Francesca Biagini
Thilo Meyer-Brandis

SS 2012

Mathematical Finance in Continuous Time
Preface

These lecture notes were written in order to support the students of the graduate course “Finanzmathematik II” at the Mathematics Department of the Ludwig Maximilian University of Munich in their first approach to financial mathematics in continuous time.
I wish to thank S. Carstens, D. Filipovic and S. Ulmer for fruitful discussions and useful remarks that helped to improve these notes. A special thanks goes to Alessandra Cretarola who with infinite patience typed the first version of this text, and to Irene Schreiber and Bernhard Karl who continued the work of Alessandra.
Contents

Part I Stochastic Calculus

1 Preliminaries .............................................. 3
  1.1 Stochastic Processes .................................. 3
  1.2 Construction of the Brownian Motion .................. 6
  1.3 Properties of the Brownian Motion ..................... 11
  1.4 Adapted Processes ................................... 14
  1.5 Martingales ........................................... 15
  1.6 Stopping Theorems ................................... 17
  1.7 Martingale Inequalities ................................ 19
  1.8 Martingale Convergence ................................ 23
  1.9 Square Integrable Martingales ......................... 24
  1.10 Local Martingales and the Sharp Bracket .............. 26

2 Stochastic Integration ..................................... 33
  2.1 Integral w.r.t. Finite Variation Processes ............... 33
  2.2 Stochastic Integral w.r.t. Brownian Motion ............. 37
  2.3 The Stochastic Integral as a continuous Process ........ 42
  2.4 Extension to Integrands in $A^2_{loc}$ .................. 45
  2.5 The Vectorial Case ................................... 47
  2.6 Itô’s Formula .......................................... 47
  2.7 The Stochastic Exponential ............................... 52
  2.8 Girsanov’s Theorem ..................................... 54
  2.9 Brownian Martingales .................................. 58

3 Stochastic Differential Equations .......................... 63
  3.1 Definition of Stochastic Differential Equations ........ 63
  3.2 Existence and Uniqueness of Solutions ................... 65
  3.3 Markov Property of Solutions ............................ 67
  3.4 Examples: Geometric Brownian Motion and Ornstein-
              Uhlenbeck Process ................................. 69
### Part II Continuous Time Finance

4 **Financial Markets in Continuous Time** .......................... 73  
   4.1 The Market Model ........................................... 73  
   4.2 Arbitrage and Equivalent Local Martingale Measures .... 79  
   4.3 Risk-neutral Pricing of Contingent Claims ................. 86  
   4.4 Pricing and Hedging in Complete Markets ................. 90  
   4.5 Pricing and Hedging in Itô Diffusions Models via Partial  
      Differential Equations ...................................... 98  

5 **The Black-Scholes Model** ....................................... 101  
   5.1 The Black-Scholes formula ................................ 102  
   5.2 Sensitivity Analysis ....................................... 106  
   5.3 Historical and Implied Volatility ......................... 107  
   5.4 Exotic Options ........................................... 108  

6 **Forward and Future Contracts** .................................. 119  
   6.1 Forward Contracts ........................................ 119  
   6.2 Put-Call Parity ........................................... 120  
   6.3 Futures Contracts ........................................ 121  

### Part III Appendix

A **Probability Essentials** ........................................ 125  
B **Conditional Expectation** ..................................... 129  
C **Finite Variation Functions** .................................. 133  
D **Lévy Processes** ............................................... 135  
References .......................................................... 145  
Index .............................................................. 147
Introduction

The first model for financial markets in continuous time can be traced back to as far as Louis Bachelier in 1900, who in his PhD thesis modeled an asset price by means of a continuous-time stochastic process, the Brownian motion. However, the use of stochastic calculus in finance became widely accepted only after the papers of Fisher Black and Myron Scholes [1973] and Robert C. Merton [1973].

In this lecture notes we provide an overview of the main concepts for mathematical modeling of financial markets, such as the notion of no-arbitrage, market completeness and of pricing and hedging of contingent claims. This theory is presented in Chapter 4, where several modeling examples, such as the well-known Black-Scholes approach, are discussed in detail.

However in order to introduce mathematical models for financial markets in continuous time, we first need to understand the main features of particular kinds of stochastic processes and the related stochastic calculus, e.g. the Brownian motion, that we will later on use to model the dynamics of asset prices. Therefore Chapters 1 to 3 contain an overview of all necessary theoretical basics and concepts. To start with in Chapter 1 we give a short overview of the relevant preliminaries, e.g. the construction of the Brownian motion and its properties, stopping theorems, basic properties of martingales and martingale inequalities. The theory of stochastic integration with respect to the Brownian motion and Itô processes is developed in Chapter 2. Chapter 3 contains basics on stochastic differential equations and discusses some prominent examples, such as the geometric Brownian motion and Ornstein-Uhlenbeck processes.
Part I

Stochastic Calculus
1 Preliminaries

The main references for this chapter are Dellacherie and Meyer [5], Priouret [15], and Jacod and Protter [8].

1.1 Stochastic Processes

To start with in this section we introduce some basic concepts of stochastic processes. Let $I$ be a set of indices, $(E, \mathcal{E})$ a measurable space and $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. We always suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is complete (i.e. every subset of a measure zero event is in $\mathcal{F}$ and has obviously measure zero).

Definition 1.1.1. A family $X = (X_t)_{t \in I}$ of random variables from $(\Omega, \mathcal{F}, \mathbb{P})$ into $(E, \mathcal{E})$ is called a (stochastic) process with values in $(E, \mathcal{E})$.

Throughout this text if not said otherwise we will set $I = \mathbb{R}_+ = [0, \infty)$ or $I = [0, T]$ for $T > 0$, and $(E, \mathcal{E}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ for $d \geq 1$.

Definition 1.1.2. For every $\omega \in \Omega$ the map $t \mapsto X_t(\omega)$ is called realization, trajectory or sample path associated to $\omega$. A process $(X_t)_{t \in \mathbb{R}_+}$ is called
- right-continuous, if for all $\omega \in \Omega$ the map $t \mapsto X_t(\omega)$ is right-continuous.
- right-continuous almost surely, if for almost every $\omega \in \Omega$ the map $t \mapsto X_t(\omega)$ is right-continuous.

Analogously one can define the terms continuous (respectively almost surely continuous) and left-continuous (respectively almost surely left-continuous) processes.
Definition 1.1.3.  
i) Two processes \(X = (X_t)_{t \in \mathbb{R}^+}\) and \(X' = (X'_t)_{t \in \mathbb{R}^+}\) defined on \((\Omega, \mathcal{F}, \mathbb{P})\) and \((\Omega', \mathcal{F}', \mathbb{P}')\) both with values in \((E, \mathcal{E})\) are equivalent if for all \(t_1, \ldots, t_n \in \mathbb{R}^+\) and \(A_1, \ldots, A_n \in \mathcal{E}\)
\[
\mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n) = \mathbb{P}'(X'_{t_1} \in A_1, \ldots, X'_{t_n} \in A_n).
\]
i) Now assume that \(X\) and \(X'\) are both defined on the same probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Then we say that \(X'\) is a modification of \(X\), if
\[
\text{for every } t \quad X'_t = X_t \quad \text{a.s.}
\]
We say that \(X\) and \(X'\) are indistinguishable, if
\[
a.s. \quad X_t = X'_t \quad \text{for every } t.
\]

Remark 1.1.4. To be indistinguishable is stronger than to be a modification. Suppose \(X'\) is a modification of \(X\) and define \(A_t := \{\omega \mid X'_t(\omega) \neq X_t(\omega)\}\) for every fixed \(t \in \mathbb{R}^+\). Then \(\mathbb{P}(A_t) = 0\), but \(\mathbb{P}\left(\bigcup_{t \in \mathbb{R}^+} A_t\right)\) may not be zero, as shown by the following example. Let \(\Omega = [0, 1], \mathcal{F} = \mathcal{B}([0, 1])\) and denote by \(\mathbb{P}\) the Lebesgue measure. For \(t \in [0, 1]\) we consider the two processes \(X_t(\omega) = 1_{\{t\}}(\omega)\) and \(X'_t(\omega) \equiv 0\). Then for every fixed \(t\) we have that
\[
\mathbb{P}(X_t \neq X'_t) = \mathbb{P}(\{t\}) = 0,
\]
hence \(X'_t\) is a modification of \(X_t\). However, \(X_t\) and \(X'_t\) are not indistinguishable, as
\[
\bigcup_{t \in [0, 1]} A_t = \bigcup_{t \in [0, 1]} \{t\} = [0, 1].
\]
A stochastic process as introduced in Definition 1.1.1 can be seen as a function of two variables \(t\) and \(\omega\), i.e.
\[
X_t(\omega) = X(t, \omega).
\]

Lemma 1.1.5. Let \(X = (X_t)_{t \in \mathbb{R}^+}\) and \(Y = (Y_t)_{t \in \mathbb{R}^+}\) be two \(\mathbb{R}^d\)-valued processes that are right-continuous a.s. If for every \(t \geq 0\),
\[
X_t = Y_t \quad \text{a.s.,}
\]
then \(X\) and \(Y\) are indistinguishable.

Proof. Since \(X\) and \(Y\) are right-continuous a.s., then there exists a set \(N\) of measure zero, such that for all \(\omega \notin N\),
\[
X_t(\omega) \text{ and } Y_t(\omega)
\]
are right-continuous as functions of \(t\). Let \(N_t\) be the set of measure zero such that for all \(\omega \notin N_t\),
1.1 Stochastic Processes

\[ X_t(\omega) = Y_t(\omega). \]

Note that since \( X \) (respectively \( Y \)) is right-continuous, we can write

\[ X_t(\omega) = \lim_{n \to \infty} X_{nt+1}^{-} \omega). \]

Consider \( \tilde{N} = N \cup \left( \bigcup_{t \in \mathbb{Q}^+} N_t \right) \). Then \( \mathbb{P}(\tilde{N}) = 0 \) and for all \( \omega \notin \tilde{N} \),

\[ X_t(\omega) = Y_t(\omega), \quad \text{for every } t. \]

\[ \square \]

**Definition 1.1.6.** A process \( (X_t)_{t \in \mathbb{R}^+} \) with values in \( (E, \mathcal{E}) \) is measurable if the map

\[ (t, \omega) \mapsto X_t(\omega) \]

from \( (\mathbb{R}^+ \times \Omega, \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}) \) into \( (E, \mathcal{E}) \) is measurable.

As an important technical condition throughout this text we will always assume that all stochastic processes are measurable.

**Definition 1.1.7.**

i) A process \( (X_t)_{t \in \mathbb{R}^+} \) with values in \( \mathbb{R}^d \) has independent increments, if for every choice of \( 0 \leq t_0 < t_1 < \cdots < t_n \) in \( \mathbb{R}^+ \) the random variables

\[ X_{t_{k+1}} - X_{t_k}, \quad k = 0, \ldots, n - 1 \]

are independent. We call this kind of process an independent increments process.

ii) A process \( (X_t)_{t \in \mathbb{R}^+} \) with values in \( \mathbb{R}^d \) has stationary increments if for every \( t, h > 0 \), the law, or equivalently the distribution function, of

\[ X_{t+h} - X_t \]

is the same as the law of \( X_h - X_0 \).

**Definition 1.1.8.** A process \( (X_t)_{t \in \mathbb{R}^+} \) with values in \( \mathbb{R}^d \) is a Gaussian process if for every \( 0 \leq t_1 < \cdots < t_n \)

\[ \text{vec}(X_{t_1}, \ldots, X_{t_n}) \]

is a Gaussian vector with values in \( \mathbb{R}^{dn} \), i.e. \( \text{vec}(X_{t_1}, \ldots, X_{t_n}) \) is multivariate normally distributed. Note that \( \text{vec}(\cdot) \) is the column stacking operator that stacks the columns of an \( d \times n \) matrix as a vector of dimension \( dn \).

**Lemma 1.1.9.** An independent increments process \( (X_t)_{t \in \mathbb{R}^+} \) such that \( X_0 = 0 \) with Gaussian increments (i.e. for all \( s < t \), \( X_t - X_s \) has Gaussian law) is a Gaussian process.
Proof. Recall, that Gaussian random variables are invariant under linear transformations. Then the proof is straightforward, as \((X_t, X_{t_2}, \ldots, X_{t_n})\) is a linear function of the independent Gaussian random variables \(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}\), since for all \(j = 1, \ldots, n\), \(X_{t_j} = \sum_{k=1}^{j} (X_{t_k} - X_{t_{k-1}})\), where we have set \(X_{t_0} := X_0 = 0\).

Definition 1.1.10. A process \(B = (B_t)_{t \in \mathbb{R}^+}\) with values in \(\mathbb{R}^d\) is a (standard) Brownian motion (or Wiener process) if

i) \(B_0 = 0\) a.s.,

ii) \(B\) has stationary independent increments,

iii) for every \(s < t\), \(B_t - B_s \sim \mathcal{N}(0, (t-s)I_d)\),

iv) \(B\) has continuous sample paths a.s.,

where as usual \(I_d \in \mathbb{R}^{d \times d}\) represents the identity matrix.

In particular for \(d = 1\), \(E[B_tB_s] = t \wedge s\).

Proposition 1.1.11. \(B = (B^1, \ldots, B^d)\) is a \(d\)-dimensional (standard) Brownian motion, if and only if \(B^1, \ldots, B^d\) are independent 1-dimensional Brownian motions.

Proof. Follows directly by the properties of the multivariate normal distribution.

The name Brownian motion is due to the biologist Robert Brown, who as early as the beginning of the 19th century used this type of process to describe the path of a particle in a glass of water. However, the mathematical theory of Brownian motion was developed only about a century later by Louis Bachelier (1900), Albert Einstein (1905) and Norbert Wiener (1923).

Before examining properties and applications, we start with a natural question: under which circumstances does a process with properties as required in Definition 1.1.10 exist? An answer to this question is provided for example by the construction of a 1-dimensional Brownian motion of Paul Lévy, which has the advantage that it directly shows how to simulate the paths of the Brownian motion. Note that once we have constructed a 1-dimensional Brownian motion we can build a \(d\)-dimensional Brownian motion for any \(d \in \mathbb{N}\) by Proposition 1.1.11. For other possible constructions we refer to Karatzas and Shreve [9].

1.2 Construction of the Brownian Motion

We start with a helpful lemma, which we will need later on for proving the existence of the Brownian motion.

