = 2ze^y$, and finally

$$z = \frac{e^y - e^{-y}}{2} = \sinh y$$

so that

$$\xi_t = \sinh\Big(\log\big(\sqrt{1+x^2}+x\big) + t + B_t\Big)\,.$$

**9.14**

a) If $\zeta_t = (\xi_t, \eta_t)$, we can write

$$d\zeta_t = -\lambda\zeta_t\,dt + \Sigma\,dB_t$$

with the initial condition $\zeta_0 = z = (x, y)$, where

$$\Sigma = \begin{pmatrix} \sigma & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{pmatrix}\,.$$

We know (Example 9.1) that the solution of this SDE is

$$\zeta_t = e^{-\lambda t}z + e^{-\lambda t}\int_0^t e^{\lambda s}\,\Sigma\,dB_s$$

whence we obtain that $\zeta_t$ is Gaussian with mean $e^{-\lambda t} z$ and covariance matrix

$$\frac{1}{2\lambda}\,(1-e^{-2\lambda t})\,\Sigma\Sigma^* \,. \tag{S.65}$$

Now

$$\Sigma\Sigma^* = \begin{pmatrix} \sigma & 0 \\ \sigma\rho & \sigma\sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} \sigma & \sigma\rho \\ 0 & \sigma\sqrt{1-\rho^2} \end{pmatrix} = \begin{pmatrix} \sigma^2 & \sigma^2\rho \\ \sigma^2\rho & \sigma^2 \end{pmatrix} ,$$

hence each of the r.v.'s $\xi_t, \eta_t$ is Gaussian with variance

$$\frac{\sigma^2}{2\lambda}\,(1-e^{-2\lambda t}) \tag{S.66}$$

and mean $e^{-\lambda t}x$ and $e^{-\lambda t}y$ respectively. In particular, their (marginal) laws coincide and do not depend on $\rho$.

b) Their joint law has already been determined: it is Gaussian with covariance matrix given by (S.65), i.e.

$$\frac{\sigma^2}{2\lambda}\,(1-e^{-2\lambda t}) \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} .$$

Their covariance is therefore equal to

$$\frac{\sigma^2}{2\lambda}\,(1-e^{-2\lambda t})\rho \,,$$

which is maximum for $\rho = 1$.

A necessary and sufficient condition for a Gaussian law to have a density with respect to Lebesgue measure is that the covariance matrix is invertible. The determinant of $\Sigma\Sigma^*$ is equal to $\sigma^2(1-\rho^2)$: the covariance matrix is invertible if and only if $\rho \neq \pm 1$.

c)

$$L = \frac{\sigma^2}{2}\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + 2\rho\frac{\partial^2}{\partial x \partial y}\Big) - \lambda x\frac{\partial}{\partial x} - \lambda y\frac{\partial}{\partial y} \cdot$$

**9.15**

a) Here the drift is $b(x) = (-\frac{1}{2}x_1, -\frac{1}{2}x_2)$ whereas the diffusion coefficient is

$$\sigma(x) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

whence the generator of the diffusion is the second-order operator

$$L = \frac{1}{2}\sum_{i,j=1}^{2} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{2} b_i(x)\frac{\partial}{\partial x_i}\,,$$

where

$$a(x) = \sigma(x)\sigma(x)^* = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \begin{pmatrix} -x_2 & x_1 \end{pmatrix} = \begin{pmatrix} x_2^2 & -x_2x_1 \\ -x_2x_1 & x_1^2 \end{pmatrix}\,,$$

i.e.

$$L = \frac{1}{2}\,x_2^2\frac{\partial^2}{\partial x_1^2} + \frac{1}{2}\,x_1^2\frac{\partial^2}{\partial x_2^2} - x_1x_2\frac{\partial^2}{\partial x_1\partial x_2} - \frac{1}{2}\,x_1\frac{\partial}{\partial x_1} - \frac{1}{2}\,x_2\frac{\partial}{\partial x_2}\,\cdot$$

It is immediate that $\det a(x) = 0$ for every $x$, hence $L$ is not elliptic. The lack of ellipticity was actually obvious from the beginning as the matrix $\sigma$ above has rank 1 so that necessarily $a = \sigma\sigma^*$ has rank 1 at most. This argument allows us to say that, in all generality, the generator cannot be elliptic if the dimension of the driving Brownian motion (1 in this case) is strictly smaller than the dimension of the diffusion.

b) Let us apply Ito's formula: if $f(x) = x_1^2 + x_2^2$, then $dY_t = df(\xi_t)$ and

$$df(\xi_t) = \frac{\partial f}{\partial x_1}(\xi_t)\,d\xi_1(t) + \frac{\partial f}{\partial x_2}(\xi_t)\,d\xi_2(t) + \frac{1}{2}\Big(\frac{\partial^2 f}{\partial x_1^2}(\xi_t)d\langle\xi_1\rangle_t + \frac{\partial^2 f}{\partial x_2^2}(\xi_t)d\langle\xi_2\rangle_t\Big)$$

as the mixed derivatives of $f$ vanish. Replacing and keeping in mind that $d\langle\xi_1\rangle_t = \xi_2(t)^2\,dt$, $d\langle\xi_2\rangle_t = \xi_1(t)^2\,dt$, we obtain

$$\begin{aligned} &df(\xi_t)\\ &= 2\xi_1(t)\big(-\tfrac{1}{2}\xi_1(t)dt - \xi_2(t)\,dB_t\big) + 2\xi_2(t)\big(-\tfrac{1}{2}\xi_2(t)dt + \xi_1(t)\,dB_t\big)\\ &\qquad + \big(\xi_2(t)^2 + \xi_1(t)^2\big)\,dt\\ &= 0 \end{aligned}$$

so that the process $Y_t = \xi_1(t)^2 + \xi_2(t)^2$ is constant and $Y_1 = 1$ a.s. The process $\xi_t = (\xi_1(t), \xi_2(t))$ takes its values in the circle of radius 1.

**9.16**

a) We note that $(Y_t)_t$ is a geometric Brownian motion and that the pair $(X_t, Y_t)$ solves the SDE

$$\begin{aligned} dX_t &= Y_t\,dt\\ dY_t &= \mu Y_t\,dt + \sigma Y_t\,dB_t \end{aligned}$$

with the initial conditions $X_0 = x$, $Y_0 = 1$. Its generator is

$$L = \frac{1}{2}\,\sigma^2 y^2\,\frac{\partial^2}{\partial y^2} + y\,\frac{\partial}{\partial x} + \mu y\,\frac{\partial}{\partial y}\,\cdot$$

b) Ito's formula (assuming for an instant that it is legitimate to use it) gives

$$d\xi_t = -\frac{X_t}{Y_t^2}\,dY_t + \frac{X_t}{Y_t^3}\,d\langle Y\rangle_t + \frac{1}{Y_t}\,dX_t = -\frac{X_t}{Y_t^2}\,Y_t\big(\mu\,dt + \sigma\,dB_t\big) + \frac{X_t}{Y_t^3}\sigma^2 Y_t^2\,dt + dt\,.$$

Actually $d\langle X, Y\rangle_t = 0$. This gives

$$d\xi_t = \big(1 + (\sigma^2 - \mu)\xi_t\big)\,dt - \sigma\xi_t\,dB_t\,. \tag{S.67}$$

However, we are not authorized to apply Ito's formula, as the function $\psi$ : $(x, y) \mapsto \frac{x}{y}$ is not defined on the whole of $\mathbb{R}^2$. We can, however, reproduce here an idea developed earlier: let, for every $\varepsilon > 0$, $\psi_\varepsilon$ be a function coinciding with $\psi$ on $\{|y| > \varepsilon\}$ and then extended so that it is $C^2$. Let $\tau_\varepsilon$ be the first time that $Y_t < \varepsilon$. Then the computation above applied to the function $\psi_\varepsilon$ gives

$$\xi_{t\wedge\tau_\varepsilon} = x + \int_0^{t\wedge\tau_\varepsilon} \big(1 + (\sigma^2 - \mu)\xi_u\big)\,du - \int_0^{t\wedge\tau_\varepsilon} \sigma\xi_u\,dB_u\,.$$

As $\mathrm{P}(Y_t > 0$ for every $t \geq 0) = 1$, we have $\tau_\varepsilon \to +\infty$ as $\varepsilon \to 0$ and, taking the limit as $\varepsilon \to 0$ in the relation above, we find

$$\xi_t = x + \int_0^t \big(1 + (\sigma^2 - \mu)\xi_u\big)\,du - \int_0^t \sigma\xi_u\,dB_u\,.$$

c) Equation (9.55) is a particular case of (S.67) with $\mu = \sigma^2 - r$. With this choice of $\mu$ we have $\mu - \frac{\sigma^2}{2} = \frac{\sigma^2}{2} - r$ and $Y_u = \mathrm{e}^{(\frac{\sigma^2}{2}-r)u+\sigma B_u}$ so that a solution is, taking $x = z$,

$$Z_t = \frac{X_t}{Y_t} = \mathrm{e}^{-(\frac{\sigma^2}{2}-r)t-\sigma B_t}\Big(z + \int_0^t \mathrm{e}^{(\frac{\sigma^2}{2}-r)u+\sigma B_u}\,du\Big)\,.$$

Recalling the expression of the Laplace transform of a Gaussian r.v. we have, for $r \neq 0$,

$$\begin{aligned}
\mathrm{E}[Z_t] &= z\,\mathrm{E}[\mathrm{e}^{-(\frac{\sigma^2}{2}-r)t-\sigma B_t}] + \int_0^t \mathrm{E}[\mathrm{e}^{-(\frac{\sigma^2}{2}-r)(t-u)-\sigma(B_t-B_u)}]\,du\\
&= z\,\mathrm{e}^{-(\frac{\sigma^2}{2}-r)t+\frac{\sigma^2}{2}t} + \int_0^t \mathrm{e}^{-(\frac{\sigma^2}{2}-r)(t-u)+\frac{\sigma^2}{2}(t-u)}\,du\\
&= z\,\mathrm{e}^{rt} + \int_0^t \mathrm{e}^{r(t-u)}\,du = z\,\mathrm{e}^{rt} + \frac{1}{r}\,(\mathrm{e}^{rt} - 1)\,.
\end{aligned}$$

If $r > 0$ we have $\mathrm{E}[Z_t] \equiv -\frac{1}{r}$ if $z = -\frac{1}{r}$ and

$$\lim_{t\to+\infty} \mathrm{E}[Z_t] = \begin{cases} +\infty & \text{if } z > -\frac{1}{r} \\ -\infty & \text{if } z < -\frac{1}{r}\,. \end{cases}$$

If $r < 0$ then $\lim_{t\to+\infty} \mathrm{E}[Z_t] = -\frac{1}{r}$. whatever the starting point $z$.

If $r = 0$ the computation of the integrals above is different and we obtain

$$\mathrm{E}[Z_t] = t + z\,,$$

which converges to $+\infty$ as $t \to +\infty$.

**9.17**

a) Ito's formula would give

$$dZ_t = -\frac{1}{\xi_t^2}\,d\xi_t + \frac{1}{\xi_t^3}\,d\langle\xi\rangle_t = -\frac{1}{\xi_t^2}\,\xi_t(a - b\xi_t)\,dt - \frac{1}{\xi_t^2}\,\sigma\xi_t\,dB_t + \frac{1}{\xi_t^3}\,\sigma^2\xi_t^2\,dt\,,$$

i.e.

$$dZ_t = \big(b - (a - \sigma^2)Z_t\big)\,dt - \sigma Z_t\,dB_t \tag{S.68}$$

at least as soon as $\xi$ remains far from 0. Of course $Z_0 = z$ with $z = \frac{1}{x} > 0$.

b) Exercise 9.11 gives for (S.68) the solution

$$Z_t = \mathrm{e}^{-(a-\frac{\sigma^2}{2})t-\sigma B_t}\Big(z + b\int_0^t \mathrm{e}^{(a-\frac{\sigma^2}{2})s+\sigma B_s}\,ds\Big)\,.$$

As $z > 0$, clearly $Z_t > 0$ for every $t$ a.s. We can therefore apply Ito's formula and compute the stochastic differential of $t \mapsto \frac{1}{Z_t}$. As the function $z \to \frac{1}{z}$ is not everywhere defined, this will require the trick already exploited in other exercises (Exercise 9.16 e.g.): let $\tau_\varepsilon$ be the first time $Z_t < \varepsilon$ and let $\psi_\varepsilon$ be a $C^2$ function coinciding with $z \mapsto \frac{1}{z}$ for $z \geq \varepsilon$. Then Ito's formula gives

$$\begin{aligned}\psi_\varepsilon(Z_{t\wedge\tau_\varepsilon}) = \psi_\varepsilon(z) &+ \int_0^{t\wedge\tau_\varepsilon}\Big(\psi_\varepsilon'(Z_s)(b - (a - \sigma^2)Z_s) + \frac{1}{2}\psi_\varepsilon''(Z_s)\sigma^2 Z_s^2\Big)\,ds \\ &+ \int_0^{t\wedge\tau_\varepsilon} -\psi_\varepsilon'(Z_s)\sigma Z_s\,dB_s\,.\end{aligned}$$

As $\psi_\varepsilon(z) = \frac{1}{z}$ for $z \geq \varepsilon$, we have for $\varepsilon < z$

$$\begin{aligned}\frac{1}{Z_{t\wedge\tau_\varepsilon}} &= \frac{1}{z} + \int_0^{t\wedge\tau_\varepsilon}\Big(-\frac{1}{Z_s^2}\,(b - (a - \sigma^2)Z_s) + \frac{1}{Z_s^3}\sigma^2 Z_s^2\Big)\,ds + \int_0^{t\wedge\tau_\varepsilon}\frac{1}{Z_s^2}\,\sigma Z_s\,dB_s \\ &= \frac{1}{z} + \int_0^{t\wedge\tau_\varepsilon}\frac{1}{Z_s}\Big(a - b\frac{1}{Z_s}\Big)\,ds + \int_0^{t\wedge\tau_\varepsilon}\frac{1}{Z_s}\,\sigma\,dB_s\,.\end{aligned}$$

As $Z_t > 0$ for every $t$, we have $\tau_\varepsilon \to +\infty$ as $\varepsilon \to 0$ and the process $\xi_t = \frac{1}{Z_t}$ satisfies (9.56). Hence (9.56) has the explicit solution

$$\xi_t = \frac{1}{Z_t} = \mathrm{e}^{(a-\frac{\sigma^2}{2})t+\sigma B_t}\Big(\frac{1}{x} + b\int_0^t \mathrm{e}^{(a-\frac{\sigma^2}{2})s+\sigma B_s}\,ds\Big)^{-1}\cdot$$

Of course, $\xi_t > 0$ for every $t$ with probability 1.

**9.18**

a) Of course the pair $(X_t, Y_t)_t$ solves the SDE

$$\begin{aligned} d\xi_t &= \gamma\eta_t\,dt + \sigma\,dB_t \\ d\eta_t &= \xi_t\,dt \end{aligned}$$

with the initial conditions $\xi_0 = x$, $\eta_0 = 0$. This is an SDE with Lipschitz continuous coefficients with a sublinear growth and therefore has a unique solution. The generator is

$$L = \frac{\sigma^2}{2}\,\frac{\partial^2}{\partial x^2} + \gamma y\,\frac{\partial}{\partial x} + x\,\frac{\partial}{\partial y}\cdot$$

The process $Z_t = (\xi_t, \eta_t)$ is the solution of the SDE

$$dZ_t = MZ_t\,dt + \Sigma\,dB_t\ ,$$

where $M$, $\Sigma$ are the matrices

$$M = \begin{pmatrix} 0 & \gamma \\ 1 & 0 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}\ .$$

We recognize the Ornstein–Uhlenbeck process of Example 9.1: a linear drift and a constant diffusion matrix. Going back to (9.5), the solution is a Gaussian process and we have the explicit formula

$$Z_t = \mathrm{e}^{-Mt}z + \mathrm{e}^{-Mt}\int_0^t \mathrm{e}^{Ms}\,\Sigma\,dB_s\ ,$$

where $z = \binom{x}{0}$.

b) We have

$$M^2 = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} = \begin{pmatrix} ab & 0 \\ 0 & ab \end{pmatrix}$$

and by recurrence

$$M^{2n} = \begin{pmatrix} (ab)^n & 0 \\ 0 & (ab)^n \end{pmatrix} \qquad M^{2n+1} = \begin{pmatrix} 0 & a(ab)^n \\ b(ab)^n & 0 \end{pmatrix} .$$

Therefore the elements in the diagonal of $e^M$ are

$$\sum_{n=0}^{\infty} \frac{(ab)^n}{(2n)!} = \sum_{n=0}^{\infty} \frac{(\sqrt{ab}\,)^{2n}}{(2n)!} = \cosh(\sqrt{ab})$$

whereas the $(1,2)$-entry is

$$a\sum_{n=0}^{\infty} \frac{(ab)^n}{(2n+1)!} = \sqrt{\frac{a}{b}} \sum_{n=0}^{\infty} \frac{(\sqrt{ab}\,)^{2n+1}}{(2n+1)!} = \sqrt{\frac{a}{b}}\, \sinh(\sqrt{ab}) .$$

A similar computation allows us to obtain the $(2,1)$-entry and we find

$$e^M = \begin{pmatrix} \cosh\sqrt{ab} & \sqrt{\frac{a}{b}}\sinh\sqrt{ab} \\ \sqrt{\frac{b}{a}}\sinh\sqrt{ab} & \cosh\sqrt{ab} \end{pmatrix}$$

and (9.57) follows from the relations $\cosh ix = \cos x$, $\sinh ix = i\sin x$.

c) The mean of $Z_t$ is $e^{Mt}z$ with $z = \binom{x}{0}$. We have

$$Mt = \begin{pmatrix} 0 & \gamma t \\ t & 0 \end{pmatrix}$$

and, thanks to b2),

$$e^{Mt} = \begin{pmatrix} \cosh(\sqrt{\gamma}\,t) & \sqrt{\gamma}\sinh(\sqrt{\gamma}\,t) \\ \frac{1}{\sqrt{\gamma}}\sinh(\sqrt{\gamma}\,t) & \cosh(\sqrt{\gamma}\,t) \end{pmatrix}$$

so that $E[\xi_t] = \cosh(\sqrt{\gamma}\,t)x$. If $\gamma > 0$ this tends to $\pm\infty$ according to the sign of $x$. If $\gamma < 0$ then $E[\xi_t] = \cos(\sqrt{-\gamma}\,t)x$. This quantity oscillates between the values $x$ and $-x$ with fast oscillations if $|\gamma|$ is large. The mean remains identically equal to 0 if $x = 0$.