Lemma 1.2.1. Let \((X_n)_{n \in \mathbb{N}}\) be a sequence of centered Gaussian random vectors with values in \(\mathbb{R}^n\). Suppose that \((X_n)_{n \in \mathbb{N}}\) converges to \(X\) in probability. Then \(X\) is Gaussian and
1.2 Construction of the Brownian Motion

\[ X_n \xrightarrow{n \to \infty} X \text{ also in } L^2. \]

**Proof.** We consider the case \( n = 1 \), the general case works analogously. Since \( X_n \overset{P}{\to} X \), then \( X_n \overset{L^2}{\to} X \) and for \( \sigma_n^2 = E[X_n^2] \) we have that

\[ \Phi_{X_n}(t) = E[e^{i X_n t}] = e^{-\frac{\sigma_n^2 t^2}{2}} \xrightarrow{n \to \infty} \Phi_X(t), \]

because of Paul Lévy’s theorem (see Theorem A.0.7). Hence

\[ \sigma_n^2 = -2 \log \Phi_{X_n}(1) \xrightarrow{n \to \infty} \sigma^2 < \infty \]

and

\[ \Phi_X(t) = e^{-\frac{\sigma^2 t^2}{2}}, \]

which implies \( X \) is Gaussian. We furthermore note that

\[ E[e^{X_n}] = \int_\mathbb{R} e^x \frac{e^{-\frac{x^2}{2\sigma_n^2}}}{\sqrt{2\pi\sigma_n^2}} \, dx = e^{\frac{\sigma^2}{2}} \int_\mathbb{R} e^{-\frac{(x-\sigma_n^2)^2}{2\sigma^2 n}} \frac{1}{\sqrt{2\pi\sigma_n^2}} \, dx \geq e^{\frac{\sigma^2}{2}}, \]

and therefore

\[ \frac{1}{3!} E[|X_n|^3] \leq E[|e^{X_n}|] \leq E[e^{X_n}] + E[e^{-X_n}] = 2 e^{\frac{\sigma^2}{2}} \leq 2 \left( \varepsilon_n + \frac{\sigma^2}{2} \right), \]

where \( \varepsilon_n \xrightarrow{n \to \infty} 0 \) since \( \sigma_n^2 \xrightarrow{n \to \infty} \sigma^2 \). Hence we conclude that

\[ \sup_{n \in \mathbb{N}} E[|X_n|^3] < \infty \]

and \( X_n \xrightarrow{n \to \infty} X \) in \( L^2 \) by Proposition A.0.5 of Appendix A. \( \square \)

We now construct an approximation of the Brownian motion for \( t \in [0, 1] \). Let \( (Y_{k,n}, 0 \leq k < 2^n, n \geq 0) \) be independent random variables with standard normal distribution \( N(0, 1) \). We recursively define a sequence of processes \( X_n(t) \) as follows:

i) \( X_0(0) = 0, \quad X_0(1) = Y_{0,0}. \)

ii) \( X_{n+1}(\frac{k+1}{2^{n+1}}) = X_n(\frac{k}{2^n}), \quad X_{n+1}(1) = X_n(1), \)

\[ X_{n+1}(\frac{2k+1}{2^{n+1}}) = \frac{1}{2} \left( X_n(\frac{k+1}{2^n}) + X_n(\frac{k-1}{2^n}) \right) + 2^{-\frac{\sigma_n^2}{2}} Y_{2k+1,n+1}. \]

iii) For \( t \in \left[ \frac{k}{2^n}, \frac{k+1}{2^n} \right] \) \( X_n(t) \) is linear over \( \left[ \frac{k}{2^n}, \frac{k+1}{2^n} \right], \) i.e.

\[ X_n(t) = 2^n \left[ X_n(\frac{k+1}{2^n}) - X_n(\frac{k}{2^n}) \right] \left( t - \frac{k}{2^n} \right) + X_n(\frac{k}{2^n}). \]

**Example 1.2.2.** We compute some of the first \( X_n(t) \):
Lemma 1.2.3. The vector \( \left( X_n \left( \frac{k}{2^n} \right), k = 0, \ldots, 2^n \right) \) is Gaussian with mean zero and
\[
E \left[ X_n \left( \frac{k}{2^n} \right) X_n \left( \frac{k'}{2^n} \right) \right] = \frac{k}{2^n} \land \frac{k'}{2^n}
\]
for \( k, k' \in \{0, \ldots, 2^n \} \).

Proof. By induction. If the statement is true until level \( n \), we obtain that 
\( \left( X_{n+1} \left( \frac{k}{2^{n+1}} \right), k = 0, \ldots, 2^{n+1} \right) \) is a linear function of the Gaussian and independent vectors 
\( \left( X_n \left( \frac{k}{2^n} \right), k = 0, \ldots, 2^n \right) \) and 
\( \left( Y_{k,n+1}, k = 0, \ldots, 2^{n+1} \right) \).

Moreover
\[
E \left[ X_{n+1}^2 \left( \frac{2k+1}{2^{n+1}} \right) \right] =
\]
\[
= E \left[ \left\{ \frac{1}{2} \left( X_n \left( \frac{k}{2^n} \right) + X_n \left( \frac{k+1}{2^n} \right) \right) + 2^{-\frac{n+2}{2}} Y_{2k+1,n+1} \right\}^2 \right]
\]
\[
= \frac{1}{4} E \left[ X_n \left( \frac{k}{2^n} \right)^2 \right] + \frac{1}{4} E \left[ X_n \left( \frac{k+1}{2^n} \right)^2 \right] + \frac{1}{2} E \left[ X_n \left( \frac{k}{2^n} \right) X_n \left( \frac{k+1}{2^n} \right) \right] + 2^{-(n+2)} E \left[ Y_{2k+1,n+1}^2 \right]
\]
\[
= \frac{1}{4} \left( \frac{k}{2^n} + \frac{k+1}{2^n} + \frac{2k}{2^n} + \frac{1}{2^n} \right) = \frac{2k+1}{2^{n+1}}.
\]

For the mixed terms the computation is similar. \( \Box \)

We can conclude that \( X_n(t) \) is centered and Gaussian since \( \left( X_n(t_1), \ldots, X_n(t_k) \right) \) is a linear combination of \( \left( X_n \left( \frac{k}{2^n} \right), k = 0, \ldots, 2^n \right) \).

We prove the following lemma.

Lemma 1.2.4. \( E[X_n(s)X_n(t)] = s \land t + \varepsilon_n \), \hspace{1cm} (1.1)

where \( |\varepsilon_n| \leq \frac{1}{2^n} \).
Proof. For \( s, t \in \{0, \frac{1}{2^n}, \ldots, 1\} \) we have already proved (1.1) in Lemma 1.2.3. For \( s, t \in (\frac{k}{2^n}, \frac{k+1}{2^n}) \) we have

\[
E[X_n(s)X_n(t)] = E\left[\left\{2^n \left[X_n\left(\frac{k+1}{2^n}\right) - X_n\left(\frac{k}{2^n}\right)\right] \left(t - \frac{k}{2^n}\right) + X_n\left(\frac{k}{2^n}\right)\right\} \cdot \left\{2^n \left[X_n\left(\frac{k+1}{2^n}\right) - X_n\left(\frac{k}{2^n}\right)\right] \left(s - \frac{k}{2^n}\right) + X_n\left(\frac{k}{2^n}\right)\right\}\right]
\]

\[
= 2^n \left(t - \frac{k}{2^n}\right) \left(s - \frac{k}{2^n}\right) E\left[\left(X_n\left(\frac{k+1}{2^n}\right) - X_n\left(\frac{k}{2^n}\right)\right) \left(X_n\left(\frac{k}{2^n}\right)\right)\right]
\]

Note that

\[
2^n \left(t - \frac{k}{2^n}\right) \left(s - \frac{k}{2^n}\right) \leq \frac{1}{2^n}
\]

and that we can rewrite

\[
k \frac{2^n}{2^n} = s \land t + \left(\frac{k}{2^n} - s \land t\right).
\]

Since \(|\frac{k}{2^n} - s \land t| \leq \frac{1}{2^n}\), (1.1) is proved for \( s, t \in (\frac{k}{2^n}, \frac{k+1}{2^n}) \).

Finally, if \( s \in (\frac{k}{2^n}, \frac{k+1}{2^n}) \) and \( t \in (\frac{j}{2^n}, \frac{j+1}{2^n}) \) with \( k + 1 \leq j \), we obtain that

\[
E[X_n(s)X_n(t)] =
\]

\[
= E\left[\left\{2^n \left[X_n\left(\frac{j+1}{2^n}\right) - X_n\left(\frac{j}{2^n}\right)\right] \left(t - \frac{j}{2^n}\right) + X_n\left(\frac{j}{2^n}\right)\right\} \cdot \left\{2^n \left[X_n\left(\frac{j+1}{2^n}\right) - X_n\left(\frac{j}{2^n}\right)\right] \left(s - \frac{j}{2^n}\right) + X_n\left(\frac{j}{2^n}\right)\right\}\right]
\]

\[
= 2^n \left(s - \frac{k}{2^n}\right) \left(t - \frac{j}{2^n}\right)
\]

\[
\cdot E\left[\left(X_n\left(\frac{k+1}{2^n}\right) - X_n\left(\frac{k}{2^n}\right)\right) \left(X_n\left(\frac{j+1}{2^n}\right) - X_n\left(\frac{j}{2^n}\right)\right)\right]
\]

\[
+ 2^n \left(s - \frac{k}{2^n}\right) E\left[\left(X_n\left(\frac{k+1}{2^n}\right) - X_n\left(\frac{k}{2^n}\right)\right) X_n\left(\frac{j}{2^n}\right)\right]
\]

\[
+ 2^n \left(t - \frac{j}{2^n}\right) E\left[\left(X_n\left(\frac{j+1}{2^n}\right) - X_n\left(\frac{j}{2^n}\right)\right) X_n\left(\frac{k}{2^n}\right)\right]
\]

\[
+ E\left[X_n\left(\frac{k}{2^n}\right) X_n\left(\frac{j}{2^n}\right)\right] = s - \frac{k}{2^n} + \frac{k}{2^n} = s \land t.
\]
Lemma 1.2.5. There exists a constant $c > 0$ such that
\[
P \left( \sup_{t \in [0, 1]} |X_{n+1}(t) - X_n(t)| > 2^{-\frac{n}{4}} \right) \leq c 2^{-n}.
\]

Proof. Recall that if $Y \sim N(0, 1)$, then for $a > 0$
\[
P(|Y| > a) \leq a - 8 E[|Y|^8].
\]
Therefore we obtain
\[
P \left( \sup_{t \in [0, 1]} |X_{n+1}(t) - X_n(t)| > 2^{-\frac{n}{4}} \right) = \sum_{k=0}^{2^{n-1}} P \left( \bigcup_{t \in \left[ \frac{2k + 1}{2^n} \right]} |X_{n+1}(t) - X_n(t)| > 2^{-\frac{n}{4}} \right)
\]
\[
= \sum_{k=0}^{2^{n-1}} P \left( 2^{-\frac{n}{4}} |Y_{2k+1,n+1}| > 2^{-\frac{n}{4}} \right)
\]
\[
\leq 2^{-n} \sum_{k=0}^{2^{n-1}} P \left( |Y| > 2^{\frac{n+4}{4}} \right) \leq c 2^{-n}.
\]

Equation (*) is obtained as follows:

(1) For $t \in \left[ \frac{k}{2^n}, \frac{2k+1}{2^n} \right]$
\[
|X_{n+1}(t) - X_n(t)| = 2^{n+1} \left( X_n \left( \frac{k+1}{2^n} \right) + X_n \left( \frac{k}{2^n} \right) \right) \left( t - \frac{k}{2^n} \right) + X_n \left( \frac{k}{2^n} \right)
\]
\[
\leq 2^n \left( X_n \left( \frac{k+1}{2^n} \right) - X_n \left( \frac{k}{2^n} \right) \right) \left( t - \frac{k}{2^n} \right) - X_n \left( \frac{k}{2^n} \right)
\]
\[
= 2^{n+1} \left( \frac{1}{2} \left( X_n \left( \frac{k}{2^n} \right) + X_n \left( \frac{k+1}{2^n} \right) \right) + 2^{-\frac{n+2}{2}} Y_{2k+1,n+1} \right) \left( t - \frac{k}{2^n} \right)
\]
\[
- 2^n \left( X_n \left( \frac{k+1}{2^n} \right) - X_n \left( \frac{k}{2^n} \right) \right) \left( t - \frac{k}{2^n} \right)
\]
\[
= 2^\frac{n}{2} Y_{2k+1,n+1} \left( t - \frac{k}{2^n} \right)
\]

(2) For $t \in \left[ \frac{2k+1}{2^n}, \frac{k+1}{2^n} \right]$
\[
|X_{n+1}(t) - X_n(t)|
\]
We conclude as follows.

i) By applying the Borel-Cantelli lemma (Theorem A.0.1) to Lemma 1.2.5 we obtain that there exists a set \( N \) of measure zero such that for every \( \omega \not\in N \) there exists \( n_0(\omega) \in \mathbb{N} \) such that

\[
\sup_{t \in [0,1]} |X_{n+1}(t,\omega) - X_n(t,\omega)| \leq 2^{-\frac{n_0}{4}}
\]

for every \( n > n_0(\omega) \). Hence \( X_n(t,\omega) \) converge uniformly on \([0,1]\) and the limit \( X(t,\omega) = X_t(\omega) \) is continuous in \( t \).

ii) \((X_t)_{t \in [0,1]}\) is a centered Gaussian process. This follows by applying Lemma 1.2.1 to the sequence of centered Gaussian random variables \( X_n(t) \). Since the convergence also holds in \( L^2 \), by Lemma 1.2.4 we also have that

\[
E[X(t)X(s)] = t \wedge s.
\]

(1.2)

iii) To conclude we need to show that \((X_t)_{t \in [0,1]}\) has independent increments. Since the vector

\[
(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}})
\]

is Gaussian as a linear function of \((X_{t_1}, X_{t_2}, \ldots, X_{t_n})\), we only need to show that for \( j < k \)

\[
E\left[(X_{t_j} - X_{t_{j-1}})(X_{t_k} - X_{t_{k-1}})\right] = 0,
\]

this, however, follows by (1.2).

1.3 Properties of the Brownian Motion

In the following we investigate some useful properties of the standard Brownian motion.

**Proposition 1.3.1.** Let \( B = (B_t)_{t \in \mathbb{R}^+} \) be a real-valued standard Brownian motion on \((\Omega, \mathcal{F}, P)\). Consider

i) for \( s \geq 0 \), \( B^{(s)} = (B_t^{(s)})_{t \in \mathbb{R}^+} \), where \( B_t^{(s)} = B_{t+s} - B_s \),

ii) for \( c > 0 \), \( Y = (Y_t)_{t \in \mathbb{R}^+} \), where \( Y_t = cB_{t/c^2} \),

iii) for \( t \neq 0 \), \( Z = (Z_t)_{t \in \mathbb{R}^+} \), where \( Z_t = tB_{1/t}, Z_0 = 0 \).
Then $-B, B^{(s)}, Y$ and $Z$ are standard Brownian motions.

Proof. It is easy to see that $-B, B^{(s)}$ and $Y$ are centered Gaussian processes with continuous paths, independent increments, and variance $t$, hence Brownian motions. To prove (iii), we check that $Z$ is a centered Gaussian process with

$$\text{cov}(Z_s, Z_t) = E[Z_s Z_t] = st \ E[B_{1/s} B_{1/t}] = st \left( \frac{1}{s} \wedge \frac{1}{t} \right) = s \wedge t,$$

thus it will be a Brownian motion if its paths are continuous. Since they are clearly continuous on $(0, \infty)$, it is sufficient to show that $\lim_{t \to 0} Z_t = 0$. Note that $\lim_{t \to 0} Z_t = \lim_{t \to \infty} B_t / t$; therefore we now prove that $\lim_{t \to \infty} B_t / t = 0$. Let $n \in \mathbb{N}$. Since $B$ is a continuous martingale, by Theorem 1.7.4 (Doob’s maximal inequalities in continuous time) we obtain that

$$E \left( \left( \sup_{2^n \leq t \leq 2^{n+1}} \left| B_t \right| / t \right)^2 \right) \leq \frac{1}{2^{2n}} E \left( \sup_{2^n \leq t \leq 2^{n+1}} \left| B_t \right|^2 \right) \leq \frac{4}{2^{2n}} \sup_{2^n \leq t \leq 2^{n+1}} E[B_t^2] = \frac{4}{2^{2n}} 2^{n+1} = 8 \cdot 2^{-n}.$$

Let $\varepsilon > 0$ and $A_n := \left\{ \sup_{2^n \leq t \leq 2^{n+1}} \left| B_t \right| / t > \varepsilon \right\}$. By Chebyshev’s inequality

$$P(A_n) \leq \frac{1}{\varepsilon^2} E \left( \left( \sup_{2^n \leq t \leq 2^{n+1}} \left| B_t \right| / t \right)^2 \right) \leq \frac{8 \cdot 2^{-n}}{\varepsilon^2}.$$

Since $\sum_{n \geq 1} P[A_n] < \infty$, we have $P\left( \lim_{n} \sup A_n \right) = 0$ by the Borel-Cantelli lemma, i.e. $P\left( \lim_{n} \sup \left( \sup_{2^n \leq t \leq 2^{n+1}} \left| B_t \right| / t \right) \right) = 0$ for any $\varepsilon > 0$. Hence

$$P \left( \lim_{t \to \infty} \frac{B_t}{t} = 0 \right) = 1,$$

and therefore $\lim_{t \to 0} Z_t = 0$ a.s. Because of the existence of a modification of $B$ with continuous trajectories, we finally have that $\lim_{t \to 0} Z_t = 0$. \qed

Proposition 1.3.2. Let $B = (B_t)_{t \in \mathbb{R}_+}$ be a real-valued standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Then
1.3 Properties of the Brownian Motion

1.3.2

\[
\begin{align*}
1) \limsup_{t \to \infty} \frac{B_t}{\sqrt{t}} &= +\infty \text{ a.s.}, \quad \liminf_{t \to \infty} \frac{B_t}{\sqrt{t}} = -\infty \text{ a.s.} \\
2) \limsup_{t \to 0} \frac{B_t}{\sqrt{t}} &= +\infty \text{ a.s.}, \quad \liminf_{t \to 0} \frac{B_t}{\sqrt{t}} = -\infty \text{ a.s.} \\
3) \limsup_{t \to -\infty} \frac{B_t}{t} &= 0 \text{ a.s.}
\end{align*}
\]

Proof. i) Let \( R = \limsup_{t \to \infty} \frac{B_t}{\sqrt{t}} \). Not that for any \( s > 0 \), we have that

\[
R = \limsup_{t \to \infty} \frac{B_t + s - B_s}{\sqrt{t}}.
\]

By Proposition 1.3.1, for \( s > 0 \) \((B_{t+s} - B_s) \in \mathbb{R}_+\) is a Brownian motion that is independent of what happened before time \( s \), i.e. \( R \) is independent of \( \sigma(B_u, u \leq s) \) for any \( s > 0 \). Hence \( R \) is also \( \sigma(B_u, u \geq 0) \)-measurable it is also independent of itself. Therefore we conclude that \( R \) must be deterministic, i.e \( \mathbb{P}(R = +\infty) = 1 \) or \( \mathbb{P}(R = -\infty) = 1 \) for some \( \alpha > 0 \). Suppose \( \mathbb{P}(R = \alpha) = 1 \) for some \( \alpha \geq 0 \). Then for any \( \beta > \alpha \) we obtain that \( \mathbb{P}\left(\frac{B_t}{\sqrt{t}} > \beta\right) \to 0 \) as \( t \to \infty \), but \( \mathbb{P}\left(\frac{B_t}{\sqrt{t}} > \beta\right) = \mathbb{P}(B_1 > \beta) > 0 \). Therefore \( R = +\infty \) a.s. by contradiction. Considering \(-R\) we conclude that \( \liminf_{t \to \infty} \frac{B_t}{\sqrt{t}} = -\infty \) a.s.

ii) \( \limsup_{t \to 0} \frac{B_t}{\sqrt{t}} = \limsup_{t \to \infty} \frac{B_{t+1/2} - B_{1/2}}{\sqrt{t+1/2}} = \limsup_{t \to \infty} \frac{B_{t+1/2}}{\sqrt{1/2}} = +\infty \) a.s., analogously \( \liminf_{t \to 0} \frac{B_t}{\sqrt{t}} = -\infty \) a.s.

iii) \( \limsup_{t \to -\infty} \frac{B_t}{t} = \lim_{t \to 0} \frac{B_{t+1/2}}{t+1/2} = \lim_{t \to 0} \frac{Z_t}{t} = 0 \) a.s. by Proposition 1.3.1. \( \square \)

From the previous propositions it follows that the trajectories of \( B \) reach every point an infinite number of times and furthermore are nowhere differentiable.

Proposition 1.3.3. Let \( B = (B_t)_{t \in \mathbb{R}_+} \) be a real-valued standard Brownian motion on \((\Omega, \mathcal{F}, \mathbb{P})\). Then

i) almost surely \( B \) reaches every point an infinite number of times

ii) for every \( t \geq 0 \), \( B \) is almost surely neither right-differentiable, nor left-differentiable (clearly for \( t = 0 \) only the right-derivative is well-defined).

Proof. i) By Proposition 1.3.2 and by the continuity of the trajectories of \( B \) it follows that \( B \) reaches every point almost surely, and clearly the same holds for \( B^{(s)} \) for any \( s > 0 \), i.e. for any \( s > 0 \) \((B_{t+s})_{t \in \mathbb{R}_+} \) reaches every point almost surely. Therefore almost surely \( B \) reaches every point an infinite number of times.

ii) From Proposition 1.3.2 it follows that \( \limsup_{t \to 0} \frac{B_t - B_0}{\sqrt{t}} = +\infty \) a.s. Hence
$B_t$ is a.s. not right-differentiable for $t = 0$. Considering first $B_t^{(s)}$ and then $Z_t$ we conclude that $B_t$ is a.s. neither right-differentiable nor left-differentiable for every $t \geq 0$. \qed

1.4 Adapted Processes

Recall that throughout the whole of this text we assume given a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

**Definition 1.4.1.**

i) An increasing family $\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}$ of sub-$\sigma$-algebras of $\mathcal{F}$ is called a filtration and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{R}^+}, \mathbb{P})$ a filtered probability space.

ii) We define $\mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \geq 0)$. If every $\mathcal{F}_t$ contains all sets of measures zero of $\mathcal{F}_\infty$, the filtration $\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}$ is said to be complete.

iii) Consider $\mathcal{F}_{t+} = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$.

Then $\mathcal{F}_{t+}$ is a $\sigma$-algebra and $\{\mathcal{F}_{t+}\}_{t \in \mathbb{R}^+}$ also is a filtration. If $\mathcal{F}_t = \mathcal{F}_{t+}$ for all $t \geq 0$, we say that $\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}$ is right-continuous.

iv) A filtration $\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}$ is standard if it is complete and right-continuous. In this case we also say that $\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}$ satisfies the usual conditions.

Note that there is a general understanding that a filtration should be interpreted as the amount of information that we know at each time $t$.

**Remark 1.4.2.** We can always complete a filtration $\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}$ by replacing $\mathcal{F}_t$ with $\mathcal{F}_t = \sigma(\mathcal{F}_t, \mathcal{N})$, where $\mathcal{N}$ is the set of all measure zero events of $\mathcal{F}_\infty$.

**Definition 1.4.3.** Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{R}^+}, \mathbb{P})$ be a filtered probability space. A stochastic process $X = (X_t)_{t \in \mathbb{R}^+}$ on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{R}^+}, \mathbb{P})$ with values in $(E, \mathcal{E})$ is called $\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}$-adapted, if $X_t$ is $\mathcal{F}_t$-measurable for all $t \geq 0$.

If it is clear which filtration we are referring to, we simply call $X$ adapted instead of $\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}$-adapted.

Often it is convenient to consider $X$ together with the filtration that is generated by $X$, $\mathcal{F}_t^X = \sigma(X_s, s \leq t)$. Then $X$ is automatically $\{\mathcal{F}_t^X\}_{t \in \mathbb{R}^+}$-adapted.

**Definition 1.4.4.**

i) A stochastic process $(X_t)_{t \in \mathbb{R}^+}$ on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{R}^+}, \mathbb{P})$ with values in $\mathbb{R}^d$ is called progressive or progressively measurable if for all $t > 0$ the map

$$(s, \omega) \mapsto X_s(\omega)$$

from $([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t)$ into $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is measurable.

* A family $\{\mathcal{F}_t\}_{t \in \mathbb{R}^+}$ of $\sigma$-algebras is increasing, if $\mathcal{F}_s \subseteq \mathcal{F}_t \forall s \leq t$. 

\qed
ii) We define
\[ \mathcal{P}_r := \{ A \subset \mathbb{R}_+ \times \Omega, (t, \omega) \mapsto \mathbb{1}_A(t, \omega) \text{ is progressive} \}. \]

Then \( \mathcal{P}_r \) is a \( \sigma \)-algebra on \( \mathbb{R}_+ \times \Omega \) called the \( \sigma \)-algebra of progressive sets.

Clearly a progressive process is also measurable and adapted, and one can show that a process is progressive if and only if it is measurable with respect to \( \mathcal{P}_r \).

**Proposition 1.4.5.** Let \( (X_t)_{t \in \mathbb{R}_+} \) be an adapted process on \( (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P}) \) with values in \( \mathbb{R}^d \). If \( (X_t)_{t \in \mathbb{R}_+} \) is left-continuous (respectively right-continuous), then it is progressive.

**Proof.** For the proof we refer to Proposition 3.4.4 of Priouret [15].

We may extend our notion of the Brownian motion as introduced in Definition 1.1.10 in the following way.

**Definition 1.4.6.** Let \( (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P}) \) be a filtered probability space. An \( (\mathcal{F}_t)_{t \in \mathbb{R}_+} \)-adapted process \( B = (B_t)_{t \in \mathbb{R}_+} \) is a standard \( (\mathcal{F}_t)_{t \in \mathbb{R}_+} \)-Brownian motion with values in \( \mathbb{R}^d \) if

i) \( B_0 = 0 \) a.s.,

ii) for every \( s < t \), \( B_t - B_s \) is independent of \( \mathcal{F}_s \),

iii) for every \( s < t \), \( B_t - B_s \sim \mathcal{N}(0, (t-s)I_d) \),

iv) \( B \) has continuous sample paths a.s..

Given a Brownian motion \( B = (B_t)_{t \in \mathbb{R}_+} \) on \( (\Omega, \mathcal{F}, \mathbb{P}) \) we can always associate to \( B \) a complete filtration such that \( B \) is adapted to this filtration. Consider \( \mathcal{F}^0_{\tau} = \sigma(B_s, s \leq \tau) \), \( \mathcal{F}^0_{\infty} = \sigma(B_t, t \geq 0) \) and let \( \mathcal{N} \) be the family of all events in \( \mathcal{F}^0_{\infty} \) with measure zero. We define

\[ \mathcal{F}^B_t = \sigma(\mathcal{F}^0_t, \mathcal{N}), \quad t \geq 0. \]

Then \( (\mathcal{F}^B_t)_{t \in \mathbb{R}_+} \) is called the natural filtration of the Brownian motion. One can prove that for all \( t \geq 0 \), \( \mathcal{F}^B_t = \mathcal{F}^B_{t+} \).

### 1.5 Martingales

For the rest of this chapter note that we assume given a filtered complete probability space \( (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P}) \), where the filtration satisfies the usual hypotheses. For the definition and properties of the conditional expectation, we refer to Appendix B.

**Definition 1.5.1.** A real-valued adapted stochastic process \( X = (X_t)_{t \in \mathbb{R}_+} \) is a (sub-, super-)martingale if

i) for every \( t \in \mathbb{R}_+ \), \( X_t \) is integrable,
ii) for every \( s < t \)
\[
E[X_t|\mathcal{F}_s] = X_s \quad \text{a.s.}
\]
(respectively \( E[X_t|\mathcal{F}_s] \geq X_s \), or \( E[X_t|\mathcal{F}_s] \leq X_s \) a.s.)

Condition ii) of Definition 1.5.1 is often referred to as the martingale condition. If we replace the integrability condition by the assumption that \( X \) is positive, we obtain the concept of generalized positive (sub-, super-)martingales.

Please take note of the following simple remarks:
- If \( X \) is a supermartingale, then \(-X\) is a submartingale.
- If \( X \) is a supermartingale, then \( E[X_t] \) is decreasing.
- If \( X \) is a submartingale, then \( E[X_t] \) is increasing.
- If \( X \) is a martingale (resp. submartingale) and \( f: \mathbb{R} \to \mathbb{R} \) is a convex (resp. increasing convex) map such that \( f(X_t) \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \) for every \( t \), then \( f(X) \) is a submartingale.

**Proposition 1.5.2.** Let \( B = (B_t)_{t \in \mathbb{R}^+} \) be a real-valued \( (\mathcal{F}_t)_{t \in \mathbb{R}^+} \)-Brownian motion. Then

i) \( B \),
ii) \( Y = (Y_t)_{t \in \mathbb{R}^+}, Y_t = B_t^2 - t \),
iii) for every \( \lambda \in \mathbb{R} \), \( Z^\lambda = (Z^\lambda_t)_{t \in \mathbb{R}^+}, Z^\lambda_t = \exp(\lambda B_t - \frac{1}{2} \lambda^2 t) \)

are \((\mathcal{F}_t)_{t \in \mathbb{R}^+}\)-martingales.

**Proof.** i) We know that for every \( t \), \( B_t \) is integrable. Since for every \( s < t \) \( B_t - B_s \) is independent of \( \mathcal{F}_s \), we also obtain by Proposition B.0.4 that
\[
E[B_t|\mathcal{F}_s] = E[B_t - B_s + B_s|\mathcal{F}_s] = E[B_t - B_s|\mathcal{F}_s] + B_s = E[B_t - B_s] + B_s = B_s.
\]

ii) We note that \( Y \) is clearly adapted and integrable. The martingale condition follows from
\[
t - s = E[(B_t - B_s)^2] = E[(B_t - B_s)^2|\mathcal{F}_s] = E[B_t^2|\mathcal{F}_s] - 2B_s E[B_t|\mathcal{F}_s] + B_s^2 = E[B_t^2|\mathcal{F}_s] - B_s^2,
\]
which implies that
\[
E[B_t^2 - t|\mathcal{F}_s] = B_s^2 - s.
\]

iii) Finally
\[
E\left[e^{\lambda(B_t - B_s)}\right|\mathcal{F}_s] = E\left[e^{\lambda(B_t - B_s)}\right] = e^{\lambda B_s^2/2}(t-s).
\]

\(\square\)
Definition 1.5.3. We say that a submartingale \((X_t)_{t \in \mathbb{R}_+}\) is closed by \(X_\infty\) if there exists a random variable \(X_\infty\) that is \(\mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \geq 0)\)-measurable and integrable such that for every \(t\)

\[
X_t \leq E[X_\infty | \mathcal{F}_t] \quad \text{a.s.}
\]

Note that if there exists \(X \in L^1(\Omega, \mathcal{F}, \mathbb{P})\) such that \(X_t \leq E[X | \mathcal{F}_t]\), then \(X_t\) is closed by \(X_\infty = E[X | \mathcal{F}_\infty]\).

1.6 Stopping Theorems

Definition 1.6.1. i) A map \(\tau : \Omega \to \mathbb{R}_+ \cup \{\infty\}\) is called a stopping time if for every \(t \in \mathbb{R}_+\)

\[
\{\tau \leq t\} \in \mathcal{F}_t.
\]

ii) Every stopping time \(\tau\) is associated with a particular \(\sigma\)-algebra, the stopping time \(\sigma\)-algebra

\[
\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty \mid A \cap \{\tau \leq t\} \in \mathcal{F}_t, \ t \geq 0\},
\]

where as before we define \(\mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \geq 0)\).

iii) Let \(X = (X_t)_{t \in \mathbb{R}_+}\) be an adapted process, \(X_\infty\) a given \(\mathcal{F}_\infty\)-measurable random variable (for closed submartingale we have a canonical candidate) and \(\tau\) a stopping time. Then the position (value) of \(X\) at \(\tau\) is defined as

\[
X_\tau(\omega) := X_{\tau(\omega)}(\omega).
\]

iv) If \((X_t)_{t \in \mathbb{R}_+}\) is an adapted process and \(\tau\) a stopping time, we define the stopped process \(X^\tau = (X^\tau_t)_{t \in \mathbb{R}_+}\) as

\[
X^\tau_t(\omega) = X_{t \wedge \tau}(\omega) = \begin{cases} X_t(\omega), & t < \tau(\omega), \\ X_{\tau(\omega)}(\omega), & t \geq \tau(\omega). \end{cases}
\]

Remark 1.6.2. • A deterministic time \(T \in \mathbb{R}\) is a stopping time.