Let us now look at the variance of the distribution of $\xi_t$: we must compute

$$\int_0^t e^{M(t-u)}\,\Sigma\Sigma^* e^{M^*(t-u)}\,du .$$

As

$$\Sigma\Sigma^* = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 0 \end{pmatrix}$$

let us first compute the integrand:

$$\mathrm{e}^{M(t-u)}\Sigma\Sigma^* = \begin{pmatrix} \cosh(\sqrt{\gamma}\,(t-u)) & \sqrt{\gamma}\sinh(\sqrt{\gamma}\,(t-u)) \\ \frac{1}{\sqrt{\gamma}}\sinh(\sqrt{\gamma}\,(t-u)) & \cosh(\sqrt{\gamma}\,(t-u)) \end{pmatrix}\begin{pmatrix} \sigma^2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^2\cosh(\sqrt{\gamma}\,(t-u)) & 0 \\ \frac{\sigma^2}{\sqrt{\gamma}}\sinh\sqrt{\gamma}\,(t-u) & 0 \end{pmatrix}$$

so that

$$\mathrm{e}^{M(t-u)}\Sigma\Sigma^*\mathrm{e}^{M^*(t-u)}$$

$$= \begin{pmatrix} \sigma^2\cosh(\sqrt{\gamma}\,(t-u)) & 0 \\ \frac{\sigma^2}{\sqrt{\gamma}}\sinh(\sqrt{\gamma}\,(t-u)) & 0 \end{pmatrix}\begin{pmatrix} \cosh(\sqrt{\gamma}\,(t-u)) & \frac{1}{\sqrt{\gamma}}\sinh(\sqrt{\gamma}\,(t-u)) \\ \sqrt{\gamma}\sinh(\sqrt{\gamma}\,(t-u)) & \cosh(\sqrt{\gamma}\,(t-u)) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^2\cosh^2(\sqrt{\gamma}\,(t-u)) & \frac{\sigma^2}{\sqrt{\gamma}}\cosh(\sqrt{\gamma}\,(t-u))\sinh(\sqrt{\gamma}\,(t-u)) \\ \frac{\sigma^2}{\sqrt{\gamma}}\cosh(\sqrt{\gamma}\,(t-u))\sinh(\sqrt{\gamma}\,(t-u)) & \frac{\sigma^2}{\gamma}\sinh^2(\sqrt{\gamma}\,(t-u)) \end{pmatrix}\cdot$$

The variance of $\xi_t$ is therefore equal to

$$\sigma^2\int_0^t \cosh^2(\sqrt{\gamma}\,(t-u))\,du = \sigma^2\int_0^t \cosh^2(\sqrt{\gamma}\,u)\,du$$

and it diverges as $t \to +\infty$. If $\gamma < 0$ the integral grows linearly fast (the integrand becomes $\cos^2(\sqrt{-\gamma}\,u)$ and is bounded), whereas if $\gamma > 0$ it grows exponentially fast.

**9.19** By Ito's formula

$$d\eta_t = 2\xi_t\,d\xi_t + \sigma^2\,dt = (\sigma^2 + 2b\xi_t^2)\,dt + 2\sigma\xi_t\,dB_t\,.$$

If we define

$$W_t = \int_0^t \Big(\frac{\xi_s}{\sqrt{\eta_s}}\,1_{\{\eta_s\neq 0\}} + 1_{\{\eta_s=0\}}\Big)\,dB_s \tag{S.69}$$

then, by Corollary 8.2, $W$ is a real Brownian motion (the integrand is a process having modulus equal to 1) and

$$\sqrt{\eta_t}\,dW_t = \sqrt{\eta_t}\Big(\frac{\xi_t}{\sqrt{\eta_t}}\,1_{\{\eta_t\neq 0\}} + 1_{\{\eta_t=0\}}\Big)\,dB_t = \xi_t\,dB_t\,.$$

Hence $\eta$ solves the SDE

$$d\eta_t = (\sigma^2 + 2b\eta_t)\,dt + 2\sigma\sqrt{\eta_t}\,dW_t \tag{S.70}$$

of course with the initial condition $\eta_0 = x^2$.

- Note that the coefficients of (S.70) do not satisfy Condition (A') (the diffusion coefficient is not Lipschitz continuous at 0). In particular, in this exercise we prove that (S.70) has a solution but are unable to discuss uniqueness.

**9.20** Given the process

$$\xi_t = \int_0^t \sigma(\xi_s)\,dB_s$$

we know that there exists a Brownian motion $W$ such that $\xi_t = W_{A_t}$, where $A_t = \int_0^t \sigma(\xi_t)^2\,dt$. By the Iterated Logarithm Law therefore there exist arbitrarily large values of $t$ such that

$$\xi_t = W_{A_t} \ge (1-\varepsilon)\sqrt{2A_t \log\log A_t)}\;.$$

As $A_t \ge c^2 t$, $\xi_t$ takes arbitrarily large values a.s. and therefore exits with probability 1 from any bounded interval.

As we are under hypotheses of sublinear growth, the process $\xi$ and also $(\sigma(\xi_t))_t$ belong to $M^2$ and therefore $\xi$ is a square integrable martingale. By the stopping theorem therefore

$$\mathrm{E}(\xi_{\tau\wedge t}) = 0$$

for every $t > 0$. Taking the limit with Lebesgue's theorem ($t \mapsto \xi_{\tau\wedge t}$ is bounded as $\xi_{\tau\wedge t} \in [-a, b]$) we find

$$0 = \mathrm{E}(\xi_\tau) = -a\mathrm{P}(\xi_\tau = -a) + b\mathrm{P}(\xi_\tau = b)$$

and, replacing $\mathrm{P}(\xi_\tau = -a) = 1 - \mathrm{P}(\xi_\tau = b)$,

$$\mathrm{P}(\xi_\tau = b) = \frac{a}{a+b}\,\cdot$$

Note that this result holds for any $\sigma$, provided it satisfies the conditions indicated in the statement. This fact has an intuitive explanation: the solution $\xi$ being a time changed Brownian motion, its exit position from $]-a, b[$ coincides with the exit position of the Brownian motion.

**9.21**

a) Let us first look for a function $u$ such that $t \mapsto u(\xi_t)$ is a martingale. We know (see Remark 9.1) that $u$ must solve $Lu = 0$, $L$ denoting the generator of the Ornstein–Uhlenbeck process, i.e.

$$\frac{\sigma^2}{2}\,u''(x) - \lambda x u'(x) = 0\,.$$

If $v = u'$, then $v$ must solve $v'(x) = \frac{2\lambda}{\sigma^2} x\,v(x)$, i.e. $v(x) = c_1 \mathrm{e}^{\frac{2\lambda}{\sigma^2}x^2}$, so that the general integral of the ODE above is

$$u(x) = c_1 \underbrace{\int_{z_0}^{x} \mathrm{e}^{\frac{2\lambda}{\sigma^2}y^2}\,dy}_{:=F(x)} + c_2 \tag{S.71}$$

for any $z_0 \in \mathbb{R}$. As $u$ is a regular function, with all its derivatives bounded in every bounded interval, $t \mapsto u(\xi_{t\wedge\tau})$ is a martingale. This already proves that $\tau < +\infty$ a.s. Actually this is a bounded martingale, therefore a.s. convergent, and convergence cannot take place on $\tau = +\infty$ (see Sect. 10.2 for another proof of this fact).

We have therefore, for every starting position $x$, $u(x) = \mathrm{E}^x[u(\xi_{t\wedge\tau})]$ and taking the limit as $t \to +\infty$ with Lebesgue's theorem, we find

$$\begin{aligned} c_1F(x) + c_2 = u(x) = \mathrm{E}^x[u(\xi_\tau)] &= u(b)\mathrm{P}^x(\xi_\tau = b) + u(-a)(1 - \mathrm{P}^x(\xi_\tau = b))\\ &= c_1\big(F(b) - F(-a)\big)\mathrm{P}^x(\xi_\tau = b) + c_1F(-a) + c_2\,. \end{aligned}$$

Let us choose $z_0 = -a$ in (S.71), so that $F(-a) = 0$. Then the previous relation becomes

$$\mathrm{P}^x(\xi_\tau = b) = \frac{F(x)}{F(b)} = \frac{\displaystyle\int_{-a}^{x} \mathrm{e}^{\frac{2\lambda}{\sigma^2}z^2}\,dz}{\displaystyle\int_{-a}^{b} \mathrm{e}^{\frac{2\lambda}{\sigma^2}z^2}\,dz}\,\cdot \tag{S.72}$$

b1) With the change of variable $y = \sqrt{\lambda}z$ we have

$$\lim_{\lambda\to+\infty} \frac{\displaystyle\int_0^a \mathrm{e}^{\lambda z^2}\,dz}{\displaystyle\int_0^b \mathrm{e}^{\lambda z^2}\,dz} = \lim_{\lambda\to+\infty} \frac{\displaystyle\int_0^{a\sqrt{\lambda}} \mathrm{e}^{y^2}\,dy}{\displaystyle\int_0^{b\sqrt{\lambda}} \mathrm{e}^{y^2}\,dy}\,\cdot$$

Taking the derivative with respect to $\lambda$ and applying L'Hospital rule this limit is equal to

$$\lim_{\lambda\to+\infty} \frac{a}{b}\,\frac{e^{\lambda a^2}}{e^{\lambda b^2}}\,,$$

which is equal to $+\infty$ if $a > b$.

b2) Using the symmetry of the integrand we can write (S.72) as

$$P^x(\xi_\tau = b) = \frac{\displaystyle\int_0^a e^{\frac{2\lambda}{\sigma^2}z^2}\,dz + \int_0^x e^{\frac{2\lambda}{\sigma^2}z^2}\,dz}{\displaystyle\int_0^a e^{\frac{2\lambda}{\sigma^2}z^2}\,dz + \int_0^b e^{\frac{2\lambda}{\sigma^2}z^2}\,dz}$$

and now one only has to divide numerator and denominator by $\int_0^a e^{\frac{2\lambda}{\sigma^2}z^2}\,dz$: as both $b$ and $|x|$ are smaller then $a$, the limit as $\lambda\to+\infty$ turns out to be equal to 1 thanks to b1).

b3) For large $\lambda$ the process is affected by a large force (the drift $-\lambda x$) taking it towards 0. Therefore its behavior is as follows: the process is attracted towards 0 and stays around there until some unusual increment of the Brownian motion takes it out of $]-a,b[$. The exit takes place mostly at $b$ because this is the closest to 0 of the two endpoints.

**9.22**

a) We know (see Example 9.1) that the law of $\xi_t^\varepsilon$ is Gaussian with mean $e^{-\lambda t}x$ and variance

$$\frac{\varepsilon^2\sigma^2}{2\lambda}\,(1-e^{-2\lambda t})\,.$$

By Chebyshev's inequality

$$P(|\xi_t^\varepsilon - e^{-\lambda t}x| \ge \delta) \le \frac{\varepsilon^2\sigma^2}{2\lambda\delta^2}\,(1-e^{-2\lambda t}) \ \underset{\varepsilon\to 0}{\to}\ 0$$

so that $\xi_t^\varepsilon$ converges to $e^{-\lambda t}x$ in probability and therefore in distribution.

b) It is sufficient to prove that

$$P\Big(\sup_{0\le t\le T}|\xi_t^\varepsilon - e^{-\lambda t}x| \ge \delta\Big) \ \underset{\varepsilon\to 0}{\to}\ 0\,.$$

Actually this entails that the probability for $\xi^\varepsilon$ to be outside of a fixed neighborhood of the path $x_0(t) = e^{-\lambda t}$ goes to 0 as $\varepsilon\to 0$. Recall the explicit expression for $\xi^\varepsilon$:

$$\xi_t^\varepsilon = \mathrm{e}^{-\lambda t}x + \varepsilon\sigma\mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{\lambda s}\,dB_s\,.$$

Therefore we are led to the following limit to be proved

$$\mathrm{P}\Big(\varepsilon \sup_{0\le t\le T}\Big|\sigma\mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{\lambda s}\,dB_s\Big| \ge \delta\Big) \underset{\varepsilon\to 0}{\to} 0$$

which is immediate since the process inside the absolute value is continuous and therefore the r.v.

$$\sup_{0\le t\le T}\Big|\sigma\mathrm{e}^{-\lambda t}\int_0^t \mathrm{e}^{\lambda s}\,dB_s\Big|$$

is finite. Note that, using the exponential inequality of martingales we might even give a speed of convergence. This will be made explicit in the general case of Exercise 9.28.

**9.23** The function $f(y) = \frac{x}{1-xy}$ is such that $f(0) = x$ and $f'(y) = f(y)^2, f''(y) = 2f(y)^3$. Therefore if it was legitimate to apply Ito's formula we would have

$$df(B_t) = f'(B_t)\,dB_t + \frac{1}{2}f''(B_t)\,dt = f(B_t)^2\,dB_t + f(B_t)^3\,dt$$

so that $\xi_t = f(B_t)$ would be a solution of (9.60). This is not correct formally because $f$ is not $C^2(\mathbb{R})$ (it is not even everywhere defined). In order to fix this point, let us assume $x > 0$ and let $f_\varepsilon$ be a function that coincides with $f$ on $]-\infty, \frac{1}{x}-\varepsilon]$ and then extended so that it is $C^2(\mathbb{R})$. If we denote by $\tau_\varepsilon$ the passage time of $B$ at $\frac{1}{x}-\varepsilon$, then Ito's formula gives, for $\xi_t = f_\varepsilon(B_t)$,

$$\xi_{t\wedge\tau_\varepsilon} = x + \int_0^{t\wedge\tau_\varepsilon} \xi_s^3\,ds + \int_0^{t\wedge\tau_\varepsilon} \xi_s^2\,dB_s\,.$$

Therefore $\xi_t = f(B_t)$ is the solution of (9.60) for every $t \le \tau_\varepsilon$. Letting $\varepsilon \searrow 0$, $\tau_\varepsilon \nearrow \tau_x$ and therefore $(\xi_t)_t$ is the solution of (9.60) on $[0, \tau_x[$.

- Note that $\lim_{t\to\tau_x-}\xi_t = +\infty$ and that, by localization, any other solution must agree with $\xi$ on $[0, \tau_\varepsilon[$ for every $\varepsilon$, hence it is not possible to have a solution defined in any interval $[0, T]$. This is an example of what can happen when the sublinear growth property of the coefficients is not satisfied.

**9.24**

a) The matrix of the second-order coefficients of $L$ is

$$a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

whose square root, $\sigma$, is of course equal to $a$ itself. $a$ is not invertible, hence $L$ is not elliptic and *a fortiori* not uniformly elliptic. The corresponding SDE is then

$$d\xi_t = b\xi_t \, dt + \sigma \, dB_t \ ,$$

where

$$\sigma = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and we know, by Example 9.1, that this SDE has the explicit solution

$$\xi_t = \mathrm{e}^{bt} x + \mathrm{e}^{bt} \int_0^t \mathrm{e}^{-bs} \sigma \, dB_s$$

and has a Gaussian law with mean $\mathrm{e}^{bt} x$ and covariance matrix

$$\Gamma_t = \int_0^t \mathrm{e}^{bs} \sigma \sigma^* \mathrm{e}^{b^* s} \, ds \ .$$

We see that $b^2 = b^3 = \cdots = 0$. Hence

$$\mathrm{e}^{bu} = \sum_{k=0}^{\infty} \frac{b^k u^k}{k!} = I + bu = \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}$$

so that

$$\mathrm{e}^{bu} \sigma \sigma^* \mathrm{e}^{b^* u} = \begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & u \\ u & u^2 \end{pmatrix}$$

and

$$\Gamma_t = \begin{pmatrix} t & \frac{t^2}{2} \\ \frac{t^2}{2} & \frac{t^3}{3} \end{pmatrix} \cdot$$

$\Gamma_t$ being invertible for $t > 0$ (its determinant is equal to $\frac{t^3}{12}$) the law of $\xi_t$ has a density with respect to Lebesgue measure.

b) The SDE now is of the same kind but with

$$b = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad e^{bu} = \begin{pmatrix} 1 & 0 \\ 0 & e^{u} \end{pmatrix}.$$

Therefore

$$e^{bu}\sigma\sigma^* e^{b^* u} = \begin{pmatrix} 1 & 0 \\ 0 & e^{u} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{u} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$\Gamma = \begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix},$$

which is not invertible, so that there is no density with respect to Lebesgue measure.

**9.25**

a) Let $B$ be an $m$-dimensional Brownian motion and $\sigma$ a square root of $a$. The SDE associated to $L$ is

$$\begin{aligned} d\xi_t &= b\xi_t\,dt + \sigma\,dB_t \\ \xi_0 &= x\,. \end{aligned}$$

We have seen in Example 9.1 that its solution is

$$\xi_t = e^{bt}x + e^{bt}\int_0^t e^{-bs}\sigma\,dB_s\,.$$

The law of $\xi_t$ is therefore Gaussian with mean $e^{bt}x$ and, recalling (S.52), covariance matrix

$$\Gamma_t = \int_0^t e^{bu} a e^{b^* u}\,du\,. \tag{S.73}$$

As the transition function $p(t, x, \cdot)$ it is nothing else than the law of $\xi_t$ with the initial condition $\xi_0 = x$, the transition function has density if and only if $\Gamma$ is invertible (see Sect. 1.7 and Exercise 1.4). If $a$ is positive definite then there exists a number $\mu > 0$ such that, for every $y \in \mathbb{R}^m$, $\langle ay, y\rangle \geq \mu|y|^2$. Therefore, if $y \neq 0$,

$$\langle \Gamma_t y, y\rangle = \int_0^t \langle e^{bu} a e^{b^* u} y, y\rangle\,du = \int_0^t \langle a e^{b^* u} y, e^{b^* u} y\rangle\,du \geq \int_0^t \mu|e^{b^* u} y|^2\,du > 0\,.$$

Actually if $y \neq 0$, we have $|e^{b^* u} y| > 0$, as the exponential of a matrix is invertible.

b) Let us first assume that there exists a non-trivial subspace contained in the kernel of $a$ and invariant with respect to the action of $b^*$. Then there exists a non-zero vector $y \in \ker a$ such that $b^{*i} y \in \ker a$ for every $i = 1, 2, \dots$ It follows that also $e^{b^* u} y \in \ker a$. Therefore for such a vector we would have

$$\langle \Gamma_t y, y \rangle = \int_0^t \langle a e^{b^* u} y, e^{b^* u} y \rangle \, du = 0$$

and $\Gamma_t$ cannot be invertible.