• For two stopping times \(\sigma\) and \(\tau\), \(\sigma \wedge \tau\) and \(\sigma \vee \tau\) are again stopping times.

• If \(\sigma, \tau\) are stopping times such that \(\sigma \leq \tau\) a.s., then \(\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau\).

• If \((X_t)_{t \in \mathbb{R}_+}\) is a progressively measurable process, then \(X_\tau\) is \(\mathcal{F}_\tau\)-measurable for every stopping time \(\tau\).

Example 1.6.3. a) Let \((X_t)_{t \in \mathbb{R}_+}\) be a right-continuous adapted process with values in \(\mathbb{R}^d\) and assume the given filtration is standard (i.e. right-continuous and complete). If \(B\) is in \(\mathcal{B}(\mathbb{R}^d)\), then the first entry time \(\tau_B(\omega) = \inf \{t \geq 0 : X_t(\omega) \in B\}\) is a stopping time.

b) Let \((X_t)_{t \in \mathbb{R}_+}\) be a continuous adapted process with values in \(\mathbb{R}^d\). If \(B \in \mathcal{B}(\mathbb{R}^d)\) is closed, then \(\tau_B(\omega) = \inf \{t \geq 0 : X_t(\omega) \in B\}\) is a stopping time (without requiring the filtration to be standard).

In particular, if for \(n \in \mathbb{N}\) we define \(\tau_n(\omega) := \inf \{t \geq 0 : X_t(\omega) \geq n\}\), then \((\tau_n)_{n \in \mathbb{N}}\) is a sequence of stopping times such that \(\tau_n \nearrow +\infty\) for \(n \to \infty\).
Proof. For a proof of a) we refer to Dellacherie and Meyer [4]. We leave b) as an exercise.

We now consider some well-known results highlighting important relations between martingales and stopping times. To keep things easy, we start with the more simple discrete-time case.

**Proposition 1.6.4.** Let $X = (X_n)_{n \in \mathbb{N}}$ be a submartingale and $\nu$ a stopping time. Then $X^\nu = (X^\nu_n)_{n \in \mathbb{N}}$ is also a submartingale.

**Proof.** We can check immediately that $X^\nu$ is integrable, since $|X_n \wedge \nu| \leq |X_0| + \cdots + |X_n|$.

Moreover, we have

$$E[X^\nu_{n+1} - X^\nu_n \mid F_n] = E[\mathbb{1}_{\{\nu > n\}} (X_{n+1} - X_n) \mid F_n]$$

$$= \mathbb{1}_{\{\nu > n\}} E[(X_{n+1} - X_n) \mid F_n] \geq 0 \text{ a.s.}$$

$$\blacksquare$$

**Theorem 1.6.5 (Doob’s optional sampling theorem).** Let $X = (X_n)_{n \in \mathbb{N}}$ be a submartingale and $\tau_1, \tau_2$ two bounded stopping times with $\tau_1 \leq \tau_2$ a.s. Then

$$E[X_{\tau_2} \mid F_{\tau_1}] \geq X_{\tau_1} \text{ a.s.}$$

**Proof.** Let $m \in \mathbb{N}$ be such that $|X_{\tau_2}| \leq |X_0| + \cdots + |X_m| \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and let $A \in \mathcal{F}_{\tau_1}$. Then we have

$$\int_A X_{\tau_2} \, d\mathbb{P} = \sum_{k=0}^m \int_{A \cap \{\tau_1 = k\}} X_{\tau_2 \wedge m} \, d\mathbb{P} = \sum_{k=0}^m \int_{A \cap \{\tau_1 = k\}} E[X_{\tau_2 \wedge m} \mid \mathcal{F}_k] \, d\mathbb{P}$$

$$\geq \sum_{k=0}^m \int_{A \cap \{\tau_1 = k\}} X_{\tau_2 \wedge k} \, d\mathbb{P} = \sum_{k=0}^m \int_{A \cap \{\tau_1 = k\}} X_k \, d\mathbb{P} = \int_A X_{\tau_1} \, d\mathbb{P},$$

where we have used that on $A \cap \{\tau_1 = k\} \in \mathcal{F}_k$, we have $X_{\tau_2 \wedge k} = X_k$. $\blacksquare$

We immediately obtain the following corollaries.

**Corollary 1.6.6.** Let $X = (X_n)_{n \in \mathbb{N}}$ be a submartingale closed by $X_\infty$ and $\tau_1, \tau_2$ two stopping times with values in $[0, m] \cup \{\infty\}$, $m \in \mathbb{N}$ and with $\tau_1 \leq \tau_2$ a.s. Then

$$E[X_{\tau_2} \mid F_{\tau_1}] \geq X_{\tau_1} \text{ a.s.}$$

**Proof.** The corollary follows from Theorem 1.6.5 applied to $Y = (Y_n)_{n \in \mathbb{N}}$, where $Y_n = X_n$ if $n \leq m$, and $Y_n = X_\infty$ if $n > m$. $\blacksquare$

**Corollary 1.6.7.** Let $X = (X_n)_{n \in \mathbb{N}}$ be a submartingale and $\tau$ a bounded stopping time. Then

$$E[X_{\tau}] \geq E[X_0].$$
We now extend these results to the continuous-time case.

**Theorem 1.6.8 (Doob’s optional sampling theorem).** Let \((X_t)_{t \in \mathbb{R}_+}\) be a right-continuous submartingale closed by \(X_\infty\) and \(\sigma, \tau\) two stopping times such that \(\sigma \leq \tau\) a.s. Then \(X_\tau \in L^1(\Omega, \mathcal{F}, \mathbb{P})\) and

\[
E[X_\tau | \mathcal{F}_\sigma] \geq X_\sigma \quad \text{a.s.}
\]

**Proof.** For the proof we refer to Dellacherie and Meyer [5].

**Corollary 1.6.9.** Let \((X_t)_{t \in \mathbb{R}_+}\) be a right-continuous submartingale and \(\sigma \leq \tau\) two bounded stopping times. Then \(X_\tau \in L^1(\Omega, \mathcal{F}, \mathbb{P})\) and

\[
E[X_\tau | \mathcal{F}_\sigma] \geq X_\sigma \quad \text{a.s.}
\]

**Proof.** Let \(\sigma \leq \tau \leq T\) for some \(T > 0\). It is sufficient to apply Theorem 1.6.8 to \(X^T = (X^T_t)_{t \in \mathbb{R}_+} = (X_{t \wedge T})_{t \in \mathbb{R}_+}\), which is closed by \(X_T\).

**Corollary 1.6.10.** Let \((X_t)_{t \in \mathbb{R}_+}\) be a right-continuous submartingale and \(\sigma, \tau\) two bounded stopping times. Then

\[
E[X_\tau | \mathcal{F}_\sigma] \geq X_{\sigma \wedge \tau} \quad \text{a.s.}
\]

**Proof.**

\[
E[X_\tau | \mathcal{F}_\sigma] = E[1_{\{\tau \leq \sigma\}} X_{\sigma \wedge \tau} + 1_{\{\tau > \sigma\}} X_{\sigma \vee \tau} | \mathcal{F}_\sigma]
\]

\[
= 1_{\{\tau \leq \sigma\}} X_{\sigma \wedge \tau} + 1_{\{\tau > \sigma\}} E[X_{\sigma \vee \tau} | \mathcal{F}_\sigma]
\]

\[
\geq 1_{\{\tau \leq \sigma\}} X_{\sigma \wedge \tau} + 1_{\{\tau > \sigma\}} X_\sigma = X_{\sigma \wedge \tau} \quad \text{a.s.}
\]

**Corollary 1.6.11.** Let \((X_t)_{t \in \mathbb{R}_+}\) be a right-continuous submartingale and \(\tau\) a stopping time. Then \(X^\tau := (X_{t \wedge \tau})_{t \in \mathbb{R}_+}\) is a submartingale.

**Proof.** For \(s < t\), we have

\[
E[X_{t \wedge \tau} | \mathcal{F}_s] \geq X_{s \wedge \tau} \quad \text{a.s.}
\]

by Corollary 1.6.10.

**1.7 Martingale Inequalities**

We now consider some important martingale inequalities. As in the previous section, we again start by considering the discrete-time case.

**Proposition 1.7.1.** i) Let \((X_n)_{n \in \mathbb{N}}\) be a positive supermartingale. Then for every \(a > 0\)

\[
a \mathbb{P}\left(\sup_n X_n \geq a\right) \leq E[X_0].
\]
ii) Let \((X_n)_{n \in \mathbb{N}}\) be a positive submartingale. Then for every \(a > 0\) and \(n \in \mathbb{N}\) we have

\[
a \mathbb{P}\left( \max_{k \leq n} X_k \geq a \right) \leq \int_{\{\max_{k \leq n} X_k \geq a\}} X_n \, d\mathbb{P} \leq E[X_n].
\]

**Proof.** We start with a preliminary remark. Consider the stopping time

\[
\tau_a = \inf \{ k \geq 0 \mid X_k \geq a \}.
\]

Then \(\{\max_{k \leq n} X_k \geq a\} = \{\tau_a \leq n\}\) and on \(\{\tau_a \leq n\}\) we have

\[X_{\tau_a} \geq a.\]

As \(X_n \geq 0\) a.s. for all \(n \in \mathbb{N}\) it follows that

\[
a \mathbb{1}_{\{\tau_a \leq n\}} \leq X_{\tau_a} \mathbb{1}_{\{\tau_a \leq n\}} \leq X_{\tau_a} \mathbb{1}_{\{\tau_a \leq n\}} + \underbrace{X_n \mathbb{1}_{\{\tau_a > n\}}}_{\geq 0} = X_{\tau_a}.
\]

(1.3)

i) Suppose now that \((X_n)_{n \in \mathbb{N}}\) is a supermartingale. By (1.3) we obtain that

\[
a \mathbb{P}(\tau_a \leq n) \leq E[X_{\tau_a \wedge n}] \leq E[X_0].
\]

But since for all \(\varepsilon > 0\)

\[
\left\{ \sup_{n \in \mathbb{N}} X_n \geq a \right\} \subseteq \lim_{n \to \infty} \{\tau_{a-\varepsilon} \leq n\}
\]

we get

\[
(a - \varepsilon) \mathbb{P}(\sup_{n \in \mathbb{N}} X_n \geq a) \leq (a - \varepsilon) \lim_{n \to \infty} \mathbb{P}(\tau_{a-\varepsilon} \leq n) \leq E[X_0].
\]

For \(\varepsilon \to 0\) the statement is proved.

ii) If \((X_n)_{n \in \mathbb{N}}\) is a submartingale, we have that

\[
a \mathbb{P}(\tau_a \leq n) \leq E[X_{\tau_a \mathbb{1}_{\{\tau_a \leq n\}}}] = \sum_{k=0}^{n} \int_{\{\tau_a = k\}} X_k \, d\mathbb{P}
\]

\[
\leq \sum_{k=0}^{n} \int_{\{\tau_a = k\}} X_n \, d\mathbb{P} = \int_{\{\tau_a \leq n\}} X_n \, d\mathbb{P}
\]

\[
\leq E[X_n].
\]

\(\square\)

We are ready to prove the following.
Theorem 1.7.2 (Doob’s maximal inequalities). Let \((X_n)_{n \in \mathbb{N}}\) be a martingale (or a positive submartingale) belonging to \(L^p(\Omega, \mathcal{F}, \mathbb{P})\), \(1 < p < \infty\). Then

\[
\left\| \max_{k \leq m} |X_k| \right\|_p = E \left[ \left( \max_{k \leq m} |X_k| \right)^p \right]^{\frac{1}{p}} \leq \frac{p}{p-1} \left\| X_m \right\|_p, \tag{1.4}
\]

\[
\left\| \sup_n |X_n| \right\|_p \leq \frac{p}{p-1} \sup_n \left\| X_n \right\|_p. \tag{1.5}
\]

Proof. Recall that if \((X_n)_{n \in \mathbb{N}}\) is a martingale, then \((|X_n|)_{n \in \mathbb{N}}\) is a positive submartingale. Therefore we now prove the statement assuming \((X_n)_{n \in \mathbb{N}}\) is a positive submartingale. Note that in this case

\[
\left\| X_n \right\|_p = E \left[ |X_n|^p \right]^{\frac{1}{p}}
\]

is increasing in \(n\). We define \(Y_m = \max_{k \leq m} X_k\). If \(\left\| Y_m \right\|_p = 0\), the statement is clear. Therefore in the following we suppose that \(\left\| Y_m \right\|_p > 0\). By Proposition 1.7.1 we have

\[
a \mathbb{P}\left(Y_m \geq a\right) \leq E \left[ X_m 1_{\{Y_m \geq a\}} \right].
\]

Now we multiply both terms by \(pa^{p-2}\) and integrate in \(da\)

\[
\int_0^\infty pa^{p-2} E \left[ 1_{\{Y_m \geq a\}} \right] da \leq p \int_0^\infty a^{p-2} E \left[ X_m 1_{\{Y_m \geq a\}} \right] da. \tag{a}
\]

By applying the Fubini-Tonelli theorem to (a) we obtain

\[
\int_0^\infty pa^{p-1} E \left[ 1_{\{Y_m \geq a\}} \right] da = E \left[ \int_0^\infty pa^{p-1} 1_{\{Y_m \geq a\}} da \right] = E \left[ Y_m \right]^{p-1} = E \left[ Y_m^p \right],
\]

and analogously for (b)

\[
\int_0^\infty pa^{p-2} E \left[ X_m 1_{\{Y_m \geq a\}} \right] da = E \left[ X_m \int_0^\infty pa^{p-2} 1_{\{Y_m \geq a\}} da \right] = \frac{p}{p-1} E \left[ X_m Y_m^{p-1} \right].
\]

Finally we make use of Hölder’s inequality which gives us

\[
E \left[ X_m Y_m^{p-1} \right] \leq E \left[ |X_m|^p \right]^{\frac{1}{p}} \cdot E \left[ (Y_m^{p-1})^{\frac{p}{p-1}} \right]^{\frac{p-1}{p}}
\]

\[
\left( \frac{1}{q} + \frac{1}{p} = 1 \Rightarrow q = \frac{p}{p-1} \right)
\]
Putting everything together, we get
\[ \|Y_m\|_p \leq \frac{p}{p-1} \|X_m\|_p \cdot \|Y_m\|_p^{p-1}, \]
and since \(0 < \|Y_m\|_p < \infty\) we can divide by \(\|Y_m\|_p^{p-1}\) to obtain (1.4). Note that (1.5) follows directly as
\[ Y_m = \max_{k \leq m} X_k \] \(\sup_n X_n\) a.s.

We extend this result to the case of (sub-, super-)martingales in continuous time. Note first that if \((X_t)_{t \in \mathbb{R}}\) is a right-continuous process, we have that
\[ \sup_{0 \leq t \leq T} |X_t| = \lim_{n \to \infty} \sup_{0 \leq k \leq 2^n} \left| \frac{X_{kT}}{2^n} \right|, \]
where the second term converges monotonously. Therefore the next Proposition 1.7.3 and Theorem 1.7.4 follow as a direct consequence of Proposition 1.7.1 and Theorem 1.7.2.

**Proposition 1.7.3.** i) Let \((X_t)_{t \in \mathbb{R}_+}\) be a right-continuous positive supermartingale. Then for every \(a > 0\) we have
\[ \mathbb{P} \left( \sup_{t \geq 0} X_t \geq a \right) \leq \frac{1}{a} E[X_0]. \]

ii) Let \((X_t)_{t \in \mathbb{R}_+}\) be a right-continuous positive submartingale. Then for every \(a > 0\) and \(T > 0\)
\[ \mathbb{P} \left( \sup_{0 \leq t \leq T} X_t \geq a \right) \leq \frac{1}{a} E[X_T]. \]

We state the continuous-time analogon of Theorem 1.7.2.

**Theorem 1.7.4 (Doob’s maximal inequalities).** Let \((X_t)_{t \in \mathbb{R}_+}\) be a right-continuous martingale (or a positive submartingale). Then for \(1 < p < \infty\)
\[ \left\| \sup_t |X_t| \right\|_p \leq \frac{p}{p-1} \sup_t \|X_t\|_p. \]

Sometimes we may want to work with finite intervals, i.e. \([0, T]\) with \(T > 0\). In this case we apply Theorem 1.7.4 to the stopped martingale \((X^T_t)_{t \in \mathbb{R}_+} = (X_{t \wedge T})_{t \in \mathbb{R}_+}\) and obtain
\[ \left\| \sup_{t \in [0, T]} |X_t| \right\|_p \leq \frac{p}{p-1} \|X_T\|_p. \]
1.8 Martingale Convergence

Lemma 1.8.1. Let \((Y_t)_{t \in \mathbb{R}_+}\) be a right-continuous adapted process. Define 
\[ Y := \limsup_{t \to \infty} Y_t \text{ and } Y := \liminf_{t \to \infty} Y_t \] 
and suppose that 
\[ E \left[ \sup_t |Y_t| \right] < \infty. \]
Then there exists an increasing sequence \((\tau_n)_{n \in \mathbb{N}}\) of bounded stopping times such that 
\[ Y_{\tau_{2n}} \xrightarrow{n \to \infty} Y, \quad Y_{\tau_{2n+1}} \xrightarrow{n \to \infty} Y \] 
a.s. and in \(L^1(\Omega, \mathcal{F}, P)\).

Proof. For the proof we refer to Priouret [15], Lemma 4.5.9.

Theorem 1.8.2 (Martingale convergence theorem). Let \(X = (X_t)_{t \in \mathbb{R}_+}\) be a right-continuous (sub-)martingale such that 
\[ E \left[ \sup_t |X_t| \right] < \infty. \] 
Then \(X\) converges a.s. and in \(L^1(\Omega, \mathcal{F}, P)\) to a random variable \(X_\infty \in L^1(\Omega, \mathcal{F}, P)\) and for every stopping time \(\tau\) we have
\[ X_\tau \leq E[X_\infty | \mathcal{F}_\tau] \quad \text{a.s.} \]

Proof. First suppose \(X\) is a submartingale. Let \(\bar{X} = \limsup_{t \to \infty} X_t, \quad \underline{X} = \liminf_{t \to \infty} X_t.\) Then by Lemma 1.8.1 there exists a sequence \((\tau_n)_{n \in \mathbb{N}}\) of bounded stopping times such that 
\[ E[X_{\tau_{2n}}] \xrightarrow{n \to \infty} E[\bar{X}] \quad \text{and} \quad E[X_{\tau_{2n+1}}] \xrightarrow{n \to \infty} E[\underline{X}]. \]
By Doob’s optional sampling theorem (Theorem 1.6.8) we obtain that \(E[X_{\tau_n}]\) is an increasing sequence. Since \(E[\sup_t |X_t|] < \infty, \quad (E[X_{\tau_n}]_{n \in \mathbb{N}}\) is also bounded. Then \((E[X_{\tau_n}]_{n \in \mathbb{N}}\) is convergent and 
\[ E[\underline{X}] = E[\bar{X}]. \]
Since \(\underline{X} \leq \bar{X}\), it follows that \(\underline{X} = \bar{X} = X_\infty\) a.s. We then conclude that 
\(X_t \xrightarrow{t \to \infty} X_\infty\) a.s. and, since \(\sup_t |X_t| \in L^1(\Omega, \mathcal{F}, P)\), also in \(L^1\). Hence 
\[ X_t \leq E[X_\infty | \mathcal{F}_t] \quad \text{a.s.} \]
(prove it by applying Fatou’s Lemma to the limsup!) and by Doob’s optional sampling theorem in continuous time (Theorem 1.6.8) it follows that for every stopping time \(\tau\) 
\[ X_\tau \leq E[X_\infty | \mathcal{F}_\tau] \quad \text{a.s.} \]
If \(X\) is a martingale, we consider the submartingales \(X\) and \(-X\). \(\square\)
Theorem 1.8.3. Let \( X = (X_t)_{t \in \mathbb{R}_+} \) be a generalized positive right-continuous supermartingale. Then \( X_t \xrightarrow{t \to \infty} X_\infty \) a.s. in \( \mathbb{R}_+ \) and for all \( t \geq 0 \), \( \{X_\infty < \infty\} \subset \{X_t < \infty\} \). In addition, for all stopping times \( \sigma, \tau \) such that \( \sigma \leq \tau \) a.s.,

\[
E[X_\tau | \mathcal{F}_\sigma] \leq X_\sigma \quad \text{a.s.}
\]

Proof. For the proof we refer to Priouret [15].

Corollary 1.8.4. Let \( (X_t)_{t \in \mathbb{R}_+} \) be a right-continuous martingale such that \( \sup_t E[|X_t|^p] < \infty \) for \( 1 < p < \infty \). Then

\[
\lim_{t \to \infty} X_t =: X_\infty \exists \text{ a.s. and in } L^p \\
\text{and for every stopping time } \tau
\]

\[
X_\tau = E[X_\infty | \mathcal{F}_\tau] \quad \text{a.s.}
\]

Proof. The proof follows by Theorem 1.7.4 (Doob’s maximal inequalities) and Proposition 1.8.2 applied to the submartingale \( Z_t = |X_t|^p \), \( t \geq 0 \).

1.9 Square Integrable Martingales

Let \( M = (M_t)_{t \in \mathbb{R}_+} \) be a square integrable martingale, i.e. \( E[M_t^2] < \infty \) for all \( t \geq 0 \). We note that \( M \) has the following useful properties. For \( s < t \) we have

\[
E[(M_t - M_s)^2 | \mathcal{F}_s] = E[M_t^2 | \mathcal{F}_s] - 2M_sE[M_t | \mathcal{F}_s] + M_s^2
\]

\[
= E[M_t^2 - M_s^2 | \mathcal{F}_s],
\]

furthermore \( M \) has orthogonal increments, i.e. for \( u < v \leq s < t \)

\[
E[(M_v - M_u)(M_t - M_s)] = E[E[(M_v - M_u)(M_t - M_s) | \mathcal{F}_s]]
\]

\[
= E \left[ (M_v - M_u) E[M_t - M_s | \mathcal{F}_s] \right] = 0.
\]

This implies that if \( 0 = t_0 < t_1 < \cdots < t_n = t \) is a partition of \( [0, t] \), then

\[
E[(M_t - M_0)^2] = \sum_{i=0}^{n-1} E[(M_{t_{i+1}} - M_{t_i})^2].
\]

The following family of square integrable continuous martingales

\[
\mathbb{H}^2_C = \left\{ M : M \text{ is a continuous martingale, } M_0 = 0, \sup_{t \in \mathbb{R}_+} E[M_t^2] < \infty \right\}
\]

turns out to be a particularly convenient class of processes, e.g. when working with stochastic integrals. The next proposition characterizes the space \( \mathbb{H}^2_C \) and typical properties of the processes in \( \mathbb{H}^2_C \).
Proposition 1.9.1. i) Let $M \in \mathbb{H}^2_C$. Then

$$\lim_{t \to \infty} M_t =: M_\infty$$

exists a.s. and in $L^2$ and we have that $M$ is closed by $M_\infty$, i.e. $M_t = E \left[ M_\infty | \mathcal{F}_t \right]$ a.s. for all $t \in \mathbb{R}^+$.  

ii) Let $(M^n)_{n \in \mathbb{N}} \subset \mathbb{H}^2_C$ be a sequence of processes such that

$$M^n_\infty \xrightarrow{n \to \infty} X$$

in $L^2$. Then there exists $M \in \mathbb{H}^2_C$ such that

$$X = M_\infty \ a.s.$$ 

and for every $t$, $M^n_t \xrightarrow{n \to \infty} M_t$ in $L^2$. 

iii) The space $\mathbb{H}^2_C$ endowed with the scalar product

$$\langle M, N \rangle_{\mathbb{H}^2_C} := E \left[ M_\infty N_\infty \right]$$

is a Hilbert space.

Proof. i) The statement follows directly by Corollary 1.8.4.

ii) Consider $X_t = E \left[ X | \mathcal{F}_t \right]$. Then by Jensen’s inequality

$$E \left[ X^2_t \right] = E \left[ (E \left[ X | \mathcal{F}_t \right])^2 \right] \leq E \left[ X^2 \right], \quad (1.6)$$

i.e.

$$\sup_t E \left[ X^2_t \right] < \infty.$$ 

First we prove that $M^n_t \xrightarrow{n \to \infty} X_t$ in $L^2$. By i) we have $M^n_t = E \left[ M^n_\infty | \mathcal{F}_t \right]$ a.s. for all $n \in \mathbb{N}$. Then by using (1.6) it follows that

$$E \left[ (M^n_t - X_t)^2 \right] \leq E \left[ (M^n_\infty - X)^2 \right], \quad (1.7)$$

hence

$$M^n_t \xrightarrow{n \to \infty} X_t \ \text{in} \ L^2. \quad (1.8)$$

We now prove that $(X_t)_{t \in \mathbb{R}^+}$ has a continuous modification. Let $n_k \xrightarrow{k \to \infty} \infty$ be such that

$$\|M^n_{\infty} - X\|_2 < 2^{-k}.$$ 

By using the Doob’s maximal inequality (Theorem 1.7.4), we obtain

$$E \left[ \sum_k \sup_t \left| M^{n_k}_t - M^{n_{k+1}}_t \right| \right] \leq \sum_k E \left[ \sup_t \left| M^{n_k}_t - M^{n_{k+1}}_t \right| \right] \leq \sum_k E \left[ \left( \sup_t \left| M^{n_k}_t - M^{n_{k+1}}_t \right| \right)^2 \right]^{1/2}$$

\[ \leq \sum_k \text{H"older} \left[ \left( \sup_t \left| M^{n_k}_t - M^{n_{k+1}}_t \right| \right)^2 \right]^{1/2} \]
\[
\sum_k \sup_t \|M_{nk} - M_{nk+1}^t\|_2 \\
\leq 2 \sum_k \|M_{nk}^t - M_{nk+1}^t\|_2 \\
\leq 2 \sum_k \|M_{nk} - M_{nk+1}^\infty\|_2 \\
\leq 2 \sum_k \|M_{nk}^\infty - X\|_2 + \|M_{nk+1}^\infty - X\|_2 < \infty.
\]

Hence \(\sum_k \sup_t |M_{nk}^t - M_{nk+1}^t| < \infty\) a.s. and \(M_{nk}^t\) converges uniformly a.s. on \(\mathbb{R}_+\). Then there exists a continuous process \(M = (M_t)_{t \in \mathbb{R}_+} \in \mathbb{H}_C^2\) such that

\[
M_{nk}^t \xrightarrow{k \to \infty} M_t \quad \text{a.s. and in } L^2.
\]

By (1.8) it follows that \(X_t = M_t\) a.s. and therefore \(M_{nk}^t \xrightarrow{n \to \infty} M_t\) in \(L^2\).

By i) \(M\) is closed by an \(\mathcal{F}_\infty\)-measurable random variable \(M_\infty\). Since

\[
E[M_\infty | \mathcal{F}_t] = M_t = X_t = E[X | \mathcal{F}_t] \quad \text{a.s.}
\]

for all \(t \geq 0\), we have that for all \(A \in \mathcal{F}_t\),

\[
\int_A M_{nk}^\infty dP = \int_A X dP.
\]

This also holds for \(A \in \mathcal{F}_\infty = \sigma(\mathcal{F}_t, t \geq 0)\), i.e. \(M_\infty = E[M_\infty | \mathcal{F}_\infty] = E[X | \mathcal{F}_\infty] \quad \text{a.s.}\)

Since \(X\) is also \(\mathcal{F}_\infty\)-measurable we obtain \(M_\infty = X\) a.s.

iii) \(\langle M, N \rangle_{\mathbb{H}_C^2} = E[M_\infty N_\infty]\) is a scalar product (prove it!) and therefore \((\mathbb{H}_C^2, \langle \cdot, \cdot \rangle_{\mathbb{H}_C^2})\) is a Hilbert space since it is complete, as we show now. Suppose \((M^n)_{n \in \mathbb{N}}\) is a Cauchy sequence in \(\mathbb{H}_C^2\). Then \((M^n_{nk})_{n \in \mathbb{N}}\) is a Cauchy sequence in \(L^2(\Omega, \mathcal{F}, \mathbb{P})\) and there exists \(X \in L^2(\Omega, \mathcal{F}, \mathbb{P})\) such that

\[
\|M^n_{nk} - X\|_2 \xrightarrow{n \to \infty} 0.
\]

To conclude, it is sufficient to apply ii).

\[\square\]

### 1.10 Local Martingales and the Sharp Bracket

**Definition 1.10.1.** A real-valued process \(X = (X_t)_{t \in \mathbb{R}_+}\) is a local martingale, if there exists an increasing sequence of stopping times \((\tau_n)_{n \in \mathbb{N}}, \tau_n \nearrow \infty\) a.s., such that \(X^{\tau_n} = (X_t \wedge \tau_n)_{t \in \mathbb{R}_+}\) is a martingale for all \(n \in \mathbb{N}\). We say such a sequence of stopping times reduces the local martingale \(X\).

**Lemma 1.10.2.** Let \(X\) be a right-continuous, positive local martingale. Then \(X\) is a generalized positive supermartingale.

**Proof.** For every \(s < t\):

\[
E[X_t | \mathcal{F}_s] = E\left[\lim_{n \to \infty} X_{t \wedge \tau_n} | \mathcal{F}_s\right] \leq \liminf_{n \to \infty} E\left[X_{t \wedge \tau_n} | \mathcal{F}_s\right] = X_s.
\]

\[\square\]
Proposition 1.10.3. Let $X$ be a right-continuous local martingale. If
\[ E \left[ \sup_{t \in \mathbb{R}_+} |X_t| \right] < \infty \]
then $X$ is a martingale.

Proof. For all $A \in \mathcal{F}_s$, $s < t$, we have $E[1_A X_{s \wedge \tau_n}] = E[1_A X_{t \wedge \tau_n}]$, with $(\tau_n)_{n \in \mathbb{N}}$ reduces $X$. Since $|X_{s \wedge \tau_n}|, |X_{t \wedge \tau_n}| \leq \sup |X_t| \in L^1$ for all $n \in \mathbb{N}$, we get Lebesgue’s dominated convergence theorem
\[ E[1_A X_t] = E \left[ \lim_{n \to \infty} 1_A X_{s \wedge \tau_n} \right] = \lim_{n \to \infty} E[1_A X_{s \wedge \tau_n}] = \lim_{n \to \infty} E[1_A X_s] \]
and it follows that $X$ is a martingale. $\square$

Remark 1.10.4. i) Under the assumptions of Proposition 1.10.3 we have that $X$ converges a.s. and in $L^1$ to a random variable $X_\infty \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and for every stopping time $\tau$ we have $X_\tau = E[X_\infty | \mathcal{F}_\tau]$ a.s. by Theorem 1.8.2.

ii) Applying Proposition 1.10.3 to the locale martingale $X_t \wedge T, T > 0$, we obtain:
\[ E[\sup_{0 \leq t \leq T} |X_t|] < \infty \]
for all $T > 0$, then $X$ is a martingale.