Conversely, if $\Gamma_t$ were not invertible, there would exist a vector $y \in \mathbb{R}^m$, $y \neq 0$, such that $\langle \Gamma_t y, y \rangle = 0$. As $\langle a e^{b^* u} y, e^{b^* u} y \rangle \geq 0$, necessarily $\langle a e^{b^* u} y, e^{b^* u} y \rangle = 0$ for every $u \leq t$ and therefore

$$a e^{b^* u} y = 0$$

for every $u \leq t$. For $u = 0$ this relation implies $y \in \ker a$. Taking the derivative with respect to $u$ at $u = 0$ we find that necessarily $a b^* y = 0$. In a similar way, taking the derivative $n$ times and setting $u = 0$ we find $a b^{*n} y = 0$ for every $n$. The subspace generated by the vectors $y, b^* y, b^{*2} y, \dots$ is non-trivial, invariant under the action of $b^*$ and contained in $\ker a$.

**9.26** Let us compute the differential of $\eta$ with Ito's formula, assuming that the function $f$ that we are looking for is regular enough. We have

$$\begin{aligned} d\eta_t &= f'(\xi_t)\, d\xi_t + \frac{1}{2} f''(\xi_t) \sigma^2(\xi_t)\, dt \\ &\Big(f'(\xi_t) b(\xi_t) + \frac{1}{2} f''(\xi_t) \sigma^2(\xi_t)\Big)\, dt + f'(\xi_t) \sigma(\xi_t)\, dB_t \,. \end{aligned} \tag{S.74}$$

Therefore, in order for (9.64) to be satisfied, we must have $f'(\xi_t)\sigma(\xi_t) = 1$. Let

$$f(z) = \int_0^z \frac{1}{\sigma(y)}\, dy \,.$$

The assumptions on $\sigma$ entail that $f \in C^2$ so that Ito's formula can be applied to $\eta_t = f(\xi_t)$. Moreover, $f$ is strictly increasing. As $f'(x) = \frac{1}{\sigma(x)}$ and $f''(x) = -\frac{\sigma'(x)}{\sigma(x)^2}$, Eq. (S.74) now becomes

$$d\eta_t = \underbrace{\Big(\frac{1}{\sigma(f^{-1}(\eta_t))} b(f^{-1}(\eta_t)) - \frac{1}{2} \sigma'(f^{-1}(\eta_t))\Big)}_{:= \widetilde{b}(\eta_t)} dt + dB_t \,,$$

with the initial condition $\eta_0 = f(x)$.

b) We have

$$B_t = \eta_t - f(x) - \int_0^t \widetilde{b}(\eta_t)\, dt\,.$$

The right-hand side is clearly measurable with respect to the $\sigma$-algebra $\sigma(\eta_u, u \le t)$, therefore $\mathscr{G}_t \subset \sigma(\eta_u, u \le t)$, where $\mathscr{G}_t = \sigma(B_u, u \le t)$ as usual. As $f$ is strictly increasing, hence invertible, $\sigma(\eta_u, u \le t) = \sigma(\xi_u, u \le t) = \mathscr{H}_t$. Hence $\mathscr{G}_t \subset \mathscr{H}_t$ and as the converse inclusion is obvious (see Remark 9.4) the two filtrations coincide.

**9.27**

a) Let us first admit the relations (9.68) and let us apply Ito's formula to the process $\xi_t = h(D_t, B_t)$. As $(D_t)_t$ has finite variation,

$$dh(D_t, B_t) = \frac{\partial h}{\partial x}(D_t, B_t) D_t'\, dt + \frac{\partial h}{\partial y}(D_t, B_t)\, dB_t + \frac{1}{2}\frac{\partial^2 h}{\partial y^2}(D_t, B_t)\, dt\,.$$

But

$$\begin{aligned}
\frac{\partial h}{\partial y}(D_t, B_t) &= \sigma(h(D_t, B_t))\\
\frac{\partial^2 h}{\partial y^2}(D_t, B_t) &= \sigma'(h(D_t, B_t))\sigma(h(D_t, B_t))\\
\frac{\partial h}{\partial x}(D_t, B_t) &= \exp\Big[\int_0^{B_t} \sigma'(h(D_t, s))\, ds\Big]\,.
\end{aligned}$$

Moreover,

$$D_t' = \Big(-\frac{1}{2}\sigma'(h(D_t, B_t))\sigma(h(D_t, B_t)) + b(h(D_t, B_t))\Big) \exp\Big[-\int_0^{B_t} \sigma'(h(D_t, s))\, ds\Big]\,.$$

Putting things together we find

$$dh(D_t, B_t) = b(h(D_t, B_t))\, dt + \sigma(h(D_t, B_t))\, dB_t\,,$$

hence $Z_t = h(D_t, B_t)$ is a solution of (9.66), which concludes the proof of a).

Let us check the two relations of the hint. We can write

$$\begin{aligned}
\frac{\partial h}{\partial y}(x, y) &= \sigma(h(x, y))\\
h(x, 0) &= x\,.
\end{aligned}$$

As we admit that $h$ is twice differentiable in $y$,

$$\frac{\partial^2 h}{\partial y^2}(x,y) = \sigma'(h(x,y))\frac{\partial h}{\partial y}(x,y) = \sigma'(h(x,y))\sigma(h(x,y)) .$$

Similarly, taking the derivative with respect to $x$,

$$\begin{aligned}\frac{\partial^2 h}{\partial y \partial x}(y,x) &= \sigma'(h(x,y))\frac{\partial h}{\partial x}(x,y)\\ \frac{\partial h}{\partial x}(x,0) &= 1 .\end{aligned}$$

Hence, if $g(y) = \frac{\partial h}{\partial x}(x,y)$, then $g$ is the solution of the linear problem

$$\begin{aligned}g'(y) &= \sigma'(h(x,y))g(y)\\ g(0) &= 1\end{aligned}$$

whose solution is

$$g(y) = \exp\Big[\int_0^y \sigma'(h(x,s))\,ds\Big] .$$

b) Let us denote by $\widetilde{f}$ the analogue of the function $f$ with $\widetilde{b}$ replacing $b$. Then clearly $\widetilde{f}(x,z) \ge f(x,z)$ for every $x, z$. If $\widetilde{D}$ is the solution of

$$\begin{aligned}\widetilde{D}'_t &= \widetilde{f}(\widetilde{D}_t, B_t)\\ \widetilde{D}_0 &= \widetilde{x} ,\end{aligned}$$

then $\widetilde{D}_t \ge D_t$ for every $t \ge 0$. As $h$ is increasing in both arguments we have

$$\widetilde{\xi}_t = h(\widetilde{D}_t, B_t) \ge h(D_t, B_t) = \xi_t .$$

**9.28**

a) We have

$$\begin{aligned}d\eta^\varepsilon_t &= d\xi^\varepsilon_t - \gamma'_t\,dt = (b(\xi^\varepsilon_t) - b(\gamma_t))\,dt + \varepsilon\sigma(\xi^\varepsilon_t)\,dB_t\\ &= (b(\eta^\varepsilon_t + \gamma_t) - b(\gamma_t))\,dt + \varepsilon\sigma(\eta^\varepsilon_t + \gamma_t)\,dB_t .\end{aligned}$$

b) If $\sup_{0\le s\le T}\big|\varepsilon\int_0^s \sigma(\eta^\varepsilon_u + \gamma_u)\,dB_u\big| < \alpha$, we have, for $t \le T$,

$$|\eta^\varepsilon_t| \le \int_0^t |b(\eta^\varepsilon_s + \gamma_s) - b(\gamma_s)|\,ds + \alpha \le L\int_0^t |\eta^\varepsilon_s|\,ds + \alpha$$

and, by Gronwall's inequality, $\sup_{0\le t\le T}|\eta^\varepsilon_t| \le \alpha e^{LT}$. Taking the complement we have

$$P\Big(\sup_{0\le s\le T}|\xi^\varepsilon_s-\gamma_s|\ge\alpha\Big)=P\Big(\sup_{0\le s\le T}|\eta^\varepsilon_s|\ge\alpha\Big)$$

$$\le P\Big(\sup_{0\le s\le T}\Big|\int_0^s\sigma(\eta^\varepsilon_u+\gamma_u)\,dB_u\Big|\ge\frac{\alpha}{\varepsilon}\,e^{-LT}\Big)\le 2m\exp\Big[-\frac{1}{\varepsilon^2}\,\frac{\alpha^2e^{-2LT}}{2mT\|\sigma\|^2_\infty}\Big]$$

thanks to the exponential inequality (8.42). As $\eta^\varepsilon_s=\xi^\varepsilon_s-\gamma_s$ and by the arbitrariness of $\alpha$, this relation implies that $\xi^\varepsilon$, as $\varepsilon\to 0$, converges in probability to a constant r.v. taking the value $\gamma$ with probability 1.

c) Let us assume that Assumption (A') holds. For fixed $t$, let $R$ be the radius of some ball centered at the origin of $\mathbb{R}^m$ and containing $\gamma_t$ for every $t\le T$. Let $\widetilde{b},\widetilde{\sigma}$ be a vector and a matrix field, respectively, bounded and Lipschitz continuous and coinciding with $b,\sigma$, respectively, for $|x|\le R+1$.

Let us denote by $\widetilde{\xi}^\varepsilon$ the solution of an SDE with coefficients $\widetilde{b}$ and $\varepsilon\widetilde{\sigma}$ and with the same initial condition $\widetilde{\xi}^\varepsilon_0=x$. We know then (Theorem 9.3) that the two processes $\xi^\varepsilon$ and $\widetilde{\xi}^\varepsilon$ coincide up to the exit out of the ball of radius $R+1$. It follows that, for $\alpha\le 1$, the two events

$$\Big\{\sup_{0\le t\le T}|\xi^\varepsilon_t-\gamma_t|\ge\alpha\Big\}\quad\text{and}\quad\Big\{\sup_{0\le t\le T}|\widetilde{\xi}^\varepsilon_t-\gamma_t|\ge\alpha\Big\}$$

are a.s. equal. Therefore

$$P\Big(\sup_{0\le t\le T}|\xi^\varepsilon_t-\gamma_t|\ge\alpha\Big)=P\Big(\sup_{0\le t\le T}|\widetilde{\xi}^\varepsilon_t-\gamma_t|\ge\alpha\Big)\le 2m\exp\Big[-\frac{1}{\varepsilon^2}\,\frac{\alpha^2e^{-2\widetilde{L}T}}{2mT\|\widetilde{\sigma}\|^2_\infty}\Big].$$

**9.29**

a) Let $k^*=\sup_{0\le t\le T}\|G_tG_t^*\|$, where $\|\ \|$ denotes the norm as an operator, so that $\langle G_tG_t^*\theta,\theta\rangle\le k^*$ for every vector $\theta$ of modulus 1. By Proposition 8.7

$$P\Big(\sup_{0\le t\le T}\Big|\int_0^t G_s\,dB_s\Big|\ge\rho\Big)\le 2me^{-c_0\rho^2}\,,$$

where $c_0=(2Tmk^*)^{-1}$. If we define $A=\big\{\sup_{0\le t\le T}\big|\int_0^t G_s\,dB_s\big|<\rho\big\}$, then on $A$ we have, for $t\le T$,

$$|X_t|\le|x|+M\int_0^t(1+|X_s|)\,ds+\rho\,,$$

i.e.

$$X_t^* \le (|x| + MT + \rho) + M\int_0^t X_s^*\,ds\;,$$

and by Gronwall's Lemma $X_T^* \le (|x| + MT + \rho)\,\mathrm{e}^{MT}$. Therefore, if $|x| \le K$,

$$\mathrm{P}\big(X_T^* > (K + MT + \rho)\,\mathrm{e}^{MT}\big) \le \mathrm{P}\Big(\sup_{0\le t\le T}\Big|\int_0^t G_s\,dB_s\Big| \ge \rho\Big) \le 2m\mathrm{e}^{-c_0\rho^2}\;.$$

Solving $R = (K + MT + \rho)\,\mathrm{e}^{MT}$ we have $\rho = R\mathrm{e}^{-MT} - (|x| + MT)$ and therefore

$$\mathrm{P}(X_T^* > R) \le 2m\exp\Big(-c_0\big(R\mathrm{e}^{-MT} - (K + MT)\big)^2\Big)\;,$$

from which we obtain that, for every constant $c = c_T$ strictly smaller than

$$c_0\mathrm{e}^{-2MT} = \frac{\mathrm{e}^{-2MT}}{2Tmk^*}\;, \tag{S.75}$$

the inequality (9.70) holds for $R$ large enough.

b) It is an obvious consequence of a), with $F_t = b(\xi_t, t)$, $G_t = \sigma(\xi_t, t)$.

c) If $u(x) = \log(1 + |x|^2)$, let us compute with patience the derivatives:

$$\begin{aligned} u_{x_i}(x) &= \frac{2x_i}{1+|x|^2} \\ u_{x_ix_j}(x) &= \frac{2\delta_{ij}}{1+|x|^2} - \frac{4x_ix_j}{(1+|x|^2)^2}\;. \end{aligned} \tag{S.76}$$

In particular, as $|x| \to +\infty$ the first-order derivatives go to 0 at least as $x \mapsto |x|^{-1}$ and the second-order derivatives at least as $x \mapsto |x|^{-2}$. By Ito's formula the process $Y_t = \log(1 + |\xi_t|^2)$ has stochastic differential

$$\begin{aligned} dY_t = \Big(\sum_{i=1}^m u_{x_i}(\xi_t)b_i(\xi_t, t) + \sum_{i,j}^m u_{x_ix_j}(\xi_t)a_{ij}(\xi_t, t)\Big)\,dt \\ + \sum_{i=1}^m\sum_{j=1}^d u_{x_i}(\xi_t)\sigma_{ij}(\xi_t, t)\,dB_j(t) \end{aligned}$$

where $a = \sigma\sigma^*$. By (S.76), as $b$ and $\sigma$ are assumed to have sublinear growth, it is clear that all the terms $u_{x_i}b_i$, $u_{x_ix_j}a_{ij}$, $u_{x_i}\sigma_{ij}$ are bounded. We can therefore apply a), which guarantees that there exists a constant $c > 0$ such that, for large $\rho$,

$$\mathrm{P}\Big(\sup_{0\le t\le T}\log(1 + |\xi_t|^2) \ge \log\rho\Big) \le \mathrm{e}^{-c(\log\rho)^2}\;,$$

i.e. $\mathrm{P}\big(\xi_T^* \ge \sqrt{\rho - 1}\big) \le \mathrm{e}^{-c(\log\rho)^2}$. Letting $R = \sqrt{\rho - 1}$, i.e. $\rho = R^2 + 1$, the inequality becomes, for large $R$,

$$\mathrm{P}(\xi_T^* \ge R) \le \mathrm{e}^{-c(\log(R^2+1))^2} \le \mathrm{e}^{-c(\log R)^2} = \frac{1}{R^{c\log R}}\,\cdot \tag{S.77}$$

d) By Exercise 1.3, if $p \ge 1$,

$$\mathrm{E}[(\xi_T^*)^p] = p\int_0^{+\infty} t^{p-1}\mathrm{P}(\xi_T^* \ge t)\,dt\,.$$

Under Assumption (A'), by (S.77), the function $t \mapsto \mathrm{P}(\xi_T^* \ge t)$ tends to zero faster than every polynomial as $t \to +\infty$ and therefore the integral converges for every $p \ge 0$. If, moreover, the diffusion coefficient $\sigma$ is bounded, then

$$\mathrm{E}[\mathrm{e}^{\alpha(\xi_T^*)^2}] = 1 + \int_0^{+\infty} 2\alpha t\mathrm{e}^{\alpha t^2}\mathrm{P}(\xi_T^* \ge t)\,dt$$

and again, by (9.70) and (S.75), for $\alpha < \mathrm{e}^{-2MT}(2Tmk^*)^{-1}$ the integral is convergent. Here $k^*$ is any constant such that $|\sigma(x)\theta|^2 \le k^*$ for every $x \in \mathbb{R}^m$ and $|\theta| = 1$. A fortiori $\mathrm{E}[\mathrm{e}^{\alpha\xi_T^*}] < +\infty$ for every $\alpha \in \mathbb{R}$.

**10.1**

a) Just apply the representation formula (10.6), recalling that here $\phi \equiv 0$, $Z_t \equiv 1$ and $f \equiv -1$.

b) If $m = 1$ then (10.39) becomes

$$\begin{cases} \frac{1}{2}u'' = -1 & \text{on } ]-1,1[\\ u(-1) = u(1) = 0 \end{cases}$$

and has the solution $u(x) = 1 - x^2$. If $m \ge 2$, as indicated in the hint, if we look for a solution of the form $u(x) = g(|x|)$, we have

$$\frac{1}{2}\triangle u(x) = \frac{1}{2}g''(|x|) + \frac{m-1}{2|x|}g'(|x|)\,. \tag{S.78}$$

We therefore have to solve

$$\frac{1}{2}g''(y) + \frac{m-1}{2y}g'(y) = -1 \tag{S.79}$$

with the boundary condition $g(1) = 0$ (plus another condition that we shall see later). Letting $v = g'$, we are led to the equation

$$\frac{1}{2}v'(y) + \frac{m-1}{2y}v(y) = -1\,. \tag{S.80}$$

Separating the variables, the homogeneous equation is

$$\frac{v'(y)}{v(y)} = -\frac{m-1}{y}\,,$$

i.e. $v(y) = c_1 y^{-m+1}$. With the method of the variation of the constants, let us look for a solution of (S.80) of the form $c(y)y^{-m+1}$. We have immediately that $c$ must satisfy

$$c'(y)y^{-m+1} = -2$$

and therefore $c(y) = -\frac{2}{m}y^m$. In conclusion, the integral of (S.80) is $v(y) = c_1 y^{-m+1} - \frac{2}{m}y$. Let us compute the primitive of $v$. If $m \ge 3$ we find that the general integral of (S.79) is

$$g(y) = c_1 y^{-m+2} - \frac{1}{m}y^2 + c_2\,.$$

The condition $g(1) = 0$ gives $c_1 + c_2 = \frac{1}{m}$. However, we also need $g$ *to be defined at 0*, i.e. $c_1 = 0$. In conclusion, the solution of (10.39) is

$$u(x) = \frac{1}{m} - \frac{1}{m}|x|^2$$

and, in particular, $u(0) = \frac{1}{m}$. If $m = 2$ the computation of the primitive of $v$ is different and the general integral of (S.79) is

$$g(y) = c_1 \log y - \frac{1}{m}y^2 + c_2$$

but the remainder of the argument is the same.

**10.2**

a) Follows from Propositions 10.2 and 10.1 (Assumption $H_2$ is satisfied).
b) It is immediate that the constants and the function $v(x) = x^{1-2\delta}$ are solutions of the ordinary equation

$$\frac{1}{2}v''(x) + \frac{\delta}{x}v'(x) = 0$$

for every $x > 0$. Therefore, if we denote by $u(x)$ the term on the right-hand side in (10.40), as it is immediate that $u(a) = 0$ and $u(b) = 1$, $u$ is the solution of the problem

$$\begin{cases} \frac{1}{2}u''(x) + \frac{\delta}{x}\,u'(x) = 0 & x \in ]a, b[ \\ u(a) = 0, u(b) = 1\ . \end{cases}$$

Therefore, by (10.7), $u(x) = \mathrm{P}^x(X_\tau = b)$.

c) Thanks to Exercise 8.24, we know that the process $\xi_t = |B_t + x|$ is a solution of the SDE

$$\begin{aligned} d\xi_t &= \frac{m-1}{2\xi_t}\,dt + dW_t \\ \xi_0 &= |x| \end{aligned}$$

where $W$ is a real Brownian motion. Therefore, denoting by $\mathrm{Q}^{|x|}$ the law of $\xi$,

$$\mathrm{P}^x(|B_\sigma| = b) = \mathrm{Q}^{|x|}(X_\tau = b)\ .$$

Comparing with the result of a), we have $\delta = \frac{m-1}{2}$ hence $\lambda = 2\delta - 1 = m - 2$ which gives

$$\mathrm{P}^x(|B_\sigma| = b) = \frac{1 - (\frac{a}{x})^{m-2}}{1 - (\frac{a}{b})^{m-2}}\ \cdot$$

As $m \to \infty$ this probability converges to 1 for every starting point $x$, $a < |x| < b$.

**10.3**

a1) The solution is clearly $\xi^x = \sigma B_t - \mu t + x$ and we can apply the usual Iterated Logarithm argument.

a2) $Lu = \frac{\sigma^2}{2}\,u'' - \mu u'$. We know that the function $u(x) = \mathrm{P}(\xi^x_\tau = b)$ is the solution of

$$\begin{cases} Lu(x) = \dfrac{\sigma^2}{2}\,u'' - \mu u' = 0 \\ u(-a) = 0, u(b) = 1\ . \end{cases} \tag{S.81}$$

Setting $v = u'$ we have

$$v' = \frac{\mu}{\sigma^2}v\ ,$$

i.e.

$$v(x) = c\mathrm{e}^{\frac{2\mu}{\sigma^2}x} ,$$

and therefore

$$u(x) = c_1 \int_{x_0}^{x} \mathrm{e}^{\frac{2\mu}{\sigma^2}z}\, dz .$$

We must enforce the constraints $u(-a) = 0, u(b) = 1$, i.e. $x_0 = -a$ and

$$c_1 = \frac{1}{\int_{-a}^{b} \mathrm{e}^{\frac{2\mu}{\sigma^2}z}\, dz} = \frac{2\mu}{\sigma^2} \frac{1}{\mathrm{e}^{\frac{2\mu}{\sigma^2}b} - \mathrm{e}^{-\frac{2\mu}{\sigma^2}a}} \cdot$$

Therefore

$$\mathrm{P}(\xi_\tau^0 = b) = u(0) = c_1 \int_{-a}^{0} \mathrm{e}^{\frac{2\mu}{\sigma^2}z}\, dz = c_1 \frac{\sigma^2}{2\mu}\big(1 - \mathrm{e}^{-\frac{2\mu}{\sigma^2}a}\big) = \frac{1 - \mathrm{e}^{-\frac{2\mu}{\sigma^2}a}}{\mathrm{e}^{\frac{2\mu}{\sigma^2}b} - \mathrm{e}^{-\frac{2\mu}{\sigma^2}a}} \cdot \tag{S.82}$$

b1) The generator of $\eta$ is

$$L_2u(x) = \frac{\sigma^2}{2(1+x^2)}\, u''(x) - \frac{\mu}{1+x^2} u'(x) = \frac{1}{1+x^2}\, Lu(x) .$$

As the coefficient of the second-order derivative is $\geq$ const $> 0$ on the interval considered and the coefficient of the first-order derivative is bounded, we know, by Assumption $H_2$ p. 308, that $\tau < +\infty$ a.s. (and is even integrable) whatever the starting point $x$.

b2) We know that the function $u(x) = \mathrm{P}(\eta_\tau^x = b)$ is the solution of

$$\begin{cases} L_2u(x) = \dfrac{\sigma^2}{2(1+x^2)}\, u''(x) - \dfrac{\mu}{1+x^2} u'(x) = 0 \\ u(-a) = 0, u(b) = 1 \end{cases} \tag{S.83}$$

but, factoring out the denominator $1 + x^2$ we see that the solution is the same as that of (S.83), so that the exit probability is as in (S.82).

**10.4** The exit law from the unit ball for a Brownian motion with starting point at $x$ has density (see Example 10.1)

$$N(x, y) = \frac{1}{2\pi} \frac{1 - |x|^2}{|x - y|^2}$$

with respect to the normalized one-dimensional measure of the circle. In this case $|x| = \frac{1}{2}$ so that $1 - |x|^2 = \frac{3}{4}$ and, in angular coordinates, we are led to the computation of the integral

$$\frac{3}{4}\frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} \frac{1}{|x - y(\theta)|^2}\, d\theta \;,$$

where $y(\theta) = (\cos\theta, \sin\theta)$. Therefore $|x - y(\theta)|^2 = (\frac{1}{2} - \cos\theta)^2 + \sin^2\theta = \frac{5}{4} - \cos\theta$. The integral is computed with the change of variable

$$\cos\theta = \frac{1-t^2}{1+t^2}\;, \qquad t = \tan\tfrac{\theta}{2}\;, \qquad d\theta = \frac{2}{1+t^2}\, dt$$

and therefore

$$\begin{aligned} \cdots = \frac{3}{8\pi}\int_{-1}^{1} \frac{1}{\frac{5}{4} - \frac{1-t^2}{1+t^2}}\, \frac{2}{1+t^2}\, dt &= \frac{3}{\pi}\int_{-1}^{1} \frac{1}{1+9t^2}\, dt = \frac{1}{\pi}\arctan(3t)\Big|_{-1}^{1} \\ &= \frac{2}{\pi}\arctan 3 \simeq 0.795\;. \end{aligned}$$

**10.5**

a1) For $\varepsilon > 0$ fixed, let us still denote by $u$ a $C^2(\mathbb{R}^m)$ function that coincides with $u$ on $D_\varepsilon$. Then, by Ito's formula,

$$\begin{aligned} dM_t &= \theta e^{\theta t} u(X_t)\, dt + e^{\theta t} u'(X_t)\, dX_t + \frac{1}{2} e^{\theta t} \sum_{i,j=1}^{m} a_{ij}(X_t) \frac{\partial^2 u}{\partial x_i \partial x_j}(X_t)\, dt \\ &= e^{\theta t}\big(\theta u(X_t) + Lu(X_t)\big)\, dt + e^{\theta t} u'(X_t)\, \sigma(X_t)\, dB_t\;. \end{aligned}$$

As $M_0 = u(x)$ $\mathrm{P}^x$-a.s. and $\theta u(X_t) + Lu(X_t) = 0$ for $t \le \tau_\varepsilon$,

$$M_{t\wedge\tau_\varepsilon} = u(x) + \int_0^{t\wedge\tau_\varepsilon} e^{\theta s} u'(X_s)\, \sigma(X_s)\, dB_s, \qquad \mathrm{P}^x\text{-a.s.}$$

This is a square integrable martingale, as $u'$ and $\sigma$ are bounded on $D_\varepsilon$.

a2) By the relation

$$\mathrm{E}^x[e^{\theta(t\wedge\tau_\varepsilon)} u(X_{t\wedge\tau_\varepsilon})] = \mathrm{E}^x(M_{t\wedge\tau_\varepsilon}) = \mathrm{E}^x(M_0) = u(x)\;,$$

taking the limit first as $t \to +\infty$ and then as $\varepsilon \searrow 0$ and using Fatou's lemma twice (recall that we assume $u \ge 0$ and that $u \equiv 1$ on $\partial D$), we have

$$\mathrm{E}^x[e^{\theta\tau}] = \mathrm{E}^x[e^{\theta\tau} u(X_\tau)] \le u(x)\;.$$

Therefore the r.v. $e^{\theta\tau}$ is $P^x$-integrable for every $x \in D$. The r.v.'s $e^{\theta(t\wedge\tau_\varepsilon)}u(X_{t\wedge\tau_\varepsilon})$ are therefore positive and bounded above by $e^{\theta\tau}\max_{x\in\overline{D}} u(x)$. Again taking the limit as $t \to +\infty$ and then as $\varepsilon \searrow 0$ and using now Lebesgue's theorem we find $E^x[e^{\theta\tau}] = u(x)$.

b1) We must solve (10.41) for $L = \frac{1}{2}\triangle$. The general integral of $\frac{1}{2}u'' + \theta u = 0$ is

$$u(x) = c_1 \cos(\sqrt{2\theta}\, x) + c_2 \sin(\sqrt{2\theta}\, x) \ .$$

With the conditions $u(a) = u(-a) = 1$ we find $c_1 = \cos(\sqrt{2\theta}\, a)^{-1}$, $c_2 = 0$. Therefore the solution of (10.41) is

$$u(x) = \frac{\cos(\sqrt{2\theta}\, x)}{\cos(\sqrt{2\theta}\, a)} \cdot$$

As $\cos y > 0$ for $|y| < \frac{\pi}{2}$, a positive solution exists if and only if $\sqrt{2\theta}\, a < \frac{\pi}{2}$ i.e. $\theta < \frac{\pi^2}{8a^2}$. $\tau$ therefore has a finite Laplace transform for these values of $\theta$. Since $E^x(e^{\theta\tau}) \nearrow +\infty$ as $\theta \nearrow \frac{\pi^2}{8a^2}$, this is the convergence abscissa.

b2) We have seen, in particular, that $\tau$ *does not* have a finite Laplace transform for every $\theta \in \mathbb{R}$ and therefore cannot be bounded $P^x$-a.s. By Markov's inequality, for every $\beta > 0$,

$$P^x(\tau > R) \le P^x(e^{\beta\tau} > e^{\beta R}) \le e^{-\beta R} E^x(e^{\beta\tau}) \ .$$

This relation gives an estimate of the tail of the distribution of $\tau$ for every $\beta < \frac{\pi^2}{8a^2}$.

b3) The mean of $\tau$ is obtained as the derivative at 0 of the Laplace transform. As

$$\frac{d}{d\theta}\frac{\cos(\sqrt{2\theta}\, x)}{\cos(\sqrt{2\theta}\, a)}$$

$$= \frac{1}{\cos^2(\sqrt{2\theta}\,)a}\Big(-\cos(\sqrt{2\theta}\, a)\sin(\sqrt{2\theta}\, x)\,\frac{x}{\sqrt{2\theta}} + \sin(\sqrt{2\theta}\, a)\cos(\sqrt{2\theta}\, x)\,\frac{a}{\sqrt{2\theta}}\Big)$$

we have, for $\theta = 0$, $E^x(\tau) = a^2 - x^2$ (the mean of the exit time has already been computed in Exercises 5.10 and 10.1).

## 10.6

a) By Theorem 10.6 a solution is given by $u(x,t) = E^{x,t}[X_T^2]$ , where $P^{x,t}$ is the law of the diffusion associated to the differential generator $Lu = \frac{1}{2}u''$ with the initial conditions $x, t$. With respect to $P^{x,t}$, $X_T^2$ has the same law as $(B_{T-t} + x)^2$, where $(B_t)_t$ is a Brownian motion. Therefore

$$u(x,t) = E(B_{T-t}^2 + 2xB_{T-t} + x^2) = T - t + x^2 \ .$$

The solution is unique in the class of solutions with polynomial growth by Theorem 10.5.

b) If $\phi(x) = x^m$ then the unique solution with polynomial growth is

$$u(x,t) = \mathrm{E}[(B_{T-t} + x)^m] = \sum_{i=1}^{m} \binom{m}{k} x^k \mathrm{E}(B_{T-t}^{m-k}) \ .$$

As $\mathrm{E}(B_{T-t}^{m-k}) = 0$ if $m-k$ is odd, whereas $\mathrm{E}(B_{T-t}^{m-k}) = (T-t)^{\ell}\mathrm{E}(B_1^{2\ell})$ if $m-k = 2\ell$, $u$ is a polynomial in the variables $x, t$. Now just observe that the solution $u$ is linear in $\phi$.

## 10.7

a) Equation (10.43) can be written in the form

$$Lu + \frac{\partial u}{\partial t} = 0 \ ,$$

where $L$ is the generator of a geometric Brownian motion. Hence a candidate solution is

$$u(x,t) = \mathrm{E}[\xi_T^{x,t}] \ ,$$

where $\xi_s^{x,t} = x\mathrm{e}^{(b-\frac{1}{2}\sigma^2)(s-t)+\sigma(B_s-B_t)}$. Hence

$$u(x,t) = x\mathrm{e}^{b(T-t)} \ . \tag{S.84}$$

We cannot apply Theorem 10.6 because the generator is not elliptic, but it is easy to check directly that $u$ given by (S.84) is a solution of (10.43).

b) Following the same idea we surmise that a solution might be

$$u(x,t) = \mathrm{E}[(\xi_T^{x,t})^2] = \mathrm{E}[x^2\mathrm{e}^{2(b-\frac{1}{2}\sigma^2)(T-t)+2\sigma(B_T-B_t)}] = x^2\mathrm{e}^{(2b+\sigma^2)(T-t)} \ .$$

Again it is easy to check that such a $u$ is a solution of the given PDE problem.

## 10.8

a) Thanks to Theorem 10.6 a solution is given by

$$u(x,t) = \mathrm{E}^{x,t}[\cos(\langle\theta, X_T\rangle)] \ ,$$

where $(\mathscr{C}, \mathscr{M}, (\mathscr{M}_t)_t, (X_t)_t, (\mathrm{P}^{x,t})_{x,t})$ denotes a realization of the diffusion process associated to the generator $L$.

It is immediate that such a diffusion is an Ornstein–Uhlenbeck process and that, with respect to $\mathrm{P}^{x,t}$, $X_T$ has a Gaussian distribution with mean $\mathrm{e}^{-\lambda(T-t)}x$ and covariance matrix $\frac{\sigma^2}{2\lambda}\,(1-\mathrm{e}^{-2\lambda(T-t)})\,I$. Hence $u(x,t)$ is equal to the expectation of $\cos(\langle\theta,Z\rangle)$ where $Z$ is a Gaussian r.v. with these parameters. This is equal to the real part of the characteristic function

$$\mathrm{E}[\mathrm{e}^{i\langle\theta,Z\rangle}]=\mathrm{e}^{i\langle\theta,\mathrm{e}^{-\lambda(T-t)}x\rangle}\exp\Big(-\frac{|\theta|^2\sigma^2}{4\lambda}\,(1-\mathrm{e}^{-2\lambda(T-t)})\Big)\,,$$

i.e.

$$u(x,t)=\cos(\langle\theta,\mathrm{e}^{-\lambda(T-t)}x\rangle)\exp\Big(-\frac{|\theta|^2\sigma^2}{4\lambda}\,(1-\mathrm{e}^{-2\lambda(T-t)})\Big)\,.$$

b) As $T\to+\infty$ the value at time 0 of $u(x,t)$ converges, for every $x,t$, to the constant $\exp(-\frac{|\theta|^2\sigma^2}{4\lambda})$.

**10.9** The SDE associated to $L$ is

$$d\xi_t=b\xi_t\,dt+\sigma\xi_t\,dB_t\,, \tag{S.85}$$

where $\sigma=\sqrt{a}$. As Assumption (A) is satisfied, we know that the fundamental solution is given by the density of the transition function, if it exists. In order to show the existence of the transition function we cannot apply Theorem 10.7, whose hypotheses are not satisfied since the coefficients are not bounded and the diffusion coefficient is not elliptic. However, we know that the solution of (S.85) with the initial condition $\xi_0=x$ is

$$\xi_t^x=x\mathrm{e}^{(b-\frac{\sigma^2}{2})t+\sigma B_t}$$

(see Example 9.2). This is a time homogeneous diffusion and its transition function $p(t,x,\cdot)$ is the law of $\xi_t^x$.

The r.v. $\mathrm{e}^{(b-\frac{\sigma^2}{2})t+\sigma B_t}$ is lognormal with parameters $(b-\frac{\sigma^2}{2})t$ and $\sigma^2t$ (see Exercise 1.11) and therefore has density

$$g(y)=\frac{1}{\sqrt{2\pi t}\,\sigma y}\exp\Big(-\frac{1}{2\sigma^2t}\Big(\log y-(b-\tfrac{\sigma^2}{2})t\Big)^2\Big)\,.$$

Its transition function $p(t,x,\cdot)$ therefore has density

$$q(t,x,y)=\frac{1}{\sqrt{2\pi t}\,\sigma xy}\exp\Big(-\frac{1}{2\sigma^2t}\Big(\log\tfrac{y}{x}-(b-\tfrac{\sigma^2}{2})t\Big)^2\Big)$$

and the fundamental solution is $\Gamma(s,t,x,y)=q(t-s,x,y)$.