We now look at continuous local martingales and define the following process families:

\[ \mathbb{M}^{loc}_C = \{ M_t : M \text{ is a a.s. continuous local martingale}, M_0 = 0 \}, \]
\[ \mathbb{A}_C = \{ A_t : A \text{ is an adapted, a.s. continuous and increasing process}, A_0 = 0 \}, \]
\[ \mathbb{V}_C = \{ V_t : V \text{ is an adapted, a.s. cont. and finite variation process}, V_0 = 0 \}. \]

We identify indistinguishable processes, and note that processes in $\mathbb{M}^{loc}_C$, $\mathbb{A}_C$ and $\mathbb{V}_C$ are indistinguishable of continuous processes because of Lemma 1.1.5.

Theorem 1.10.5. Let $M \in \mathbb{M}^{loc}_C$. Then

i) there exists a unique (modulo indistinguishability) $A \in \mathbb{A}_C$, such that
\[ (M^2 - A_t)_{t \in \mathbb{R}_+} \]
is a local martingale. If $M$ is a square integrable martingale, then we have $E[A_t] < \infty$ and $(M^2 - A_t)_{t \in \mathbb{R}_+}$ is a martingale. We denote $A_t$ by $\langle M \rangle_t$ and call $A$ quadratic variation or sharp bracket of $M$.

ii) Let $\pi_n : 0 = t_0^n < \cdots < t_{k_n}^n = t$ be a sequence of partitions of $[0, t]$ with $|\pi_n| \to 0$. Then
\[ V_n(M) := \sum_{i=0}^{k_n-1} \left( M_{t_{i+1}^n} - M_{t_i^n} \right)^2 \quad \text{as } |\pi_n| \to 0 \]
in probability.
Proof. We only show \( ii \). Suppose first \( |M_s| + A_s \leq k \) for all \( s > 0 \). It follows by \( i \) that \( M^2_t - A_t \) is a bounded martingale and

\[
E[V_n] = \sum_{i=0}^{k_n-1} E\left[ (M_{t_{i+1}n} - M_{tn})^2 \right] = E[A_t] \leq k \quad (1.10)
\]

and

\[
E\left[ (M_{t_{i+1}n} - M_{tn})^2 \mid F_{tn} \right] = E\left[ (M_{tn} - M_{t_{i+1}n})^2 \mid F_{tn} \right] = E\left[ A_{t_{i+1}n} - A_{tn} \mid F_{tn} \right].
\]

So we get

\[
E\left[ (V_n - A_t)^2 \right] = E\left[ \left( \sum_{i=0}^{k_n-1} (M_{t_{i+1}n} - M_{tn})^2 - (A_{t_{i+1}n} - A_{tn}) \right)^2 \right]
\]

\[
= E\left[ \sum_{i=0}^{k_n-1} (M_{t_{i+1}n} - M_{tn})^2 \right] + 2E\left[ \sum_{i=0}^{k_n-1} (A_{t_{i+1}n} - A_{tn})^2 \right]
\]

\[
\leq 2E\left[ \sup_{i} (M_{t_{i+1}n} - M_{tn})^2 \right] + 2E\left[ \sup_{i} (A_{t_{i+1}n} - A_{tn}) \right].
\]

(1.11)

Together with (1.10) we get

\[
E\left[ (V_n - A_t)^2 \right] \leq 8k^3 + 2k^2
\]

and thus

\[
\sup_n E[(V_n)^2] < +\infty.
\]

From (1.11) we get

\[
E\left[ (V_n - A_t)^2 \right] \leq 2 \left\{ \sup_n E\left[ (V_n)^2 \right] E\left[ \sup_i (M_{t_{i+1}n} - M_{tn})^4 \right] \right\}^{\frac{1}{2}}
\]

\[
+ 2kE\left[ \sup_i (A_{t_{i+1}n} - A_{tn}) \right].
\]

Since

\[
\sup_i (A_{t_{i+1}n} - A_{tn}) \leq k \quad \text{and} \quad \sup_i (M_{t_{i+1}n} - M_{tn})^4 \xrightarrow{a.s.} 0 \quad \text{as} \quad n \to \infty
\]

\[
\leq 32k^4
\]
we get by dominated convergence $V_n \to A_t$ in $L^2$.
In the unbounded case put
\[ \rho_p = \inf \{ s \geq 0 \mid |M_s| + A_s \geq p \} . \]
Then $\rho_p$ is a stopping time and $\rho_p \not\to \infty$ for $p \to \infty$ by Example 1.6.3. Further $M_{i \land \rho_p}$ and $M_{i \land \rho_p}^2 - A_{i \land \rho_p}$ are bounded martingales.

Now, for $\alpha, \varepsilon > 0$ given, there exists a $p$ such that $\mathbb{P}(t > \rho_p) < \varepsilon$.
It follows
\[
\mathbb{P}(|V_n(M) - A_t| > \alpha) \leq \mathbb{P}(|V_n(M) - A_t| > \alpha, t \leq \rho_p) + \varepsilon \\
\leq \mathbb{P}(|V_n(M^{\rho_p}) - A_{i \land \rho_p}| > \alpha) + \varepsilon \\
\xrightarrow{n \to \infty} 0
\]
and $V_n(M) \to A_t$ in probability. 

Example 1.10.6. Let $B$ be a $(F_t)_{t \in \mathbb{R}_+}$-Brownian motion. Then we have seen that $B_t^2 - t$ is a martingale. By Theorem 1.10.5 we thus get
\[ \langle B \rangle_t = t. \]

We extend Theorem 1.10.5:

**Theorem 1.10.7.** For all $M, N \in M_C$ exists a unique (modulo indistinguishability) $V \in V_C$ such that
\[ (M_tN_t - V_t) \in \mathbb{R}_+ \]
is a locale martingale. If $M, N$ are square integrable martingales then $MN - V$ is a martingale. $V$ is called sharp bracket or quadratic covariation of $M,N$ and is denoted by $\langle M, N \rangle$.

Further let $\pi_n : 0 = t_0^n < \cdots < t_k^n = t$ be a sequence of partitions of $[0, t]$ with $|\pi_n| \xrightarrow{n \to \infty} 0$. Then
\[
\sum_{i=0}^{k_n-1} (M_{i+1}^t - M_i^t)(N_{i+1}^t - N_i^t) \xrightarrow{n \to \infty} \langle M, N \rangle_t
\]
in probability. Moreover the map
\[ (M, N) \mapsto \langle M, N \rangle \]
is bilinear and for every stopping time $\tau$, $t \geq 0$,
\[ \langle M^\tau, N^\tau \rangle_t = \langle M, N \rangle_t = \langle M^\tau, N \rangle_t = \langle M, N \rangle_{t \land \tau}. \]

**Proof.** Uniqueness: Follows from (1.12) and Lemma 1.1.5.
Existence: Rewrite for $t \geq 0$
\[ M_tN_t = \frac{1}{4} \{(M_t + N_t)^2 + (M_t - N_t)^2\} \]

So there exists \( \langle M + N \rangle, \langle M - N \rangle \) and \( \langle M, N \rangle \) is given by
\[
\langle M, N \rangle_t = \frac{1}{4} ((M + N)_t + (M - N)_t) .
\]

From (1.9) in Theorem 1.10.5 we obtain (1.12).

Further, since local martingales are stable under stopping \( (M^\tau) \) is again a local martingale for all stopping times \( \tau \) by Corollary 1.6.11), we have that
\[
\langle M^\tau, N^\tau \rangle_t = \langle M, N \rangle_{t \wedge \tau} .
\]

Moreover \( \langle M^\tau, N^\tau \rangle_t = \langle M, N \rangle_{t \wedge \tau} \) follows by (1.9). \( \square \)

Attention: There exist integrable, or even square integrable, locale martingales, which are not martingales! However, the integrability of \( \langle M \rangle_t \) gives information about \( M_t \):

**Lemma 1.10.8.** Let \( M \in \mathcal{M}_{C}^{loc} \). Then
\[
E \left[ \sup_t |M_t| \right] \leq 3 E \left[ \sqrt{\langle M \rangle_\infty} \right] .
\]

**Proof.** Let \( A_t := \langle M \rangle_t \) and
\[
\rho_p := \inf \{ s \geq 0 : |M_s| + A_s \geq p \} .
\]

Then \( M_{t \wedge \rho_p} \) and \( M^2_{t \wedge \rho_p} - A_{t \wedge \rho_p} \) are bounded martingales. Let further
\[
\sigma_c := \inf \{ t \geq 0 : A_t \geq c^2 \} .
\]

By proposition 1.7.1
\[
\mathbb{P} \left( \sup_t M^2_{t \wedge \sigma_c \land \rho_p} > c^2 \right) \leq \frac{1}{c^2} E \left[ M^2_{\sigma_c \land \rho_p} \right] = \frac{1}{c^2} E \left[ A_{\sigma_c \land \rho_p} \right] 
\leq \frac{1}{c^2} E \left[ A_{\sigma_c} \right] \leq \frac{1}{c^2} E \left[ c^2 \wedge A_\infty \right] .
\]

Then
\[
\mathbb{P} \left( \sup_t M^2_{t \wedge \sigma_c} > c^2 \right) \leq \frac{1}{c^2} E \left[ c^2 \wedge A_\infty \right] ;
\]

and
\[
\mathbb{P} \left( \sup_t M^2_t > c^2 \right) \leq \mathbb{P}(\sigma_c < +\infty) + \mathbb{P} \left( \sup_t M^2_{t \wedge \sigma_c} > c^2 \right) 
\leq \mathbb{P}(A_\infty \geq c^2) + e^{-c^2} E \left[ c^2 \wedge A_\infty \right] .
\]
We get

\[
E \left[ \sup_t |M_t| \right] = \int_0^\infty \mathbb{P} \left( \sup_t M_t^2 > c^2 \right) dc \\
\leq \int_0^\infty \mathbb{P} (A_\infty \geq c^2) dc + \int_0^\infty c^{-2} E \left[ c^2 \wedge A_\infty \right] dc \\
= E \left[ \int_0^{\sqrt{A_\infty}} dc + \int_0^{\sqrt{A_\infty}} dc + A_\infty \int_0^{\sqrt{A_\infty}} c^{-2} dc \right] \\
= 3E \left[ \sqrt{A_\infty} \right].
\]

From Lemma 1.10.8 we obtain:

**Theorem 1.10.9.** Let \( M \in M_{C}^{\text{loc}} \). Then

i) \( M_\infty = \lim_{t \to \infty} M_t \) exists a.s. on \( \{ \langle M \rangle_\infty < \infty \} \).

ii) if \( E[\sqrt{\langle M \rangle_t}] < \infty \ \forall t \geq 0 \), then \( M \) is a martingale.

iii) if \( E[\sqrt{\langle M \rangle_\infty}] < \infty \), then \( M_t \xrightarrow{a.s. \ & \ L^1} M_\infty \in L^1 \) and for all stopping times \( \tau \) we have \( M_\tau = E[M_\infty | \mathcal{F}_\tau] \) a.s.

iv) if \( E[\langle M \rangle_t] < \infty \ \forall t \geq 0 \), then \( M \) is a square integrable martingale and \( M_t^2 - \langle M \rangle_t \) is a martingale.

v) if \( E[\langle M \rangle_\infty] < \infty \), then \( M \in H_2^C \).

**Proof.** For the proof we refer to Priouret [15], Theorem 4.6.10. □
Stochastic Integration

The main references for this chapter are Priouret [15] and Protter [17]. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$ be a complete, filtered probability space on which we identify indistinguishable processes.

2.1 Integral w.r.t. Finite Variation Processes

In this section we will see how the concept of finite variation functions (see Appendix D) and the connected idea of Stieltjes integration may (under certain conditions) be extended to stochastic processes.

Suppose $A = (A_t)_{t \in \mathbb{R}_+} \in h_C$, i.e. $A$ is an adapted, a.s. continuous and increasing process with $A_0 = 0$, i.e. its sample paths are a.s. increasing continuous functions from $\mathbb{R}_+$ into $\mathbb{R}_+$. By Theorem C.0.2 we have that for almost every $\omega \in \Omega$, there exists a unique measure

$$\mu(\cdot, \cdot) \text{ on } \mathcal{B}(\mathbb{R}_+),$$

such that

$$A_t(\omega) = \mu(\omega, [0, t]).$$

If $X = (X_t)_{t \in \mathbb{R}_+}$ is an a.s. positive measurable process, we can define a new process

$$(X \cdot A)_t = \int_0^t X_s dA_s, \quad t \geq 0,$$

where for almost all $\omega$

$$(X \cdot A)_t(\omega) := \int_{[0,t]} X_s(\omega) \mu(\omega, ds), \quad t \geq 0. \quad (2.1)$$

If $X$ is real-valued we can analogously define $(X \cdot A)_t = \int_0^t X_s dA_s$, if for every $t \geq 0$, 

\[
\int_0^t |X_s|dA_s < \infty \text{ a.s.}
\]

\((X \cdot A)_t\) is then called the path-wise or Stieltjes integral of \(X\) with respect to \(A\).

Let now \(V = (V_t)_{t \in \mathbb{R}^+} \in \mathcal{V}_C\), i.e. \(V\) is an adapted process with a.s. continuous sample paths of finite variation with \(V_0 = 0\) (we say that \(V\) is a continuous finite variation process or FV process). As in the deterministic case, if \(V = (V_t)_{t \in \mathbb{R}^+}\) is an FV process, we may define its total variation process \(S^V = (S^V_t)_{t \in \mathbb{R}^+}\) by

\[
S^V_t(\omega) = \sup_\pi \sum_{k=1}^n |V_{t_k}(\omega) - V_{t_{k-1}}(\omega)|,
\]

where the supremum is taken over all possible partitions \(\pi: 0 = t_0 < t_1 < \cdots < t_n = t\) of \([0, t]\).

The total variation \(S^V\) of \(V\) has the following useful properties.

- \(S^V \in \mathcal{A}_C\).
- \(|V_t| \leq S^V_t\) for all \(t \geq 0\).
- The processes \(A^1 = (A^1_t)_{t \in \mathbb{R}^+}\) and \(A^2 = (A^2_t)_{t \in \mathbb{R}^+}\), with

\[
A^1_t = \frac{S^V_t + V_t}{2}, \\
A^2_t = \frac{S^V_t - V_t}{2},
\]

are a.s. increasing, adapted, continuous processes, i.e. they belong to \(\mathcal{A}_C\).

Clearly \(S^V_t = A^1_t + A^2_t\) and \(V_t = A^1_t - A^2_t\) for all \(t \geq 0\).

The next definition shows, how we can extend the integral as defined in (2.1) from increasing continuous processes to the case of finite variation continuous processes as integrators.