**10.10**

a) Let us apply Ito's formula to the function log, with the usual care, as it is not defined on the whole of $\mathbb{R}$. Let us denote by $\tau_\varepsilon$ the exit time of $\xi^{x,s}$ from the half line $]\varepsilon, +\infty[$; then, writing $\eta_t = \eta_t^{y,s}$, $y = \log x$ and $\xi_t = \xi_t^{x,s}$ for simplicity, we have

$$\begin{aligned}\eta_{t\wedge\tau_\varepsilon} &= y + \int_s^{t\wedge\tau_\varepsilon} \big(b(\xi_u, u) - \frac{1}{2}\sigma(\xi_u, u)^2\big)\, du + \int_s^{t\wedge\tau_\varepsilon} \sigma(\xi_u, u)\, dB_u \\ &= y + \int_s^{t\wedge\tau_\varepsilon} \big(b(\mathrm{e}^{\eta_u}, u) - \frac{1}{2}\sigma(\mathrm{e}_u^{\eta_u}, u)^2\big)\, du + \int_s^{t\wedge\tau_\varepsilon} \sigma(\mathrm{e}_u^{\eta_u}, u)\, dB_u\,.\end{aligned} \tag{S.86}$$

If $\widetilde{b}(y, u) = b(\mathrm{e}^y, u) - \frac{1}{2}\sigma(\mathrm{e}^y, u)^2$, $\widetilde{\sigma}(y, u) = \sigma(\mathrm{e}^y, u)$, the process $\eta$ coincides therefore up to time $\tau_\varepsilon$ with the solution $Y$ of the SDE

$$\begin{aligned} dY_t &= \widetilde{b}(Y_t, t)\, dt + \widetilde{\sigma}(Y_t, t)\, dB_t \\ Y_s &= y\,. \end{aligned} \tag{S.87}$$

As the coefficients $\widetilde{b}, \widetilde{\sigma}$ are bounded and locally Lipschitz continuous, by Theorem 9.2, the SDE (S.87) has a unique solution. Moreover, as $\tau_\varepsilon$ coincides with the exit time of $Y$ from the half line $]\log\varepsilon, +\infty[$, by Remark 9.3 $\tau_\varepsilon \to +\infty$ as $\varepsilon \to 0$ and, taking the limit as $\varepsilon \to 0$ in (S.86), we find that $\eta$ is a solution of (S.87) and that $\xi_t^{x,s} = \mathrm{e}^{\eta_t^{x,s}} > 0$ a.s. for every $t \geq 0$.

b) The generator $\widetilde{L}_t$ of the diffusion $\eta$ is

$$\widetilde{L}_t = \frac{1}{2}\widetilde{\sigma}(y, t)^2 \frac{\partial^2}{\partial y^2} + \widetilde{b}(y, t)\frac{d}{dy}\,\cdot$$

Moreover, if $\widetilde{\phi}(y) = \phi(\mathrm{e}^y)$, $\widetilde{f}(y, s) = f(\mathrm{e}^y, s)$ and $\widetilde{c}(y, s) = c(\mathrm{e}^y, s)$, we can write

$$u(\mathrm{e}^y, t) = \mathrm{E}\big[\widetilde{\phi}(\eta_T^{y,t})\, \mathrm{e}^{-\int_t^T \widetilde{c}(\eta_v^{y,t}, v)\, dv}\big] - \mathrm{E}\Big[\int_t^T \widetilde{f}(\eta_s^{y,t}, s)\, \mathrm{e}^{-\int_t^s \widetilde{c}(\eta_v^{y,t}, v)\, dv}\, ds\Big]\,.$$

As $\phi$ and $f$ have polynomial growth in $x$, $\widetilde{\phi}$ and $\widetilde{f}$ have exponential growth in $y$. The operator $\widetilde{L}_t$ satisfies Assumption (A') and has a bounded diffusion coefficient. By Theorem 10.6, the function $\widetilde{u}(y, t) = u(\mathrm{e}^y, t)$ is a solution of

$$\begin{cases} \widetilde{L}_t\widetilde{u} + \dfrac{\partial \widetilde{u}}{\partial t} - \widetilde{c}\,\widetilde{u} = \widetilde{f} & \text{on } \mathbb{R}\times[0, T[ \\ \widetilde{u}(y, T) = \widetilde{\phi}(y)\,. \end{cases}$$

We easily see that

$$\Big(\widetilde{L}_t + \frac{\partial}{\partial t} - \widetilde{c}\Big)\widetilde{u}(y,t) = \Big(L_t + \frac{\partial}{\partial t} - c\Big)u(\mathrm{e}^y, t)$$

and therefore $u$ is a solution of (10.45).

**10.11**

a) We have

$$u(x,t) = \frac{1}{(2\pi(T-t))^{m/2}} \int \phi(y)\,\mathrm{e}^{-\frac{1}{2(T-t)}|x-y|^2}\,dy$$

thus $u$ is $C^\infty(\mathbb{R}^m \times [0,T[)$ as we can take the derivative under the integral sign (this is a repetition of Remark 6.4 or of Proposition 10.4).

b) This is a consequence of the Markov property (6.13):

$$u(x,s) = \mathrm{E}^{x,s}[\mathrm{E}^{x,s}(\phi(X_T)\,|\,\mathscr{F}_t^s)] = \mathrm{E}^{x,s}[\mathrm{E}^{X_t,t}(\phi(X_T))] = \mathrm{E}^{x,s}[u(X_t,t)]\;.$$

c) Thanks to a), the function $x \mapsto u(x,T')$ is continuous for every $T' < T$. Moreover, $u$ is clearly bounded, as $\phi$ is bounded itself. By point b) and the Feynman–Kac formula (Theorem 10.6) $u$ is therefore a solution of (10.46) on $\mathbb{R}^m \times [0,T'[$ and therefore, by the arbitrariness of $T'$, on $\mathbb{R}^m \times [0,T[$.

In order to prove (10.47), let $x$ be a continuity point of $\phi$ and let $\varepsilon > 0$ be fixed. Let $\delta > 0$ be such that $|\phi(x)-\phi(y)| \le \varepsilon$ if $|x-y| \le \delta$. Let us first prove that there exists a number $\bar{t} > 0$ such that $\mathrm{P}^{x,t}(|X_T - x| > \delta) = \mathrm{P}^{x,0}(|X_{T-t} - x| > \delta) < \varepsilon$ for every $T - \bar{t} < t < T$. Actually with respect to $\mathrm{P}^{x,0}$, $(X_t - x)_t$ is a Brownian motion and therefore $X_{T-t} - x$ has the same law as $\sqrt{T-t}\,Z$ where $Z \sim N(0,I)$, so that

$$\mathrm{P}^{x,0}(|X_{T-t} - x| > \delta) = \mathrm{P}(|\sqrt{T-t}\,Z| > \delta) \underset{t\to T-}{\to} 0\;.$$

If $T - \bar{t} < t < T$ we then have

$$\begin{aligned} &|u(x,t) - \phi(x)| \le \mathrm{E}^{x,t}[|\phi(X_T) - \phi(x)|] \\ &= \mathrm{E}^{x,t}\big[|\phi(X_T) - \phi(x)|1_{\{|X_T - x| > \delta\}}\big] + \mathrm{E}^{x,t}\big[|\phi(X_T) - \phi(x)|1_{\{|X_T - x| \le \delta\}}\big] \\ &\le \|\phi\|_\infty \mathrm{P}^{x,t}(|X_T - x| > \delta) + \varepsilon < \varepsilon(\|\phi\|_\infty + 1) \end{aligned}$$

and we can conclude the proof thanks to the arbitrariness of $\varepsilon$.

d) By a) and c) for $\phi = 1_{[0,+\infty[}$,

$$u(x,t) = \mathrm{E}^{x,t}[1_{[0,+\infty[}(X_T)] = \frac{1}{(2\pi)^{1/2}} \int_{-x/\sqrt{T-t}}^{+\infty} \mathrm{e}^{-y^2/2}\,dy$$

is a solution of (10.46),(10.47); $x = 0$ is the only point of discontinuity of $\phi = 1_{[0,+\infty[}$ and it is immediate that $u(0,t) = \frac{1}{2}$ for every $t \le T$.

**10.12**

a) Let $\tau' = \inf\{t; t > 0, X_t \notin D\}$. We must prove that, if $\partial D$ has a local barrier for $L$ at $x$, then $\mathrm{P}^x(\tau' = 0) = 1$. Let us still denote by $w$ a bounded $C^2(\mathbb{R}^m)$ function coinciding with $w$ in a neighborhood of $x$, again denoted by $W$. Let $\sigma = \tau_W \wedge \tau'$, where $\tau_W$ is the exit time from $W$. As $w(x) = 0$, Ito's formula gives, for $t > 0$,

$$w(X_{t\wedge\sigma}) = \int_0^{t\wedge\sigma} Lw(X_s)\, ds + \int_0^{t\wedge\sigma} w'(X_s)\, dB_s \qquad \mathrm{P}^x\text{-a.s.}$$

The stochastic integral has mean equal to zero, since the gradient $w'$ is bounded in $W$. Therefore

$$\mathrm{E}^x[w(X_{t\wedge\sigma})] = \mathrm{E}^x\Big(\int_0^{t\wedge\sigma} Lw(X_s)\, ds\Big) \le -\mathrm{E}^x(t\wedge\sigma)\,.$$

As $w \ge 0$ on $W \cap \overline{D}$, the left-hand side is $\ge 0$ and necessarily $\mathrm{E}^x(t\wedge\sigma) = 0$ for every $t > 0$, i.e. $\sigma = 0$ $\mathrm{P}^x$-a.s. As $\tau_W > 0$ $\mathrm{P}^x$-a.s. by the continuity of the paths, necessarily $\tau' = 0$ $\mathrm{P}^x$-a.s. and therefore $x$ is regular for the diffusion $X$.

b) Let $z$ be the center of the sphere and let us look for a barrier of the form $w(y) = k(|x-z|^{-p} - |y-z|^{-p})$ where $k, p$ are numbers $> 0$ to be made precise later. First it is clear that $w(x) = 0$ and $w > 0$ on $D$, as $S \subset D^c$, and therefore, for $y \in D$, $|y-z|$ is larger than the radius of the sphere, which is equal to $|x-z|$. Let us compute the derivatives of $w$.

$$\begin{aligned}\frac{\partial w}{\partial y_i}(y) &= kp|y-z|^{-p-2}(y_i - z_i)\\ \frac{\partial^2 w}{\partial y_i \partial y_i}(y) &= -kp(p+2)|y-z|^{-p-4}(y_i-z_i)(y_j-z_j) + kp|y-z|^{-p-2}\delta_{ij}\,.\end{aligned}$$

Therefore

$$\begin{aligned}Lw(y) =& -kp(p+2)|y-z|^{-p-4}\sum_{i,j=1}^m a_{ij}(y)(y_i-z_i)(y_j-z_j)\\ &+ kp|y-z|^{-p-2}\sum_{i=1}^m a_{ii}(y) + kp|y-z|^{-p-2}\sum_{i=1}^m b_i(y)(y_i-z_i)\,.\end{aligned}$$

Let now $\lambda$ be a positive number such that $\langle a(y)\xi,\xi\rangle \ge \lambda|\xi|^2$ for every $y \in D$ and $\xi \in \mathbb{R}^m$ and $M$ a number majorizing the norm of $b(y)$ and the trace of $a(y)$ for every $y \in W$. Then

$$\begin{aligned} Lw(y) &\le -kp(p+2)\lambda|y-z|^{-p-2} + kpM|y-z|^{-p-2} + kpM|y-z|^{-p-1} \\ &= -kp|y-z|^{-p-2}(\lambda(p+2) - M - M|y-z|)\ . \end{aligned}$$

If $W$ is a bounded neighborhood of $x$, then the quantity $|y-z|$ remains bounded for $y \in W$. One can therefore choose $p$ large enough so that $\lambda(p+2)-M-M|y-z| > 1$ on $W$. With this choice we have, for $y \in W \cap D$,

$$Lw(y) \le -kp|y-z|^{-p-2} \le -kp|x-z|^{-p-2}$$

and, choosing $k$ large enough, we have $Lw(y) \le -1$, as requested.

**11.1**

a1) We have

$$\overline{\xi}_{t_{k+1}} = \overline{\xi}_{t_k} + b\overline{\xi}_{t_k}h + \sigma\overline{\xi}_{t_k}\sqrt{h}\,Z_k = \overline{\xi}_{t_k}(1 + bh + \sigma\sqrt{h}\,Z_k)$$

from which (11.44) is immediate.

a2) Recall that for the geometric Brownian motion of (11.43) $\mathrm{E}[\xi_T] = \mathrm{e}^{bT}$ and $\mathrm{E}[\xi_T^2] = \mathrm{e}^{(2b+\sigma^2)T}$. We have

$$\mathrm{E}[\overline{\xi}_T] = x\prod_{i=1}^n \mathrm{E}[1 + bh + \sigma\sqrt{h}\,Z_k] = x(1+bh)^n = x\Big(1 + \frac{bT}{n}\Big)^n \underset{n\to\infty}{\to} x\mathrm{e}^{bT}$$

and

$$\mathrm{E}[\overline{\xi}_T^2] = x\prod_{i=1}^n \mathrm{E}[(1 + bh + \sigma\sqrt{h}\,Z_k)^2]$$

$$= x\prod_{i=1}^n \mathrm{E}\big[1 + b^2h^2 + \sigma^2hZ_k^2 + 2bh + 2\sigma\sqrt{h}\,Z_k + 2bh\sigma\sqrt{h}\,Z_k\big]$$

$$= x\prod_{i=1}^n \big(1+(2b+\sigma^2)h+b^2h^2\big) = x\Big(1+(2b+\sigma^2)\frac{T}{N}+b\frac{T^2}{N^2}\Big)^n \underset{n\to\infty}{\to} \mathrm{e}^{(2b+\sigma^2)T}\ .$$

a3) We have

$$x(1+bh)^n = x\mathrm{e}^{n\log(1+bh)}\ .$$

Recall the developments

$$\log(1+z) = z - \frac{1}{2}z^2 + O(z^3)$$
$$e^y = e^{y_0} + e^{y_0}(y - y_0) + O(|y - y_0|^2)$$

that give

$$x e^{n\log(1+bh)} = x e^{bT - \frac{1}{2}b^2 Th + O(h^2)} = x e^{bT} - \frac{1}{2} x e^{bT} b^2 Th + O(h^2)\ .$$

Hence

$$\big|E[\xi_T] - E[\overline{\xi}_T]\big| = c_1 h + O(h^2)\ ,$$

with $c_1 = \frac{1}{2} x e^{bT} b^2 T$. The computation for (11.46) is quite similar.

b1) We have

$$\begin{aligned}\widetilde{\xi}_{t_{k+1}} &= \widetilde{\xi}_{t_k} + b\widetilde{\xi}_{t_k} h + \sigma\widetilde{\xi}_{t_k}\sqrt{h}\, Z_{k+1} + \frac{1}{2}\,\sigma^2\widetilde{\xi}_{t_k}(hZ_{k+1}^2 - h)\\ &= \widetilde{\xi}_{t_k}\Big(1 + bh + \sigma\sqrt{h}\, Z_k + \frac{1}{2}\,\sigma^2 h(Z_{k+1}^2 - 1)\Big)\ ,\end{aligned}$$

hence

$$\widetilde{\xi}_T = x\prod_{k=0}^{n-1}\Big(1 + h\Big(b + \frac{1}{2}\,\sigma^2(Z_{k+1}^2 - 1)\Big) + \sigma\sqrt{h}\, Z_{k+1}\Big)\ .$$

b2) We have

$$E[\widetilde{\xi}_T] = x\prod_{k=0}^{n-1} E\Big[1 + h\Big(b + \frac{1}{2}\,\sigma^2(Z_{k+1}^2 - 1)\Big) + \sigma\sqrt{h}\, Z_{k+1}\Big] = x(1+bh)^n\ .$$

The mean of the Milstein approximation in this case is exactly the same as that of the Euler approximation. Hence (11.47) follows by the same computation leading to (11.45).

**12.1**

a) By Girsanov's theorem, if

$$\widetilde{Z}_t = e^{cB_t - \frac{1}{2}c^2 t}\ ,$$

then $W_s = B_s - cs$ is a Brownian motion for $s \le t$ with respect to the probability $dQ = \widetilde{Z}_t\, dP$. Therefore, with respect to Q, writing $B_s = W_s + cs$, $(B_s)_{0\le s\le t}$ has the same law as $(Y_s)_{0\le s\le t}$. We deduce that $P^Y$ is the law of $(B_t)_t$ with respect to Q. If $0 \le t_1 < \dots < t_m \le t$ and $f_1, \dots, f_m$ are bounded Borel functions, we have

$$E^{P^Y}[f_1(X_{t_1})\dots f_m(X_{t_m})] = E^Q[f_1(B_{t_1})\dots f_m(B_{t_m})]$$
$$= E[f_1(B_{t_1})\dots f_m(B_{t_m})\, e^{cB_t-\frac{1}{2}c^2t}] = E^{P^B}[f_1(X_{t_1})\dots f_m(X_{t_m})\, e^{cX_t-\frac{1}{2}c^2t}]\,.$$

Therefore the two probabilities $P^Y$ and $e^{cX_t-\frac{1}{2}c^2t}dP^B$ have the same finite-dimensional distributions on $\mathscr{M}_t$ and hence coincide on $\mathscr{M}_t$. This proves that $P^Y$ has a density with respect to $P^B$ so that $P^Y \ll P^B$. Let $Z_t = e^{cX_t-\frac{1}{2}c^2t}$. As $Z_t > 0$, we also have $P^B \ll P^Y$ on $\mathscr{M}_t$.

If $c > 0$, for instance, we have $\lim_{t\to+\infty} Y_t = +\infty$ hence with respect to $P^Y$ $\lim_{t\to+\infty} X_t = +\infty$ a.s. Therefore the event $\{\lim_{t\to+\infty} X_t = +\infty\}$ has probability 1 with respect to $P^Y$ but probability 0 with respect to $P^B$, as under $P^B$ $(X_t)_t$ is a Brownian motion and therefore $\underline{\lim}_{t\to+\infty} X_t = -\infty$ a.s. by the Iterated Logarithm Law.

b) As in the argument developed in the second part of a) we must find an event in $\mathscr{M}_t$ which has probability 1 for $P^B$ and 0 for $P^Z$. For instance, by the Iterated Logarithm Law,

$$P^B\Big(\overline{\lim_{t\to0+}} \frac{X_t}{\big(2t\log\log\frac{1}{t}\big)^{1/2}} = 1\Big) = 1$$

whereas, as $(X_t)_t$ under $P^Z$ has the same law as $(\sigma B_t)_t$, we have, considering separately the cases $\sigma > 0$ and $\sigma < 0$,

$$\overline{\lim_{t\to0+}} \frac{X_t}{\big(2t\log\log\frac{1}{t}\big)^{1/2}} = |\sigma| \qquad P^Z\text{–a.s.}$$

hence, if $|\sigma| \ne 1$, the event

$$\Big\{\overline{\lim_{t\to0+}} \frac{X_t}{\big(2t\log\log\frac{1}{t}\big)^{1/2}} = 1\Big\}$$

has probability 1 for $P^B$ and probability 0 for $P^Z$.

Another event having probability 1 for $P^B$ and probability 0 for $P^Z$ can be easily produced using P. Lévy's modulus of continuity (Theorem 3.1).

**12.2**

a) Of course it is sufficient to consider the case $\nu \ll \mu$. As $f(x) = x\log x$, with the understanding $f(0) = 0$, is convex and lower semi-continuous, by Jensen's inequality,

$$H(\nu;\mu) = \int_E \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu}\, d\mu = \int_E f\Big(\frac{d\nu}{d\mu}\Big)\, d\mu \ge f\Big(\int_E \frac{d\nu}{d\mu}\, d\mu\Big)$$
$$= f\Big(\int_E d\nu\Big) = f(1) = 0\,.$$

As, moreover, $f$ is strictly convex, the inequality is strict, unless the function $\frac{d\nu}{d\mu}$ is $\mu$-a.s. constant. As the integral of $\frac{d\nu}{d\mu}$ with respect to $\mu$ must be equal to 1, this happens if and only if $\frac{d\nu}{d\mu} = 1$ $\mu$-a.s. and therefore if and only if $\nu = \mu$.

A simple approach to finding an example showing that the entropy is not symmetric in its two arguments is to consider a measure $\mu$ such that there exists a set $A$ with $0 < \mu(A) < 1$ and then setting $d\nu = \mu(A)^{-1} 1_A\, d\mu$. The measure $\mu$ is not absolutely continuous with respect to $\nu$, as $\nu(A^c) = 0$ whereas $\mu(A^c) > 0$; therefore $H(\mu;\nu) = +\infty$. Conversely, $\nu \ll \mu$ and one computes immediately $H(\nu;\mu) = -\log \mu(A) < +\infty$.

b1) The process $(X_t)_{0\le t\le T}$, under the Wiener measure P, is a Brownian motion. If

$$Z = \exp\Big(\int_0^T \gamma'_s\, dX_s - \frac{1}{2}\int_0^T |\gamma'_s|^2\, ds\Big)$$

then, as the r.v. $\int_0^T \gamma'_s\, dX_s$ is Gaussian, we know that $\mathrm{E}[Z] = 1$ and by Girsanov's Theorem 12.1 under the probability $d\mathrm{Q} = Z\, d\mathrm{P}$, the process $W_t = X_t - \gamma_t$ is a Brownian motion for $0 \le t \le T$. Hence $X_t = W_t + \gamma_t$ and, with respect to Q, $X$ is, up to time $T$, a Brownian motion with the deterministic drift $\gamma$. Hence $\mathrm{Q} = \mathrm{P}_1$ and $\mathrm{P}_1$ is absolutely continuous on $\mathscr{M}_T$ with respect to the Wiener measure P and

$$\frac{d\mathrm{P}_1}{d\mathrm{P}} = \exp\Big(\int_0^T \gamma'_s\, dX_s - \frac{1}{2}\int_0^T |\gamma'_s|^2\, ds\Big)\,.$$

Using the expression (12.20) the entropy $H(\mathrm{P}_1;\mathrm{P})$ is the mean with respect to $\mathrm{P}_1$ of the logarithm of this density. But, as with respect to $\mathrm{P}_1$ for $t \le T$ $W_t = X_t - \gamma_t$ is a Brownian motion, we have,

$$H(\mathrm{P}_1;\mathrm{P}) = \mathrm{E}^{\mathrm{P}_1}\Big(\log \frac{d\mathrm{P}_1}{d\mathrm{P}}\Big) = \mathrm{E}^{\mathrm{P}_1}\Big(\int_0^T \gamma'_s\, dX_s - \frac{1}{2}\int_0^T |\gamma'_s|^2\, ds\Big)$$
$$\mathrm{E}^{\mathrm{P}_1}\Big(\int_0^T \gamma'_s\, dW_s + \int_0^T |\gamma'_s|^2\, ds - \frac{1}{2}\int_0^T |\gamma'_s|^2\, ds\Big) = \frac{1}{2}\int_0^T |\gamma'_s|^2\, ds = \frac{1}{2}\|\gamma'\|_2^2\,.$$

Very similar to this is the computation of the entropy $H(\mathrm{P};\mathrm{P}_1)$. We have, clearly,

$$\frac{d\mathrm{P}}{d\mathrm{P}_1} = Z^{-1} = \exp\Big(-\int_0^T \gamma'_s\, dX_s + \frac{1}{2}\int_0^T |\gamma'_s|^2\, ds\Big)$$

and, as with respect to P $(X_t)_{0\le t\le T}$ is a Brownian motion,

$$H(\mathrm{P};\mathrm{P}_1)=\mathrm{E}^{\mathrm{P}}\Big(-\int_0^T\gamma'_s\,dX_s+\frac12\int_0^T|\gamma'_s|^2\,ds\Big)=\frac12\int_0^T|\gamma'_s|^2\,ds=\frac12\|\gamma'\|_2^2\,.$$

Here the two entropies $H(\mathrm{P};\mathrm{P}_1)$ and $H(\mathrm{P}_1;\mathrm{P})$ coincide.

b2) Developing the square we have in general, if $\nu\ll\mu$,

$$\chi^2(\nu;\mu)=\int_E\Big[\Big(\frac{d\nu}{d\mu}\Big)^2-2\frac{d\nu}{d\mu}+1\Big]\,d\mu=\int_E\Big(\frac{d\nu}{d\mu}\Big)^2\,d\mu-1=\int_E\frac{d\nu}{d\mu}\,d\nu-1\,.$$

In our case, i.e. $\nu=\mathrm{P}$ and $\mu=\mathrm{P}_1$,

$$\begin{aligned}\int\frac{d\mathrm{P}}{d\mathrm{P}_1}\,d\mathrm{P}&=\mathrm{E}^{\mathrm{P}}\Big[\exp\Big(-\int_0^T\gamma'_s\,dX_s+\frac12\int_0^T|\gamma'_s|^2\,ds\Big)\Big]\\&=\mathrm{E}^{\mathrm{P}}\Big[\exp\Big(-\int_0^T\gamma'_s\,dX_s\Big)\Big]\exp\Big(\frac12\int_0^T|\gamma'_s|^2\,ds\Big)\\&=\exp\Big(\int_0^T|\gamma'_s|^2\,ds\Big)=\mathrm{e}^{\|\gamma'\|_2^2}\,.\end{aligned}$$

Therefore

$$\chi^2(\mathrm{P}_1;\mathrm{P})=\int\frac{d\mathrm{P}}{d\mathrm{P}_1}\,d\mathrm{P}-1=\mathrm{e}^{\|\gamma'\|_2^2}-1\,.$$

$\chi^2(\mathrm{P};\mathrm{P}_1)$ can be computed similarly, giving the same result.

**12.3**

a) $(Z_t)_t$ is a martingale and an old acquaintance (see Example 5.2). $\widetilde B$ is a Brownian motion with respect to Q by Girsanov's Theorem 12.1 (here $\Phi_s\equiv 2\theta$).

b) We have

$$Z_t^{-1}=\mathrm{e}^{-2\theta B_t+2\theta^2t}=\mathrm{e}^{-2\theta\widetilde B_t-2\theta^2t}\,.$$

Again $(Z_t^{-1})_t$ is a Q-martingale as $\widetilde B$ is a Q-Brownian motion. Alternatively, simply observe that on $\mathscr{F}_t$ we have $d\mathrm{P}=Z_t^{-1}\,d\mathrm{Q}$, and therefore necessarily $(Z_t^{-1})_t$ is a Q-martingale (Exercise 5.27). Note also that $Z_t^{-1}=\mathrm{e}^{-2\theta B_t+2\theta^2t}=\mathrm{e}^{-2\theta X_t}$.

c) As $\{\tau_R\le T\}\in\mathscr{F}_T$,

$$\mathrm{P}(\tau_R\le T)=\mathrm{E}^{\mathrm{Q}}(1_{\{\tau_R\le T\}}Z_T^{-1})\,,$$

but also $\{\tau_R \le T\} \in \mathscr{F}_{\tau_R}$ and therefore $\{\tau_R \le T\} \in \mathscr{F}_{T\wedge\tau_R}$. Conditioning with respect to $\mathscr{F}_{T\wedge\tau_R}$, by the stopping theorem of martingales (Theorem 5.13),

$$\begin{aligned} \mathrm{E}^{\mathrm{Q}}(1_{\{\tau_R\le T\}}Z_T^{-1}) = \mathrm{E}^{\mathrm{Q}}\big[\mathrm{E}^{\mathrm{Q}}(1_{\{\tau_R\le T\}}Z_T^{-1}\,|\,\mathscr{F}_{T\wedge\tau_R})\big] &= \mathrm{E}^{\mathrm{Q}}\big[1_{\{\tau_R\le T\}}\mathrm{E}^{\mathrm{Q}}(Z_T^{-1}\,|\,\mathscr{F}_{T\wedge\tau_R})\big] \\ &= \mathrm{E}^{\mathrm{Q}}(1_{\{\tau_R\le T\}}Z_{T\wedge\tau_R}^{-1})\ . \end{aligned}$$

As $Z_{T\wedge\tau_R}^{-1} = \mathrm{e}^{-2\theta X_{T\wedge\tau_R}}$ and $X_{T\wedge\tau_R} = R$ on $\{\tau_R \le T\}$, we have $Z_{T\wedge\tau_R}^{-1}1_{\{\tau_R\le T\}} = \mathrm{e}^{-2\theta R}1_{\{\tau_R\le T\}}$ hence (12.21).

d) As, with respect to Q, $X_t = \widetilde{B}_t + \theta t$, we have $\lim_{t\to+\infty} X_t = +\infty$ Q-a.s. and $\mathrm{Q}(\tau_R < +\infty) = 1$. Taking the limit as $T \to +\infty$ in (12.21) we have

$$\mathrm{P}\Big(\sup_{t>0} X_t \ge R\Big) = \mathrm{P}(\tau_R < +\infty) = \mathrm{e}^{-2\theta R}$$

so that the r.v. $\sup_{t>0} X_t$ has an exponential law with parameter $2\theta$.

- The observant reader has certainly noticed in the last limit a certain carelessness (the probability Q itself actually depends on $T$). However, it is not difficult to complete the argument with care.

**12.4** By Girsanov's theorem, the process

$$W_t = B_t + \int_0^t \frac{B_s}{1-s}\,ds,$$

$t \le T$, is a Brownian motion with respect to Q. Hence, with respect to Q, $B$ is a solution of the SDE

$$dB_t = -\frac{B_t}{1-t}\,dt + dW_t\ .$$

As clearly $B_0 = 0$ Q-a.s., we know (Exercise 9.2 b)) that the solution of this SDE is a Brownian bridge.

**12.5**

a) In order to show that $\mathrm{E}(Z_t) = 1$, by Corollary 12.1 b) it is enough to prove that $\mathrm{E}(\mathrm{e}^{\mu\theta^2|X_s|^2}) < +\infty$ for some $\mu > 0$ for every $s \le t$. If $X_s = (X_1(s),\dots,X_m(s))$, then, as the components of $X$ are independent,

$$\mathrm{E}(\mathrm{e}^{\mu\theta^2|X_s|^2}) = \mathrm{E}(\mathrm{e}^{\mu\theta^2X_1(s)^2})\dots\mathrm{E}(\mathrm{e}^{\mu\theta^2X_m(s)^2}) = \mathrm{E}(\mathrm{e}^{\mu\theta^2X_1(s)^2})^m = \mathrm{E}(\mathrm{e}^{\mu s\theta^2W^2})^m\ ,$$

where $W \sim N(0, 1)$. This quantity, thanks to Exercise 1.12, is finite for every $s \le t$ if $\mu < (2\theta^2 t)^{-1}$. Actually this is a repetition of the argument of Example 12.2.

b) By Girsanov's Theorem 12.1, with respect to Q the process

$$W_s = X_s - \theta \int_0^s X_u \, du$$

is an ($m$-dimensional) Brownian motion for $s \le t$. Therefore, for $s \le t$, $X$ is the solution of

$$dX_s = \theta X_s \, ds + dW_s$$

and is therefore, under Q, an Ornstein–Uhlenbeck process. As $X_0 = W_0 = 0$, as seen in Sect. 9.2 we have

$$X_t = \int_0^t e^{\theta(t-s)} \, dW_s$$

and $X$ is a Gaussian process; $E^Q(X_t) = 0$ and, by Proposition 8.5, the covariance matrix of $X_t$ with respect to Q is equal to the identity matrix multiplied by

$$\int_0^t e^{2\theta(t-s)} \, ds = \frac{1}{2\theta}(e^{2\theta t} - 1) \, . \tag{S.88}$$

c) By Ito's formula, with respect to P we have $d|X_t|^2 = 2X_t \, dX_t + m \, dt$, i.e.

$$\int_0^t X_s \, dX_s = \frac{1}{2} \, (|X_t|^2 - mt) \, .$$

Therefore, for $\lambda = -\frac{\theta^2}{2}$,

$$\begin{aligned} E^Q[e^{-\frac{\theta}{2}(|X_t|^2 - mt)}] &= E[Z_t e^{-\frac{\theta}{2}(|X_t|^2 - mt)}] \\ = E\Big[\exp\Big(\theta \int_0^t X_s \, dX_s - \frac{\theta^2}{2} \int_0^t |X_s|^2 \, ds - \frac{\theta}{2}(|X_t|^2 - mt)\Big)\Big] \\ = E\Big[\exp\Big(-\frac{\theta^2}{2} \int_0^t |X_s|^2 \, ds\Big)\Big] = J_\lambda \, . \end{aligned}$$

As the components of $X$ are independent with respect to Q, thanks to b) we have

$$J_\lambda = E^Q[e^{-\frac{\theta}{2}(|X_t|^2 - mt)}] = e^{\frac{1}{2}m\theta t} E^Q[e^{-\frac{\theta}{2}X_1(t)^2}]^m \, .$$

With respect to Q, $X_1(t)$ is a centered Gaussian r.v. with a variance given by (S.88); therefore, recalling Exercise 1.12,

$$\mathrm{E}^{\mathrm{Q}}[\mathrm{e}^{-\frac{\theta}{2}X_1(t)^2}] = \big(1+\theta\mathrm{Var}_{\mathrm{Q}}(X_1(t))\big)^{-1/2} = \Big(1+\frac{1}{2}\,(\mathrm{e}^{2\theta t}-1)\Big)^{-1/2}$$
$$= \Big(\frac{1}{2}\,(\mathrm{e}^{2\theta t}+1)\Big)^{-1/2}\,.$$

Putting all pieces together we obtain

$$J_\lambda = \mathrm{e}^{\frac{1}{2}m\theta t}\Big(\frac{1}{2}\,(\mathrm{e}^{2\theta t}+1)\Big)^{-m/2} = \Big(\frac{\mathrm{e}^{2\theta t}+1}{2\mathrm{e}^{\theta t}}\Big)^{-m/2}$$
$$= \cosh(\theta t)^{-m/2} = \cosh(\sqrt{-2\lambda}\,t)^{-m/2}\,.$$

d) Let us denote by $\psi$ the Laplace transform of the r.v. $\int_0^t |X_s|^2\,ds$ so that $J_\lambda = \psi(\lambda)$. We have seen that, for $\lambda \le 0$, $\psi(\lambda) = \cosh(\sqrt{-2\lambda}\,t)^{-m/2}$. Recalling that the Laplace transform is an analytic function (see Sect. 5.7), for $\lambda \ge 0$ we have $\psi(\lambda) = \cosh(i\sqrt{2\lambda}\,t)^{-m/2} = \cos(\sqrt{2\lambda}\,t)^{-m/2}$, up to the convergence abscissa. Keeping in mind that the first positive zero of the cosine function is $\frac{\pi}{2}$, the convergence abscissa is $\lambda = \frac{\pi^2}{8t^2}$ and in conclusion

$$J_\lambda = \begin{cases} \cosh(\sqrt{-2\lambda}\,t)^{-m/2} & \text{if } \lambda \le 0 \\ \cos\,(\sqrt{2\lambda}\,t)^{-m/2} & \text{if } 0 < \lambda < \frac{\pi^2}{8t^2} \\ +\infty & \text{if } \lambda \ge \frac{\pi^2}{8t^2}\,. \end{cases}$$

**12.6**

a1) By Girsanov's theorem, $W_t = X_t - \int_0^t b(X_s + x)\,ds$ is a Brownian motion for $t \le T$ with respect to the probability Q defined by $d\mathrm{Q} = Z_T\,d\mathrm{P}$, where

$$Z_t = \exp\Big(\int_0^t b(X_s + x)\,dX_s - \frac{1}{2}\int_0^t b^2(X_s + x)\,ds\Big)\,. \tag{S.89}$$

Such a process $(Z_t)_t$ is a martingale thanks to Proposition 12.1, $b$ being bounded.