**Definition 2.1.1.** Let \(V = (V_t)_{t \in \mathbb{R}^+} \in \mathcal{V}_C\) and let \(X = (X_t)_{t \in \mathbb{R}^+}\) be progressive, such that for every \(t \geq 0\), \(\int_0^t |X_s|dS^V_s < \infty \text{ a.s.}\). We define

\[
(X \cdot V)_t = \int_0^t X_s dV_s := \int_0^t X_s dA^1_s - \int_0^t X_s dA^2_s, \quad t \geq 0.
\]

**Proposition 2.1.2.** The integral \((X \cdot V)_t, \ t \geq 0\) is an adapted, continuous FV process with \((X \cdot V)_0 = 0\), i.e. \(X \cdot V \in \mathcal{V}_C\). In particular,

\[
\left| \int_0^t X_s dV_s \right| \leq \int_0^t |X_s|dS^V_s.
\]

**Proof.** Exercise. \(\square\)
Example 2.1.3. Let \( \Psi = (\Psi_t)_{t \in \mathbb{R}_+} \) be progressive, such that for every \( t \geq 0 \),
\[
\int_0^t |\Psi_s| \, ds < \infty \quad \text{a.s.}
\]
Then \( (V_t)_{t \in \mathbb{R}_+} = (\int_0^t \Psi_s \, ds)_{t \in \mathbb{R}_+} \in \mathbb{V}_C \), and \( (S_t^V)_{t \in \mathbb{R}_+} = (\int_0^t X_s \Psi_s \, ds)_{t \in \mathbb{R}_+} \). Furthermore,
\[
(\int_0^t X_s \, dV_s)_{t \in \mathbb{R}_+} = (\int_0^t X_s \Psi_s \, ds)_{t \in \mathbb{R}_+}
\]
is well-defined for every measurable process \( (X_t)_{t \in \mathbb{R}_+} \), such that
\[
\int_0^t |X_s| \cdot |\Psi_s| \, ds < \infty \quad \text{for all } t \geq 0.
\]
Note that in order to obtain that the integral process \( \int_0^t X_s \, dV_s \) is adapted, we have to require that the integrand \( X \) is actually progressive, the measurability of the process alone, as assumed in (2.1), is not sufficient.

If, in addition, the integrand process \( X \) has continuous trajectories, then the integral in (2.2) is known as the path-wise Riemann-Stieltjes integral, and the following approximation result holds.

Theorem 2.1.4. Let \( V = (V_t)_{t \in \mathbb{R}_+} \) be a continuous FV process and let \( X = (X_t)_{t \in \mathbb{R}_+} \) be continuous and adapted. Let \( (\pi_n)_{n \in \mathbb{N}} \) be a sequence of partitions \( \pi_n : 0 = t^n_0 < t^n_1 < \cdots < t^n_k^n = t \) of \([0, t]\), such that
\[
|\pi_n| := \sup_i |t^n_i - t^n_{i-1}| \xrightarrow{n \to \infty} 0.
\]
Then for \( t^n_i \leq s^n_n \leq t^n_{i+1} \),
\[
\lim_{n \to \infty} \sum_{i, t^n_i < s^n_n \leq t^n_{i+1} \in \pi_n} X_{s^n_n} \left( V_{t^n_{i+1}} - V_{t^n_i} \right) = \int_0^t X_s \, dV_s \quad \text{a.s.}
\]

Proof. See Protter and Morrey [16], pp. 316–317.

Unfortunately, there are bad news concerning the Brownian motion!

Proposition 2.1.5. For almost all \( \omega \), the sample paths \( t \mapsto B_t(\omega) \) of a standard Brownian motion \( B \) are of unbounded variation on any interval.

Note that this proposition (that we will prove later, see Corollary 2.2.2) does not mean that it is always impossible to define the integral \( H \cdot B \) path by path. For special, good choices of \( H \), the path-wise definition may still work out fine. If, for example, \( H \) is a process whose sample paths are differentiable with bounded derivatives on \([0, 1]\) (or more generally, whose sample paths are of bounded \( q \)-variation for \( q < 2 \)), then the Stieltjes integral
\[
\int_0^1 H_s(\omega) \, dB_s(\omega)
\]
exists for every sample path \( (B_t(\omega))_{t \in \mathbb{R}_+} \) of \( B \).

Example 2.1.6. The following Riemann-Stieltjes integrals are well-defined:
\[
\int_0^1 e^t \, dB_t, \quad \int_0^1 \sin t \, dB_t, \quad \int_0^1 t^p \, dB_t, \quad p \geq 0.
\]
See Mikosch [12], Chapter 2, for further details.
In contrast to the integrals that we have seen in the preceding example, if one tries to define
\[ \int_0^1 B_t(\omega)dB_t(\omega) \]
as path-wise Riemann-Stieltjes integral, then this is impossible due to a more general result that is a consequence of the Banach-Steinhaus Theorem, see e.g. Mikosch [12], Chapter 2. Recall that the path-wise integral is defined \( \omega \)-per \( \omega \), hence for a fixed \( \omega \) it is actually given by the integral of a deterministic function with respect to a measure on \( B(\mathbb{R}_+) \). Now consider a deterministic continuous function \( f : [0,1] \rightarrow \mathbb{R} \) and let \( (\pi_n)_{n \in \mathbb{N}} \) denote a refining sequence of dyadic rational partitions of \( [0,1] \), i.e. \( \pi_n : 0 = t^n_0 < t^n_1 < \cdots < t^n_{k_n} = 1 \), with \( \lim_{n \rightarrow \infty} |\pi_n| = 0 \). Define
\[ S_n(h) = \sum_{t^n_l, t^n_{l+1} \in \pi_n} h(t^n_l) \left( f(t^n_{l+1}) - f(t^n_l) \right), \quad (2.3) \]
for an arbitrary continuous function \( h : [0,1] \rightarrow \mathbb{R} \). Which conditions are now required for \( f \) so that \( S_n(h) \) converges for every continuous function \( h \)? While Theorem 2.1.4 tells us that \( f \) continuous and of finite variation is sufficient, the next theorem implies, that this condition in fact is really necessary.

**Theorem 2.1.7.** If the sums \( (S_n)_{n \in \mathbb{N}} \) as defined in (2.3) converge to a finite limit for every continuous function \( h : [0,1] \rightarrow \mathbb{R} \), then \( f \) has finite variation.

**Proof.** Apply the Banach-Steinhaus theorem to
\[ X = \{ h : [0,1] \rightarrow \mathbb{R} : h \text{ continuous} \}, \]
equipped with the supremum norm and \( Y = \mathbb{R} \), equipped with the absolute value as the norm. Define \( (T_n)_{n \in \mathbb{N}} = (S_n)_{n \in \mathbb{N}} \), where \( S_n : X \rightarrow Y \), \( h \mapsto S_n(h) \).

Returning to stochastic processes, we may hope to avoid the limitations imposed on the integrator by this theorem by asking the convergence of
\[ S_n(H) = \sum_{t^n_l, t^n_{l+1} \in \pi_n} H_{t^n_l} \left( X_{t^n_{l+1}} - X_{t^n_l} \right) \quad (2.4) \]
to hold only in probability (instead of almost surely), for an integrator process \( X \) and an integrand process \( H \), both adapted and continuous. However, most unfortunately, this does not help, and we still obtain that if the sequence in (2.4) converges to a finite limit for every continuous process \( H \), then \( X \) must be an FV process.

---

*Banach-Steinhaus theorem.* Let \( X \) be a Banach space and \( Y \) be a normed linear space. Let \( \{T_n\} \) be a family of bounded linear operators from \( X \) into \( Y \). If for every \( x \in X \) the set \( \{T_n x\} \) is bounded, then the set \( \{\|T_n\|\} \) is bounded.
This means that for the moment it is impossible, to coherently define $H \cdot X$ as a path-wise Stieltjes integral if $X$ is a process with paths of infinite variation, e.g. a Brownian motion. Therefore, in the sections to come we will need to develop a rather different notion of stochastic integration. The Brownian motion will serve as integrator, and as integrands we first consider so-called simple processes, later extending this class of integrands to a wider class of stochastic processes.

2.2 Stochastic Integral w.r.t. Brownian Motion

Let $B = (B_t)_{t \in \mathbb{R}_+}$ be a real-valued standard $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$-Brownian motion on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, P)$. We want to define for $t \geq 0$

$$\int_0^t \Phi_s dB_s$$

(2.5)

for a suitable class of integrands $(\Phi_t)_{t \in \mathbb{R}_+}$. However, we cannot use the concept of pathwise or Stieltjes integration because Brownian motion is not a FV process as we will show now. We recall from Section 1.10 that $\mathcal{M}^\text{loc}_C$ denotes the family of continuous local martingales that start in zero.

**Theorem 2.2.1.**

$$\mathcal{M}^\text{loc}_C \cap \mathcal{V}_C = \{0\}.$$

**Proof.** Suppose $M \in \mathcal{M}^\text{loc}_C \cap \mathcal{V}_C$, and let $S^M$ be the total variation of $M$. Then $S^M$ is a.s. a continuous process (we have already stated this without proof). We now distinguish two cases.

i) Suppose $S^M_t \leq K$ for all $t \geq 0$. Then $|M_t| \leq K$, and therefore by Proposition 1.10.3 $M$ is not only a local martingale, but actually a martingale. Consider a sequence $(\pi_n)_{n \in \mathbb{N}}$ of partitions of $[0, t]$

$$\pi_n : 0 = t_0^n < \cdots < t_{k_n}^n = t,$$

such that $|\pi_n| \xrightarrow{n \to \infty} 0$. By using results from Section 1.9 we have that

$$E \left[ M_t^2 \right] = E \left[ \sum_{i=0}^{k_n-1} \left( M_{t_{i+1}^n}^2 - M_{t_i^n}^2 \right) \right]$$

$$= E \left[ \sum_{i=0}^{k_n-1} \left( M_{t_{i+1}^n} - M_{t_i^n} \right)^2 \right] \leq$$

$$\leq E \left[ \sup_i \left| M_{t_{i+1}^n} - M_{t_i^n} \right| \sum_{i=0}^{k_n-1} \left| M_{t_{i+1}^n} - M_{t_i^n} \right| \right] \cdot S^M_t$$

$$\leq E \left[ \sup_i \left| M_{t_{i+1}^n} - M_{t_i^n} \right| \cdot S^M_t \right]$$
\[
\begin{align*}
\leq K \cdot E \left[ \sup_{t} \left| M_{\eta+1}^n - M_{\eta}^n \right| \right] \xrightarrow{n \to \infty} 0
\end{align*}
\]
by Lebesgue’s theorem, since \( M \) is continuous and bounded. Hence for every \( t \geq 0 \), \( M_t = 0 \) a.s., and by Lemma 1.1.5 we conclude that \( M \) is indistinguishable from 0.

ii) In the general case where \( S^M \) is not necessarily bounded, we consider the stopping time
\[
\rho_n = \inf \{ t > 0 : S^M_t \geq n \}
\]
and apply (i) to \( M_{\rho_n}^0 \). \( \square \)

**Corollary 2.2.2.** The Brownian motion is not an FV process.

**Proof.** Since the Brownian motion \( B \) is a continuous martingale with \( B_0 = 0 \), then \( B \in M^\infty_{C} \), and by Theorem 2.2.1 it cannot have sample paths of finite variation. \( \square \)

In order to define the integral in (2.5), let us first consider the relatively small class of simple processes, which we will extend subsequently to more general classes of integrands.

**Definition 2.2.3.** A simple process \( \Phi = (\Phi_t)_{t \in \mathbb{R}^+} \) is a process of the form
\[
\Phi_t = \sum_{i=0}^{n-1} U_i \mathbb{1}_{(t_i, t_{i+1}]}(t),
\]
where \( 0 = t_0 < t_1 < \cdots < t_n < \infty \), and \( U_i \in \mathcal{B} \mathcal{F}_{t_i} \), i.e. \( U_i \) is a bounded \( \mathcal{F}_{t_i} \)-measurable random variable. We denote by \( \mathcal{E} \) the family of all simple processes. We define the stochastic integral for an integrand \( \Phi \in \mathcal{E} \) with respect to \( B \) by
\[
\int_0^t \Phi_s dB_s := \sum_{i=0}^{n-1} U_i \left( B_{t_{i+1} \wedge t} - B_{t_i \wedge t} \right), \quad 0 < t \leq \infty.
\]
Recall, that we denote by \( \mathbb{H}^2_C \) the space of all continuous martingales \( M \), such that \( \sup_t E \left[ M_t^2 \right] < \infty \) and \( M_0 = 0 \). The following proposition reveals the properties of the stochastic integral for \( \Phi \in \mathcal{E} \).

**Proposition 2.2.4.** Let \( \Phi \in \mathcal{E} \). Then \( M = (M_t)_{t \in \mathbb{R}^+} = (\int_0^t \Phi_s dB_s)_{t \in \mathbb{R}^+} \in \mathbb{H}^2_C \), and the sharp bracket \( \langle M \rangle \) of \( M \) is given by
\[
\langle M \rangle_t = \int_0^t \Phi_s^2 ds.
\]
In particular
\[
E \left[ \int_0^\infty \Phi_s dB_s \right] = 0 \quad \text{and} \quad E \left[ \left( \int_0^\infty \Phi_s dB_s \right)^2 \right] = E \left[ \int_0^\infty \Phi_s^2 ds \right].
\]
For the proof we will use the following lemma, that can easily be verified.

**Lemma 2.2.5.** Suppose \( X = (X_t)_{t \in \mathbb{R}_+} \) is an integrable adapted process and 
\( 0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = \infty \). If for any \( s, t \) such that \( t_i \leq s < t \leq t_{i+1} \),
\[
E \left[ X_t - X_s \mid \mathcal{F}_s \right] = 0,
\]
then \( X \) is a martingale.

**Proof (Proposition 2.2.4).** By using this result, we may restrict our attention to the case of \( t_i \leq s < t \leq t_{i+1} \). First we note that
\[
M_t - M_s = U_i(B_t - B_s),
\]
with \( U_i \in \mathcal{b}\mathcal{F}_t \subset \mathcal{b}\mathcal{F}_s \), and therefore
\[
E \left[ M_t - M_s \mid \mathcal{F}_s \right] = U_iE \left[ B_t - B_s \mid \mathcal{F}_s \right] = 0.
\]
Now set \( A_t = \int_0^t \Phi_u^2 \, du \). Then \( A_t - A_s = U_i^2(t - s) \). Since
\[
M_t^2 - M_s^2 = (M_t - M_s)^2 + 2M_s(M_t - M_s)
\]
\[
= U_i^2(B_t - B_s)^2 + 2M_s(M_t - M_s),
\]
we have
\[
E \left[ M_t^2 - M_s^2 \mid \mathcal{F}_s \right] = U_i^2 E \left[ (B_t - B_s)^2 \mid \mathcal{F}_s \right] + 2M_s E \left[ M_t - M_s \mid \mathcal{F}_s \right]
\]
\[
= U_i^2(t - s) = A_t - A_s,
\]
and since \( A_t \) is \( \mathcal{F}_s \)-measurable, this implies that \( \langle M_t \rangle = \int_0^t \Phi_u^2 \, du \).