With respect to Q, $(X_t)_{0\le t\le T}$ satisfies

$$X_t = \int_0^t b(X_s + x)\,ds + W_t\,,$$

hence $Y_t = X_t + x$ is a solution of

$$Y_t = x + \int_0^T b(X_s + x)\,ds + W_t = x + \int_0^T b(Y_s)\,ds + W_t\;.$$

a2) If $U$ is a primitive of $b$, then by Ito's formula

$$dU(X_t + x) = b(X_t + x)\,dX_t + \frac{1}{2}b'(X_t + x)\,dt$$

so that

$$\int_0^t b(X_s + x)\,dX_s = U(X_t + x) - U(x) - \frac{1}{2}\int_0^t b'(X_s + x)\,ds\;,$$

whence substituting into (S.89) we find that $Z_t$ is given by (12.22).

b1) We have

$$\begin{aligned}\tanh^2(z) &= \frac{\sinh^2(z)}{\cosh^2(z)} = 1 - \frac{1}{\cosh^2(z)}\\ \tanh'(z) &= \frac{1}{\cosh^2(z)}\;,\end{aligned}$$

hence

$$b'(z) = \frac{k^2}{\cosh^2(kz + c)}\;, \qquad b^2(z) = k^2\Big(1 - \frac{1}{\cosh^2(kz + c)}\Big)\;,$$

so that $b'(z) + b^2(z) \equiv k^2$.

b2) Let $t \le T$. The law of $Y_t$ coincides with the law of $X_t + x$ with respect to the probability Q defined as $d\mathrm{Q} = Z_T\,d\mathrm{P}$, where $Z_T$ is as in (12.22) with $b(z) = k\tanh(kz + c)$. Hence the Laplace transform of $Y_t$ is given by

$$\mathrm{E}[\mathrm{e}^{\theta Y_t}] = \mathrm{E}^{\mathrm{Q}}[\mathrm{e}^{\theta(X_t + x)}] = \mathrm{E}[Z_T\mathrm{e}^{\theta(X_t + x)}] = \mathrm{E}[\mathrm{E}[Z_T\mathrm{e}^{\theta(X_t + x)}\,|\,\mathscr{F}_t]] = \mathrm{E}[Z_t\mathrm{e}^{\theta(X_t + x)}]\;.$$

A primitive of $z \mapsto k\tanh(kz + c)$ is $U(z) = \log\cosh(kz + c)$. Therefore

$$\begin{aligned}&U(X_t + x) - U(x) - \frac{1}{2}\int_0^t [b'(X_s + x) + b^2(X_s + x)]\,ds\\ &= \log\cosh(kX_t + kx + c) - \log\cosh(kx + c) - \frac{1}{2}\,k^2t\end{aligned}$$

and

$$Z_t = \frac{\cosh(kX_t + kx + c)}{\cosh(kx + c)}\, e^{-\frac{1}{2}k^2 t}\,.$$

Hence, as $X$ is a Brownian motion,

$$\begin{aligned}
E[Z_t e^{\theta(X_t+x)}] &= \frac{e^{-\frac{1}{2}k^2 t}}{\cosh(kx+c)}\, E\big[\cosh(kX_t + kx + c)\, e^{\theta(X_t+x)}\big]\\
&= \frac{e^{-\frac{1}{2}k^2 t} e^{\theta x}}{2\cosh(kx+c)}\, E\big[e^{(\theta+k)X_t + kx + c} + e^{(\theta-k)X_t - kx - c}\big]\\
&= \frac{e^{-\frac{1}{2}k^2 t} e^{\theta x}}{2\cosh(kx+c)}\big(e^{\frac{1}{2}(\theta+k)^2 t} e^{kx+c} + e^{\frac{1}{2}(\theta-k)^2 t} e^{-kx-c}\big)\\
&\frac{1}{2\cosh(kx+c)}\big(\underbrace{e^{\frac{1}{2}\theta^2 t} e^{\theta(kt+x)}}_{\widehat{\mu}_1(\theta)}\, e^{kx+c} + \underbrace{e^{\frac{1}{2}\theta^2 t} e^{\theta(-kt+x)}}_{\widehat{\mu}_2(\theta)}\, e^{-kx-c}\big)\,,
\end{aligned}$$

where we recognize that $\widehat{\mu}_1$ and $\widehat{\mu}_2$ are respectively the Laplace transform of an $N(kt+x, t)$ and of an $N(-kt+x, t)$ law. The law of $Y_t$ is therefore a mixture of these two laws with weights (independent of $t$)

$$\alpha = \frac{e^{kx+c}}{e^{kx+c} + e^{-kx-c}}, \qquad 1-\alpha = \frac{e^{-kx-c}}{e^{kx+c} + e^{-kx-c}}\cdot$$

b3) Of course

$$\begin{aligned}
E[Y_t] = \alpha \int x\, d\mu_1(x) + (1-\alpha)\int x\, d\mu_2(x) &= \frac{e^{kx+c}(kt+x) + e^{-kx-c}(-kt+x,)}{e^{kx+c} + e^{-kx-c}}\\
&= x + kt\tanh(kx+c)\,.
\end{aligned}$$

**12.7**

a) Let $B = (\Omega, \mathscr{F}, (\mathscr{F}_t)_t, (B_t)_t, P)$ be an $m$-dimensional Brownian motion. Recall that, by definition of Wiener measure, for every Borel set $A \subset \mathscr{C}$ we have $P(B \in A) = P^W(A)$. Hence $P^W(\mathscr{C}_0) = P(B_0 = 0) = 1$.

b) If $A$ is an open set containing the path 0, then it contains a neighborhood of 0 of the form $U = \{w; \sup_{0\le t\le T} |w(t)| < \eta\}$; let us show that $P^W(U) > 0$. In fact

$$P^W(U) = P\Big(\sup_{0\le t\le T} |B_t| < \eta\Big) \ge \prod_{i=1}^m P\Big(\sup_{0\le t\le T} |B_i(t)| < \frac{\eta}{\sqrt{m}}\Big)\,.$$

As

$$\mathrm{P}\Big(\sup_{0\le t\le T}|B_i(t)|<\frac{\eta}{\sqrt{m}}\Big)=\mathrm{P}(\tau_{\eta/\sqrt{m}}>T)>0$$

we have $\mathrm{P}^W(U)>0$.

c) Let us denote by $H_1\subset\mathscr{C}_0$ the set of the paths of the form $\gamma_t=\int_0^t\Phi_s\,ds$, $\Phi\in L^2$. As $H_1$ is dense in $\mathscr{C}_0$, if $A\subset\mathscr{C}_0$ is an open set (with respect to the topology induced by $\mathscr{C}$), then it contains a path $\gamma\in H_1$. Hence $\widetilde{A}=A-\gamma$ is an open set containing the origin. Let $V=\{w;\sup_{0\le t\le T}|w(t)|<\eta\}$ be a neighborhood of the origin of $\mathscr{C}_0$ contained in $\widetilde{A}$. We have

$$\mathrm{P}(B\in A)=\mathrm{P}(B-\gamma\in\widetilde{A})\ge\mathrm{P}(B-\gamma\in V)\,.$$

Let

$$Z=\exp\Big(\int_0^T\gamma'_s\,dB_s-\frac{1}{2}\int_0^T|\gamma'_s|^2\,ds\Big)\,,$$

and let Q be the probability defined by $d\mathrm{Q}=Z\,d\mathrm{P}$. Then, by Girsanov's theorem, for $t\le T$ the process $W_t=B_t-\gamma_t$ is a Brownian motion with respect to Q and we have

$$\mathrm{P}(B-\gamma\in V)=\mathrm{E}^{\mathrm{Q}}[Z^{-1}1_{\{B-\gamma\in V\}}]=\mathrm{E}^{\mathrm{Q}}[Z^{-1}1_{\{W\in V\}}]\,.$$

Now observe that $Z^{-1}>0$ Q-a.s. and that, thanks to b), $\mathrm{Q}(W\in V)=\mathrm{P}(B\in V)>0$ ($W$ is a Brownian motion with respect to Q). Hence $\mathrm{P}(B\in A)\ge\mathrm{E}^{\mathrm{Q}}[Z^{-1}1_{\{B\in V\}}]>0$.

**12.8**

a) Let us prove that $(Z_t)_t$ is a martingale for $t\le T$. We can take advantage of Corollary 12.1, which requires us to prove that for some value of $\mu>0$ $\mathrm{E}(\mathrm{e}^{\mu\theta^2|B_t+x|^2})<+\infty$ for every $t\le T$. But this is immediate as

$$\mathrm{E}(\mathrm{e}^{\mu\theta^2|B_t+x|^2})\le\mathrm{e}^{2\mu\theta^2|x|^2}\mathrm{E}(\mathrm{e}^{2\mu\theta^2|B_t|^2})=\mathrm{e}^{2\mu\theta^2|x|^2}\mathrm{E}(\mathrm{e}^{2\mu\theta^2B_1(t)^2})^m$$

and the right-hand side is finite for every $t\le T$ as soon as $\mu<(4\theta^2T)^{-1}$ (Exercise 1.12).

b) By Girsanov's theorem, with respect to Q the process

$$W_t=B_t-\theta\int_0^t(B_s+x)\,ds$$

is a Brownian motion for $t\le T$; therefore, if $Y_t=B_t+x$,

$$Y_t=B_t+x=x+\theta\int_0^t Y_s\,ds+W_s$$

and $(Y_t)_t$ is, with respect to Q, an Ornstein–Uhlenbeck process having $x$ as its initial value. Going back to the computations of Sect. 9.2, $Y_t$ is a Gaussian r.v. with mean $b = \mathrm{e}^{\theta t}x$ and covariance matrix $\Gamma = \frac{1}{2\theta}(\mathrm{e}^{2\theta t} - 1)I$.

c) By Ito's formula, with respect to P,

$$\int_0^T (B_s + x)\,dB_s = \frac{1}{2}\,(|B_T|^2 - mT) + \langle x, B_T\rangle = \frac{1}{2}\,(|B_T + x|^2 - |x|^2 - mT)\;,$$

hence (12.25). Therefore

$$\begin{aligned}
\mathrm{e}^{\frac{\theta}{2}(mT+|x|^2)}\mathrm{E}^{\mathrm{Q}}\Big[\mathrm{e}^{-\frac{\theta}{2}|B_T+x|^2}\Big] &= \mathrm{e}^{\frac{\theta}{2}(mT+|x|^2)}\mathrm{E}\Big[Z_T\mathrm{e}^{-\frac{\theta}{2}|B_T+x|^2}\Big]\\
= \mathrm{E}\Big[\exp\Big(\theta\int_0^T (B_s+x)\,dB_s - \frac{\theta^2}{2}\int_0^T |B_s+x|^2\,ds - \frac{\theta}{2}(|B_T+x|^2 - |x|^2 - mT)\Big)\Big]&\\
= \mathrm{E}\Big[\exp\Big(-\frac{\theta^2}{2}\int_0^T |B_s+x|^2\,ds\Big)\Big]\,.&
\end{aligned}$$

d) We can write $W = b + \sigma Z$ with $Z \sim N(0, I)$. Therefore, thanks to the independence of the components of $Z$,

$$\begin{aligned}
\mathrm{E}[\mathrm{e}^{\theta|W|^2}] = \mathrm{E}[\mathrm{e}^{\theta|b+\sigma Z|^2}] &= \mathrm{E}\Big[\exp\Big(\theta\sum_{i=1}^m (b_i + \sigma Z_i)^2\Big)\Big]\\
= \mathrm{E}\Big[\exp\Big(\theta\sum_{i=1}^m (b_i^2 + 2\sigma Z_i b_i + \sigma^2 Z_i^2)\Big)\Big] &= \prod_{i=1}^m \mathrm{e}^{\theta b_i^2}\mathrm{E}\big[\mathrm{e}^{2\theta\sigma b_i Z_i + \theta\sigma^2 Z_i^2}\big]\,.
\end{aligned}$$

Now

$$\begin{aligned}
\mathrm{E}\big[\mathrm{e}^{2\theta\sigma b_i Z_i + \theta\sigma^2 Z_i^2}\big] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{e}^{2\theta\sigma b_i z + \theta\sigma^2 z^2}\mathrm{e}^{-z^2/2}\,dz\\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{1}{2}((1-2\theta\sigma^2)z^2 - 4\theta\sigma b_i z)}\,dz\,.
\end{aligned}$$

This type of integral can be computed by writing the exponent in the form of the square of a binomial times the exponential of a term not depending on $z$, i.e.

$$\begin{aligned}
\cdots = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \exp\Big[-\frac{1-2\theta\sigma^2}{2}\Big(z^2 - \frac{4\theta\sigma b_i z}{1-2\theta\sigma^2}\Big)\Big]&\\
= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \exp\Big[-\frac{1-2\theta\sigma^2}{2}\Big(z - \frac{2\theta\sigma b_i}{1-2\theta\sigma^2}\Big)^2\Big]\,dz \times \exp\Big(\frac{2\theta^2\sigma^2 b_i^2}{1-2\theta\sigma^2}\Big)\,.&
\end{aligned}$$

The integral diverges if $2\sigma^2\theta \ge 1$, otherwise, recognizing in the integrand (but for the normalizing constant) the density of a Gaussian r.v. with variance $(1 - 2\theta\sigma^2)^{-1}$ we find

$$\begin{aligned} \mathrm{E}[\mathrm{e}^{\theta|W|^2}] &= \frac{1}{(1-2\sigma^2\theta)^{m/2}} \prod_{i=1}^{m} \mathrm{e}^{\theta b_i^2} \exp\Big(\frac{2\theta^2\sigma^2 b_i^2}{1-2\theta\sigma^2}\Big) \\ &= \frac{1}{(1-2\sigma^2\theta)^{m/2}} \exp\Big(\frac{\theta}{1-2\theta\sigma^2}\,|b|^2\Big)\,. \end{aligned} \tag{S.90}$$

e) Thanks to b) $B_T + x$ has, with respect to Q, a normal law with mean $b = \mathrm{e}^{\theta T} x$ and covariance matrix $\Gamma = \sigma^2 I$, with $\sigma^2 = \frac{1}{2\theta}(\mathrm{e}^{2\theta T} - 1)$. By c) and d), replacing $\theta$ with $-\frac{\theta}{2}$ in (S.90),

$$\begin{aligned} \mathrm{E}\Big[\exp\Big(-\frac{\theta^2}{2}\int_0^T |B_s + x|^2\,ds\Big)\Big] &= \mathrm{e}^{\frac{\theta}{2}\,(mT+|x|^2)}\mathrm{E}^{\mathrm{Q}}\Big[\mathrm{e}^{-\frac{\theta}{2}|B_T+x|^2}\Big] \\ &= \mathrm{e}^{\frac{\theta}{2}(mT+|x|^2)}\,\frac{1}{(1+\sigma^2\theta)^{m/2}}\,\exp\Big(-\frac{\theta}{2(1+\sigma^2\theta)}|x|^2\mathrm{e}^{2\theta T}\Big)\,. \end{aligned}$$

As $1 + \sigma^2\theta = \frac{1}{2}(\mathrm{e}^{2\theta T} + 1)$ we have

$$\begin{aligned} \frac{\mathrm{e}^{\frac{\theta}{2}mT}}{(1+\sigma^2\theta)^{m/2}} &= (\mathrm{e}^{\theta T})^{m/2}\Big(\frac{1}{2}\,(\mathrm{e}^{2\theta T}+1)\Big)^{-m/2} \\ &= \Big(\frac{1}{2}\,(\mathrm{e}^{\theta T} + \mathrm{e}^{-\theta T})\Big)^{-m/2} = \cosh(\theta T)^{-m/2} \end{aligned}$$

whereas

$$\begin{aligned} \mathrm{e}^{\frac{\theta}{2}|x|^2}\exp\Big(-\frac{\theta}{2(1+\sigma^2\theta)}|x|^2\mathrm{e}^{2\theta T}\Big) &= \exp\Big(\frac{\theta|x|^2}{2}\Big(1 - \frac{2\mathrm{e}^{2\theta T}}{\mathrm{e}^{2\theta T}+1}\Big)\Big) \\ &= \exp\Big(-\frac{\theta|x|^2}{2}\tanh(\theta T)\Big)\,, \end{aligned}$$

i.e.

$$\mathrm{E}\Big[\exp\Big(-\frac{\theta^2}{2}\int_0^T |B_s + x|^2\,ds\Big)\Big] = \cosh(\theta T)^{-m/2}\exp\Big[-\frac{\theta|x|^2}{2}\tanh(\theta T)\Big]\,. \tag{S.91}$$

f) We have from c) above and (S.91) with $\theta = \sqrt{2\lambda}$ and $T - t$ instead of $T$

$$\begin{aligned} u(x,t) &= \mathrm{E}\Big[\mathrm{e}^{-\lambda\int_0^{T-t}|B_s+x|^2\,ds}\Big] \\ &= \cosh\big(\sqrt{2\lambda}\,(T-t)\big)^{-m/2}\exp\Big[-\frac{\sqrt{2\lambda}\,|x|^2}{2}\tanh\big(\sqrt{2\lambda}\,(T-t)\big)\Big]\,. \end{aligned}$$

This solution, which is bounded, is unique among the functions having polynomial growth.

**12.9**

a) This is a typical application of Lemma 4.1 (the freezing lemma) as developed in Example 4.4, where it is explained that

$$\mathrm{E}[f(B_T)\,|\,\mathscr{F}_t] = \psi(B_t, t) := \frac{1}{(2\pi(T-t))^{m/2}} \int_{\mathbb{R}^m} f(z)\, \mathrm{e}^{-\frac{|z-B_t|^2}{2(T-t)}}\, dz\,.$$

Differentiability properties of $\psi$ (for $t < T$) follow from the theorem of differentiation under the integral sign. Moreover, $\psi(0,0) = \mathrm{E}[f(B_T)]$.

b) Ito's formula gives

$$dZ_t = d\psi(B_t, t) = \Big(\frac{\partial \psi}{\partial t}(B_t, t) + \frac{1}{2} \triangle \psi(B_t, t)\Big)\, dt + \psi'_x(B_t, t)\, dB_t$$

the term in $dt$, however, must vanish as $\psi(B_t, t) = \mathrm{E}[f(B_T)\,|\,\mathscr{F}_t]$ is a continuous square integrable martingale!

c1) The candidate integrand is of course $X_s = \psi'_x(B_s, s)$. We must prove that such a process belongs to $M^2([0,T])$, which is not granted in advance as $\psi$ is not necessarily differentiable in the $x$ variable for $t = T$ (unless $f$ were itself differentiable, as we shall see in d)). We have, however, for every $t < T$, by Jensen's inequality,

$$\mathrm{E}[f(B_T)^2] \ge \mathrm{E}\big[\mathrm{E}[f(B_T)\,|\,\mathscr{F}_t]^2\big] = \mathrm{E}[\psi(B_t, t)^2] = \mathrm{E}\Big[\Big(\psi(0,0) + \int_0^t \psi'_x(B_s, s)\, dB_s\Big)^2\Big].$$

Observing that $\psi(0,0) = \mathrm{E}[f(B_T)]$ we have for $t < T$

$$\begin{aligned} &\mathrm{E}\Big[\Big(\psi(0,0) + \int_0^t \psi'_x(B_s, s)\, dB_s\Big)^2\Big] \\ &= \mathrm{E}[f(B_T)]^2 + \mathrm{E}\Big[\int_0^t \psi'_x(B_s, s)^2\, ds\Big] + 2\mathrm{E}[f(B_T)]\, \underbrace{\mathrm{E}\Big[\int_0^t \psi'_x(B_s, s)\, dB_s\Big]}_{=0} \end{aligned}$$

so that, for every $t \le T$,

$$\int_0^t \mathrm{E}[|\psi'_x(B_s, s)|^2]\, ds \le \mathrm{E}[f(B_T)^2] - \mathrm{E}[f(B_T)]^2$$

and, taking the limit as $t \to T-$, we obtain

$$\int_0^T \mathrm{E}[|\psi_x'(B_s,s)|^2]\,ds \le \mathrm{E}[f(B_T)^2] - \mathrm{E}[f(B_T)]^2 < +\infty\,.$$

Hence $s \mapsto \psi_x'(B_s,s)$ belongs to $M^2([0,T])$. Now (12.27) follows from (12.28), as $f(B_T) = \lim_{t\to T-} \psi(B_t,t)$ a.s.

c2) If $f(x) = 1_{\{x>0\}}$, then, using once more the fact that $B_T - B_t \sim \sqrt{T-t}\,Z$ with $Z \sim N(0,1)$,

$$\psi(x,t) = \mathrm{E}[f(x+B_T-B_t)] = \mathrm{E}[1_{\{x+B_T-B_t>0\}}]=\mathrm{P}(B_T-B_t > -x) = \Phi\Big(\frac{x}{\sqrt{T-t}}\Big),$$

where we denote by $\Phi$ the partition function of an $N(0,1)$ law. Therefore

$$X_s = \frac{\partial\psi}{\partial x}(B_s,s) = \frac{1}{\sqrt{2\pi(T-s)}}\,\mathrm{e}^{-\frac{B_s^2}{2(T-s)}}\,.$$

d) We have

$$\begin{aligned}\psi(x,t) &= \frac{1}{(2\pi(T-t))^{m/2}}\int_{\mathbb{R}^m} f(z)\,\mathrm{e}^{-\frac{|z-x|^2}{2(T-t)}}\,dz\\ &= \frac{1}{(2\pi(T-t))^{m/2}}\int_{\mathbb{R}^m} f(x+z)\,\mathrm{e}^{-\frac{|z|^2}{2(T-t)}}\,dz\,.\end{aligned}$$

If $f$ is differentiable with bounded derivatives then the theorem of differentiation under the integral sign gives

$$\begin{aligned}\psi_x'(x,t) &= \frac{1}{(2\pi(T-t))^{m/2}}\int_{\mathbb{R}^m} f_x'(x+z)\,\mathrm{e}^{-\frac{|z|^2}{2(T-t)}}\,dz\\ &= \frac{1}{(2\pi(T-t))^{m/2}}\int_{\mathbb{R}^m} f'(z)\,\mathrm{e}^{-\frac{|z-x|^2}{2(T-t)}}\,dz\end{aligned}$$

hence, again by the freezing lemma,

$$\psi_x'(B_t,t) = \mathrm{E}[f_x'(B_T)\,|\,\mathscr{F}_t]\,.$$

**13.1** Recall (this is (13.22)) that under $\mathrm{P}^*$ the prices follow the SDE

$$dS_i(t) = r_t S_i(t)\,dt + \sum_{j=1}^d \sigma_{ij}(S_t,t)S_i(t)\,dB_j^*(t)\,,$$

where $(B_t^*)_t$ is a $\mathrm{P}^*$-Brownian motion. As in this model the processes $S_i$, $i = 1, \dots, m$, are in $M^2_{B^*}$, the stochastic component of the differential is a martingale and

$$\mathrm{E}^*[S_i(t)] = \int_0^t \mathrm{E}^*[r_s S_i(s)]\,ds + \mathrm{E}[S_i(0)] = \int_0^t \mathrm{E}^*[r_s]\mathrm{E}^*[S_i(s)]\,ds + \mathrm{E}^*[S_i(0)]\ .$$

Therefore the function $v(t) = \mathrm{E}^*[S_i(t)]$ satisfies the differential equation

$$v'(t) = \mathrm{E}^*[r_t]v(t)$$

with the initial condition $v(0) = \mathrm{E}^*[S_i(0)]$, so that $v(t) = \mathrm{e}^{\int_0^t \mathrm{E}^*[r_s]\,ds}\mathrm{E}^*[S_i(0)]$.

**13.2**

a) If $s \le t$ we have

$$\mathrm{E}^*[S_i(t)\,|\,\mathscr{F}_s] = \mathrm{e}^{\int_0^t r_u\,du}\mathrm{E}^*[\widetilde{S}_i(t)\,|\,\mathscr{F}_s] = \mathrm{e}^{\int_0^t r_u\,du}\widetilde{S}_i(s) = \mathrm{e}^{\int_s^t r_u\,du}S_i(s)\ ,$$

whereas, of course, $\mathrm{E}^*[S_i(t)\,|\,\mathscr{F}_s] = S_i(t)$ if $s \ge t$.

b) We have

$$\begin{aligned} V_t &:= \mathrm{E}^*[\mathrm{e}^{-\int_t^T r_u\,du} Z\,|\,\mathscr{F}_t] = \mathrm{e}^{-\int_t^T r_u\,du}\int_0^T \big(\alpha\mathrm{E}^*[S_i(s)\,|\,\mathscr{F}_t] + \beta\big)\,ds \\ &= \mathrm{e}^{-\int_t^T r_u\,du}\Big(\beta T + \int_0^t \alpha S_i(s)\,ds + \int_t^T \alpha\mathrm{E}^*[S_i(s)\,|\,\mathscr{F}_t]\,ds\Big) \\ &= \mathrm{e}^{-\int_t^T r_u\,du}\Big(\beta T + \int_0^t \alpha S_i(s)\,ds + \alpha S_i(t)\int_t^T \mathrm{e}^{\int_t^s r_u\,du}\,ds\Big)\ . \end{aligned} \tag{S.92}$$

As $S_0(t) = \mathrm{e}^{\int_0^t r_u\,du}$, we can write

$$V_t = \underbrace{\mathrm{e}^{-\int_0^T r_u\,du}\Big(\int_0^t \alpha S_i(s)\,ds + \beta T\Big)}_{:=H_0(t)} S_0(t) + \underbrace{\alpha\int_t^T \mathrm{e}^{-\int_s^T r_u\,du}\,ds}_{:=H_i(t)}\, S_i(t)\ .$$

Let us prove that $H$ is admissible. The condition $V_t \ge 0$ is verified by construction. In order to check that it is self-financing we shall verify condition (13.8) of Proposition 13.1. From (S.92) we have

$$\widetilde{V}_t = \mathrm{e}^{-\int_0^t r_u\,du}V_t = \mathrm{e}^{-\int_0^T r_u\,du}\Big(\beta T + \int_0^t \alpha S_i(s)\,ds + \alpha S_i(t)\int_t^T \mathrm{e}^{\int_t^s r_u\,du}\,ds\Big)\ . \tag{S.93}$$

Hence, taking the differential,

$$\begin{aligned}d\widetilde{V}_t &= \mathrm{e}^{-\int_0^T r_u\,du}\Big\{\alpha S_i(t)\,dt + \alpha S_i(t)\Big(-1-\int_t^T r_t \mathrm{e}^{\int_t^s r_u\,du}\,ds\Big)\,dt\\ &\qquad + \alpha\Big(\int_t^T \mathrm{e}^{\int_t^s r_u\,du}\,ds\Big)\,dS_i(t)\Big\}\\ &= \mathrm{e}^{-\int_0^T r_u\,du}\alpha\int_t^T \mathrm{e}^{\int_t^s r_u\,du}\,ds\big(-r_tS_i(t)\,dt + dS_i(t)\big)\ .\end{aligned}$$

Recall now, this is (13.7), that

$$d\widetilde{S}_i(t) = \mathrm{e}^{-\int_0^t r_u\,du}\big(-r_tS_i(t)\,dt + dS_i(t)\big)$$

so that

$$\begin{aligned}d\widetilde{V}_t &= \mathrm{e}^{-\int_t^T r_u\,du}\alpha\Big(\int_t^T \mathrm{e}^{\int_t^s r_u\,du}\,ds\Big)\,d\widetilde{S}_i(t) = \alpha\Big(\int_t^T \mathrm{e}^{-\int_s^T r_u\,du}\,ds\Big)\,d\widetilde{S}_i(t)\\ &= H_i(t)\,d\widetilde{S}_i(t)\ .\end{aligned}$$

c) The arbitrage price of the option at time 0 is obtained by taking $t = 0$ in (S.93): denoting $S_1(0)$ by $x$,

$$V_0 = \widetilde{V}_0 = \mathrm{e}^{-\int_0^T r_u\,du}\beta T + \alpha x\int_0^T \mathrm{e}^{-\int_s^T r_u\,du}\,ds\ .$$

- This is an example of an option that does not depend on the value of the underlying at time $T$ only, as is the case for calls and puts. Note also that the price of the option at time $t$, $V_t$, is a functional of the price of the underlying $S_i$ up to time $t$ (and not just of $S_i(t)$, as is the case with calls and puts).

**13.3**

a) The portfolio that we need to investigate enjoys two properties, first that it is self-financing, i.e. is such that

$$dV_t = H_0(t)\,dS_0(t) + H_1(t)\,dS_1(t) \tag{S.94}$$

and, second, that

$$H_1(t) = \frac{M}{S_1(t)}\,\cdot$$

Now $V_t = H_0(t)S_0(t) + H_1(t)S_1(t) = H_0(t)S_0(t) + M$. Hence

$$dV_t = H_0(t)\,dS_0(t) + S_0(t)\,dH_0(t)\;, \tag{S.95}$$

so (S.94) and (S.95) together give

$$H_0(t)\,dS_0(t) + H_1(t)\,dS_1(t) = H_0(t)\,dS_0(t) + S_0(t)\,dH_0(t)\;,$$

i.e. $S_0(t)\,dH_0(t) = H_1(t)\,dS_1(t)$, from which we obtain

$$dH_0(t) = \frac{H_1(t)}{S_0(t)}\,dS_1(t) = \frac{M}{S_0(t)S_1(t)}\,dS_1(t) = \frac{M\mathrm{e}^{-rt}}{S_1(t)}\,dS_1(t)\;.$$

As in the Black–Scholes model

$$dS_1(t) = S_1(t)(b\,dt + \sigma\,dB_t)$$

we obtain

$$dH_0(t) = M\mathrm{e}^{-rt}\,(b\,dt + \sigma\,dB_t)$$

hence

$$H_0(t) = H_0(0) + \frac{Mb}{r}\,(1 - \mathrm{e}^{-rt}) + \sigma M \int_0^t \mathrm{e}^{-rs}\,dB_s\;,$$

which gives for the value of the portfolio

$$\begin{aligned} V_t &= H_0(t)S_0(t) + H_1(t)S_1(t) = H_0(t)S_0(t) + M \\ &= \mathrm{e}^{rt}\Big(H_0(0) + \frac{Mb}{r}\,(1 - \mathrm{e}^{-rt}) + \sigma M \int_0^t \mathrm{e}^{-rs}\,dB_s\Big) + M\;. \end{aligned}$$

Of course,

$$\mathrm{E}[V_t] = \mathrm{e}^{rt}\Big(H_0(0) + \frac{Mb}{r}\,(1-\mathrm{e}^{-rt})\Big) + M = \mathrm{e}^{rt}\Big(V_0 - M + \frac{Mb}{r}\,(1-\mathrm{e}^{-rt})\Big) + M\;.$$

b) In order for $V$ to be admissible it is necessary that $V_t > 0$ a.s. for every $t$. This is not true in this case as $V_t$ has a Gaussian distribution with a strictly positive variance and such a r.v. is strictly negative with a strictly positive probability.

**13.4** Let us denote by $(S_t)_t$ the price process.

a) Obviously

$$Z = C\, 1_{\{\sup_{0\le u\le T} S_u > K\}}\ .$$

b) By definition (see Proposition 13.4) the arbitrage price of $Z$ at time 0 is

$$V_0 = \mathrm{e}^{-rT}\mathrm{E}^*[Z] = \mathrm{e}^{-rT} C\, \mathrm{P}^*\Big(\sup_{0\le u\le T} S_u > K\Big)\ .$$

Under the equivalent martingale measure $\mathrm{P}^*$ the price process $S$ with the starting condition $S_0 = x$ is the geometric Brownian motion

$$S_t = x\mathrm{e}^{(r-\frac{\sigma^2}{2})t+\sigma B_t}\ .$$

Therefore

$$\begin{aligned}
V_0 &= \mathrm{e}^{-rT} C\, \mathrm{P}^*\Big(\sup_{0\le u\le T} S_u > K\Big)\\
&= \mathrm{e}^{-rT} C\, \mathrm{P}^*\Big(\sup_{0\le u\le T} \big(x\mathrm{e}^{(r-\frac{\sigma^2}{2})u+\sigma B_u}\big) > K\Big)\\
&= \mathrm{e}^{-rT} C\, \mathrm{P}^*\Big(\sup_{0\le u\le T} \big((r-\frac{\sigma^2}{2})u + \sigma B_u\big) > \log K - \log x\Big)\\
&= \mathrm{e}^{-rT} C\, \mathrm{P}^*\Big(\sup_{0\le u\le T} \big((\frac{r}{\sigma}-\frac{\sigma}{2})u + B_u\big) > \frac{1}{\sigma}\log\tfrac{K}{x}\Big)\ .
\end{aligned}$$

This quantity is of course equal to $\mathrm{e}^{-rT}C$ if $x > K$, as in this case the prices are already larger than the level $K$ at time 0. If $x < K$, hence $\log\frac{K}{x} > 0$, the probability above (probability of crossing a positive level before a time $T$ for a Brownian motion with drift) has been computed in Example 12.5. Going back to (12.11) (we must replace $\mu$ with $-(\frac{r}{\sigma}-\frac{\sigma}{2})$ and $a$ with $\frac{1}{\sigma}\log\frac{K}{x}$) we have, denoting by $\Phi$ the partition function of an $N(0,1)$-distributed r.v.

$$\begin{aligned}
V_0 = \mathrm{e}^{-rT} C\,\Big\{&\mathrm{e}^{\frac{2}{\sigma}(\frac{r}{\sigma}-\frac{\sigma}{2})\log\frac{K}{x}}\Big[1-\Phi\Big(\frac{1}{\sqrt{T}}\big(\frac{1}{\sigma}\log\frac{K}{x}+(\frac{r}{2}-\frac{\sigma}{2})T\big)\Big)\Big]\\
&+1-\Phi\Big(\frac{1}{\sqrt{T}}\big(\frac{1}{\sigma}\log\frac{K}{x}-(\frac{r}{2}-\frac{\sigma}{2})T\big)\Big)\Big\}\ .
\end{aligned}$$

- This is another example where the payoff is not a function of the underlying asset at the final time $T$ (as in Exercise 13.2). One may wonder if in this case the price of the option can also be obtained as the solution of a PDE problem. Actually the answer is yes, but unfortunately this is not a consequence of Theorem 10.4

because not all of its assumptions are satisfied (the diffusion coefficient can vanish and the boundary data are not continuous at the point $(K, T)$).

**13.5** Let us first prove that there exist no equivalent martingale measure Q such that the discounted price processes $\widetilde{S}_1, \widetilde{S}_2$ are both martingales. Actually, Eq. (13.18) here becomes the system

$$\begin{cases} \sigma\gamma_t = r - \mu_1 \\ \sigma\gamma_t = r - \mu_2 \end{cases}$$

and, if $\mu_1 \neq \mu_2$, a solution $\gamma$ cannot exist. Note that here we have $d = 1$ (dimension of the Brownian motion) smaller than $m = 2$ (number of underlying assets). In this case the existence of an equivalent martingale measure is not ensured, as remarked on p. 416.

Also note that, if $\mu_1 > \mu_2$, then the strategy $H_0 \equiv 0$, $H_1 \equiv 1$, $H_2 \equiv -\frac{x_1}{x_2}$ (recall that $x_1 = S_1(0)$, $x_2 = S_2(0)$) produces an arbitrage portfolio. In words, it is a portfolio whose composition is constant in time and is obtained by buying at time 0 a unit of asset $S_1$ and short selling $\frac{x_1}{x_2}$ units of asset $S_2$. This operation requires no capital. The value at time $t$ of such a portfolio is

$$\begin{aligned} V_t(H) = S_1(t) - \frac{x_1}{x_2} S_2(t) &= x_1 \mathrm{e}^{(\mu_1 - \frac{\sigma^2}{2})t + \sigma B_t} - \frac{x_1}{x_2} x_2 \mathrm{e}^{(\mu_2 - \frac{\sigma^2}{2})t + \sigma B_t} \\ &= x_1(\mathrm{e}^{\mu_1 t} - \mathrm{e}^{\mu_2 t})\mathrm{e}^{-\frac{\sigma^2}{2}t + \sigma B_t}, \end{aligned}$$

which proves simultaneously that it is admissible ($V_t(H) \geq 0$ for every $t$) and with arbitrage, as $V_T(H) > 0$ a.s.

# References


Asmussen, S. (2000). *Ruin probabilities*, volume 2 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co., Inc., River Edge, NJ.

Azencott, R. and al. (1981). *Géodésiques et Diffusions en temps petit*. Astérisque 84–85.

Doss, H. (1977). Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré, B*, 13:99–126.

Fishman, G. S. (1996). *Monte Carlo*. Springer Series in Operations Research. Springer-Verlag, New York. Concepts, algorithms, and applications.

Friedman, A. (1964). *Partial Differential Equations of Parabolic Type*. Prentice Hall.

Friedman, A. (1975). *Stochastic Differential Equations and Applications, I and II*. Academic Press, New York.

Gallardo, L. (1981). Au sujet du contenu probabiliste d'un lemme d'Henri Poincaré. *Ann. Sci. Univ. Clermont-Ferrand II Math.*, (19):185–190. Saint-Flour Probability Summer Schools (Saint-Flour, 1979/1980).

Gilbarg, D. and Trudinger, N. S. (2001). *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin. Reprint of the 1998 edition.

Glasserman, P. (2004). *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics (New York)*. Springer-Verlag, New York. Stochastic Modelling and Applied Probability.

Graham, C. and Talay, D. (2013). *Stochastic simulation and Monte Carlo methods*, volume 68 of *Stochastic Modelling and Applied Probability*. Springer, Heidelberg. Mathematical foundations of stochastic simulation.

Han, Q. and Lin, F. (1997). *Elliptic partial differential equations*, volume 1 of *Courant Lecture Notes in Mathematics*. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI.

Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. North Holland.

Karatzas, I. and Shreve, S. (1991). *Brownian Motion and Stochastic Calculus, $2^{nd}$ edition*. Springer, Berlin, Heidelberg, New York.

Kloeden, P. E. and Platen, E. (1992). *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin.

Levi, E. E. (1907). Sulle equazioni lineari totalmente ellittiche alle derivate parziali. *Rend. Circolo. Mat. Palermo*, 24:275–317.

Maruyama, G. (1955). Continuous Markov processes and stochastic equations. *Rend. Circ. Mat. Palermo (2)*, 4:48–90.

Musiela, M. and Rutkowski, M. (2005). *Martingale methods in financial modelling*, volume 36 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, second edition.


© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2

Neveu, J. (1964). *Bases Mathématiques du Calcul des Probabilités*. Masson et Cie.

Priouret, P. (1974). Diffusions et équations différentielles stochastiques. In *Ecole d'été de Probabilités de St. Flour III-1973*, Lect. Notes Math. 390. Springer.

Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition.

Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge. Itô calculus, Reprint of the second (1994) edition.

Stroock, D. and Varadhan, S. (1979). *Multidimensional Diffusion Processes*. Grundlehren der Mathematisches Wissenschaften 233. Springer, Berlin, Heidelberg, New York.

Sussmann, H. (1978). On the gap between deterministic and stochastic differential equations. *Ann. Probab.*, 6:19–41.

Talay, D. and Tubaro, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1990).

# Index



© Springer International Publishing AG 2017

P. Baldi, *Stochastic Calculus*, Universitext, DOI 10.1007/978-3-319-62226-2