By the previous calculations we then know that for all \( t \geq 0 \),
\[
E \left[ M_t^2 \right] = E \left[ A_t \right].
\]
Since \( \Phi_t = \sum_{i=0}^{n-1} U_i \mathbb{1}_{(t_i, t_{i+1})}(t) \), for \( t > t_n \) we have
\[
M_t = M_{t_n} \text{ and } A_t = A_{t_n},
\]
which implies
\[
\sup_t E \left[ M_t^2 \right] = \sup_t E \left[ A_t \right] = E \left[ A_{t_n} \right] < \infty,
\]
and finally
\[
E \left[ M_\infty^2 \right] = E \left[ M_{t_n}^2 \right] = E \left[ A_{t_n} \right] = E \left[ A_\infty \right].
\]
\[\square\]
We now extend the definition of the stochastic integral to a more general class of integrands. Consider the Hilbert space

$$\Lambda^2 = L^2(\mathbb{R}_+ \times \Omega, \mathcal{F}_t, dx \otimes \mathbb{P}),$$

where $dx$ denotes the Lebesgue measure on $\mathbb{R}_+$, endowed with the norm

$$\|\Phi\|_{\Lambda^2} = \mathbb{E}\left[\int_0^{\infty} \Phi_s^2 ds\right]^{\frac{1}{2}}.$$

Note that clearly $\mathcal{E} \subset \Lambda^2$ (with the usual abuse of language, since $\Lambda^2$ is a space of equivalence classes), and the map

$$I : \Phi \mapsto I(\Phi) := \int_0^{+\infty} \Phi_s dB_s$$

is a linear isometry from $\mathcal{E}$ (endowed with the norm $\| \cdot \|_{\Lambda^2}$) into $L^2(\Omega, \mathcal{F}, \mathbb{P})$. As we will show now, $I$ can be uniquely extended to a linear isometry from $\Lambda^2$ into $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

**Proposition 2.2.6.** The space $\mathcal{E}$ is dense in $(\Lambda^2, \| \cdot \|_{\Lambda^2})$.

**Proof.** Let $\Phi \in \Lambda^2$ be orthogonal to $\mathcal{E}$. We need only to show that $\Phi \equiv 0 \in \Lambda^2$. Let $X = (X_t)_{t \in \mathbb{R}_+}$ be given by $X_t = \int_0^t \Phi_s ds$. Since $|\Phi_t| < 1 + \Phi_t^2$,

$$\mathbb{E}\left[\int_0^\infty |\Phi_s| ds\right] = \int_0^\infty \int_0^\infty |\Phi_s| ds d\mathbb{P}(\omega) < \infty,$$

which means $X_t \in L^1$ for all $t \geq 0$, and $X \in \mathcal{V}_C$ (see Proposition 2.1.3). Now let $s < t$, assume $A \in \mathcal{F}_t$ and consider $\Psi = (\Psi_t)_{t \in \mathbb{R}_+} \in \mathcal{E}$ defined by $\Psi_u = \mathbb{1}_A \mathbb{1}_{[s,t]}(u)$. Since $\Phi \in \mathcal{E}^\perp$, we have

$$0 = \langle \Phi, \Psi \rangle_{\Lambda^2} = \mathbb{E}\left[\int_0^{+\infty} \Phi_u \Psi_u du\right] = \mathbb{E}\left[\mathbb{1}_A \int_s^t \Phi_u du\right] = \mathbb{E}\left[\mathbb{1}_A (X_t - X_s)\right],$$

i.e. $X$ is a martingale. From Theorem 2.2.1 we know that $\mathcal{M}^{loc} \cap \mathcal{V}_C = \{0\}$, hence we have $X_t = 0$ a.s. for all $t \geq 0$, which implies $\Phi_t = 0$ almost everywhere and therefore $\|\Phi\|_{\Lambda^2} = 0$, i.e. $\Phi \equiv 0 \in \Lambda^2$.

Since the space $\mathcal{E}$ of simple processes is dense in $\Lambda^2$, i.e. $\mathcal{E} = \Lambda^2$, the map $I$ can be extended from integrands in $\mathcal{E}$ to the whole space $\Lambda^2$ because $I$ is closable. These are the main steps:

1. Let $\Phi \in \Lambda^2$. Then by Proposition 2.2.6 there exists a sequence of processes $(\Phi_n)_{n \in \mathbb{N}} \subseteq \mathcal{E}$, with

$$\|\Phi_n - \Phi\|_{\Lambda^2} \xrightarrow{n \to \infty} 0.$$
ii) Since for \( m, n \in \mathbb{N} \), \( \|I(\Phi_n) - I(\Phi_m)\|_2 = \|\Phi_n - \Phi_m\|_{A^2} \), clearly \((I(\Phi_n))_{n \in \mathbb{N}}\) is a Cauchy sequence in the Hilbert space \( L^2 \). Hence there exists \( X = \lim_{n \to \infty} I(\Phi_n) \) in \( L^2 \). Further, it is easy to see that \( X \) is independent of the approximating sequence \( \Phi_n \) (i.e. \( I \) is closable).

iii) We therefore define \( I(\Phi) := X \), and call it the stochastic integral of \( \Phi \) over \( \mathbb{R}_+ \), i.e. \( I(\Phi) = \int_0^{+\infty} \Phi_s dB_s \).

We summarize our findings and the properties of the stochastic integral in the following theorem.

**Theorem 2.2.7.** There exists a unique linear map
\[
\mathbb{I} : A^2 \longrightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}),
\]
such that

i) if \( \Phi = U \mathbb{I}_{(s,t]} \), \( U \in \mathcal{B} \mathcal{F}_s \), then
\[
I(\Phi) = \int_0^\infty \Phi_u dB_u = U(B_t - B_s).
\]

ii) for every \( \Phi \in A^2 \), \( \|I(\Phi)\|_2 = \|\Phi\|_{A^2} \), i.e.
\[
E \left[ \int_0^\infty \Phi_s dB_s \right] = E \left[ \int_0^\infty \Phi_s^2 ds \right].
\]

Moreover, we have
\[
E[\I(\Phi)] = E \left[ \int_0^{+\infty} \Phi_s dB_s \right] = 0
\]
and
\[
E[I(\Phi)\I(\Psi)] = E \left[ \int_0^\infty \Phi_s dB_s \int_0^\infty \Psi dB_s \right]
= E \left[ \int_0^\infty \Phi_s \Psi_s ds \right]
\]
for \( \Phi_s, \Psi_s \in A^2 \).

We now define the integral for a finite interval \([0, T] \), \( T > 0 \). Consider
\[
A^2(T) = L^2 \left( [0, T] \times \Omega, \mathcal{P}_r, dx|_{[0, T]} \otimes \mathbb{P} \right),
\]
where \( dx|_{[0, T]} \) denotes the Lebesgue measure restricted to \([0, T] \). If \( \Phi \in A^2(T) \), then \((\Phi_t \mathbb{I}_{[0,T]}(t))_{t \in \mathbb{R}_+} \in A^2 \), and consequently we set
\[
\int_0^T \Phi_s dB_s := \int_0^{+\infty} \Phi_s \mathbb{I}_{[0,T]}(s) dB_s.
\]

Then \( \int_0^T \Phi_s dB_s \) is called the stochastic integral of \( \Phi \) on \([0, T]\) and it has the following properties.
Corollary 2.2.8. Let $\Phi \in \Lambda^2(T)$. Then
\[
E \left[ \int_0^T \Phi_s dB_s \right] = 0,
\]
\[
E \left[ \left( \int_0^T \Phi_s dB_s \right)^2 \right] = E \left[ \int_0^T \Phi_s^2 ds \right].
\]

Let $\Phi, \Psi \in \Lambda^2(T)$. Then
\[
E \left[ \int_0^T \Phi_s dB_s \cdot \int_0^T \Psi_s dB_s \right] = E \left[ \int_0^T \Phi_s \Psi_s ds \right].
\]

Remark 2.2.9. We would like to give a short comment on the choice of the measurability conditions for the integrand processes. We have defined the stochastic integral with respect to the Brownian motion for integrands $\Phi$ that are progressively measurable. However, this requirement is not very intuitive, especially when it comes to applications. What are alternative hypotheses?

(i) By Proposition 1.4.5, we know that if $\Phi$ is right-continuous (resp. left-continuous) and adapted, then $\Phi$ is progressive.

(ii) According to Chung and Williams [2], if $\Phi$ is an adapted measurable process such that $E \left[ \int_0^{+\infty} \Phi_s^2 ds \right] < \infty$, then there exists $(\Phi^n)_{n \in \mathbb{N}} \subseteq \mathcal{E}$ such that
\[
E \left[ \int_0^{+\infty} (\Phi_s - \Phi^n_s)^2 ds \right] \xrightarrow{n \to \infty} 0.
\]

Since $(\Phi^n)_{n \in \mathbb{N}}$ converges in $\Lambda^2$ to a progressive process $\Psi \in \Lambda^2$, then $\Phi$ belongs to the equivalence class of $\Psi$ in $\Lambda^2$, since $\Phi = \Psi \, dx \otimes \mathbb{P}$-a.s.

2.3 The Stochastic Integral as a continuous Process

In this section (more precisely, in Theorem 2.3.1) we show that for an integrand $\Phi \in \Lambda^2$, the stochastic integral process $(\int_0^t \Phi_s dB_s)_{t \in \mathbb{R}_+}$ is a martingale belonging to $\mathbb{H}^2_{\mathbb{C}}$.

Theorem 2.3.1. Let $\Phi \in \Lambda^2$. There exists $M = (M_t)_{t \in \mathbb{R}_+} \in \mathbb{H}^2_{\mathbb{C}}$ such that for all $t \geq 0$,
\[
M_t = \int_0^t \Phi_s dB_s \quad a.s.,
\]

i.e. $M$ is a modification of $(\int_0^t \Phi_s dB_s)_{t \in \mathbb{R}_+}$. Moreover, $(M_t^2 - \int_0^t \Phi_s^2 ds)_{t \in \mathbb{R}_+}$ is a martingale.
Proof. Consider \( X = \int_0^\infty \Phi_s dB_s \) and \((\Phi^n)_{n \in \mathbb{N}} \subseteq \mathcal{E}\) such that

\[
E \left[ \int_0^\infty (\Phi^n_s - \Phi_s)^2 \, ds \right] \xrightarrow{n \to \infty} 0.
\]

Since \( \Phi^n \in \mathcal{E} \), by Proposition 2.2.4 we have that \( M^n = (M^n_t)_{t \in \mathbb{R}^+} \in \mathbb{H}^2_{\mathcal{C}} \), where \( M^n_t = \int_0^t \Phi^n_s dB_s \) and \( (M^n)_t = \int_0^t (\Phi^n_s)^2 \, ds, \ t \geq 0 \). Then \( M^n_n \xrightarrow{n \to \infty} X \) in \( L^2 \) since

\[
E \left[ |M^n_n - X|^2 \right] = E \left[ \left( \int_0^\infty \Phi^n_s dB_s - \int_0^\infty \Phi_s dB_s \right)^2 \right] = E \left[ \int_0^\infty (\Phi^n_s - \Phi_s)^2 \, ds \right] \xrightarrow{n \to \infty} 0.
\]

By Proposition 1.9.1 there exists \( M = (M_t)_{t \in \mathbb{R}^+} \in \mathbb{H}^2_{\mathcal{C}} \) such that for every \( t \geq 0 \),

\[
M^n_t = \int_0^t \Phi^n_s dB_s \xrightarrow{n \to \infty} M_t \quad \text{in } L^2 \quad \text{and a.s.}
\]

Since \( \int_0^t \Phi^n_s dB_s \xrightarrow{n \to \infty} \int_0^t \Phi_s dB_s \) in \( L^2 \), then for all \( t \geq 0 \), \( M_t = \int_0^t \Phi_s dB_s \) a.s.

It remains to prove that \((M^2_t - \int_0^t \Phi^2_s ds)_{t \in \mathbb{R}^+}\) is a martingale. Consider \( U^n_t = (U^n_t)_{t \in \mathbb{R}^+} \), where

\[
U^n_t = (M^n_t)^2 - \int_0^t (\Phi^n_s)^2 \, ds.
\]

Then \( U^n \) is a martingale, and since for all \( t \geq 0 \), \((M^n_t)^2 \xrightarrow{n \to \infty} M^2_t \) in \( L^1 \) and \( \int_0^t (\Phi^n_s)^2 \, ds \xrightarrow{n \to \infty} \int_0^t \Phi^2_s \, ds \) in \( L^1 \), then \( U^n \xrightarrow{n \to \infty} U_t = M^2_t - \int_0^t \Phi^2_s \, ds \) in \( L^1 \) and \( U = (U_t)_{t \in \mathbb{R}^+} \) is a martingale. \( \square \)

Due to this theorem, from now on we may always assume that we are working with the continuous modification \( M \in \mathbb{H}^2_{\mathcal{C}} \) of \( (\int_0^t \Phi_s dB_s)_{t \in \mathbb{R}^+} \) for \( \Phi \in \Lambda^2 \), and we call \( M \) the stochastic integral of \( \Phi \).

In the following we investigate further properties of the stochastic integral.

**Corollary 2.3.2.** Let \( \Phi, \Psi \in \Lambda^2 \), and define \( M = (M_t)_{t \in \mathbb{R}^+} \) and \( N = (N_t)_{t \in \mathbb{R}^+} \) by \( M_t = \int_0^t \Phi_s dB_s \) and \( N_t = \int_0^t \Psi_s dB_s \). Then for all \( t \geq 0 \),

\[
(M, N)_t = \int_0^t \Phi_s \Psi_s \, ds
\]

and \((M_t N_t - \int_0^t \Phi_s \Psi_s \, ds)_{t \in \mathbb{R}^+}\) is a martingale.

**Proof.** We only need to write \( \Phi_s \Psi_s = \frac{1}{4} \left((\Phi_s + \Psi_s)^2 - (\Phi_s - \Psi_s)^2\right) \), \( s \geq 0 \). \( \square \)
Corollary 2.3.3. Let $\Phi, \Psi \in A^2$, let $\tau$ be a stopping time and define $M = (M_t)_{t \in \mathbb{R}_+}$ and $N = (N_t)_{t \in \mathbb{R}_+}$ by $M_t = \int_0^t \Phi_s dB_s$ and $N_t = \int_0^t \Psi_s dB_s$. Then for all $t \geq 0$,

$$E[M_{t \wedge \tau} N_{t \wedge \tau}] = E\left[\int_0^{t \wedge \tau} \Phi_s \Psi_s ds\right] \ a.s. \quad (2.6)$$

Proof. By Corollary 2.3.2 we know that $X = (X_t)_{t \in \mathbb{R}_+}$ defined by

$$X_t = M_t N_t - \int_0^t \Phi_s \Psi_s ds$$

is a continuous martingale closed by $X_\infty$, hence (2.6) follows by applying Theorem 1.6.8 (Doob’s optional sampling theorem).

Corollary 2.3.4. Let $\Phi \in A^2$, let $\tau$ be a stopping time and define $M = (M_t)_{t \in \mathbb{R}_+}$ by $M_t = \int_0^t \Phi_s dB_s$. Then for all $t \geq 0$,

$$M_{t \wedge \tau} = \int_0^t \Phi_s 1_{[0,\tau]}(s) dB_s \ a.s. \quad (2.7)$$

Proof. Define $N = (N_t)_{t \in \mathbb{R}_+}$ by $N_t = \int_0^t \Phi_s 1_{[0,\tau]}(s) dB_s$. To have (2.7) we only need to prove that

$$E[(M_{t \wedge \tau} - N_t)^2] = E[M_{t \wedge \tau}^2] + E[N_t^2] - 2E[M_{t \wedge \tau} N_t] = 0. \quad (2.8)$$

We compute each term of (2.8). By Corollary 2.3.3 we have

$$E[M_{t \wedge \tau}^2] = E\left[\int_0^{t \wedge \tau} \Phi_s^2 ds\right],$$

and by the property of the stochastic integral

$$E[N_t^2] = E\left[\int_0^t \Phi_s^2 1_{[0,\tau]}(s) ds\right].$$

Finally, by Theorem 1.10.7 and Corollary 2.3.3 it follows that

$$E[M_{t \wedge \tau} N_t] = E[M_{t \wedge \tau} N_{t \wedge \tau}] = E\left[\int_0^{t \wedge \tau} \Phi_s \Psi_s ds\right]$$

$$= E\left[\int_0^{t \wedge \tau} \Phi_s^2 1_{[0,\tau]}(s) ds\right].$$

$\square$
2.4 Extension to Integrands in $A^{2}_{loc}$

We now wish to even more generalize the space of stochastic integrands. For this purpose we consider the integrand space

$$A^{2}_{loc} = \{ \Phi : \Phi \text{ progressive and for all } t, \int_{0}^{t} \Phi^2_s ds < +\infty \text{ a.s.} \}.$$ 

We will need the following lemma.

**Lemma 2.4.1.** Let $(X^n)_{n \in \mathbb{N}}$ be a sequence of continuous processes and assume $(\tau_n)_{n \in \mathbb{N}}$ is a sequence of stopping times such that $\tau_n \nearrow +\infty$ a.s. If for all $t \in \mathbb{R}^+$ and $n \in \mathbb{N}$, $X^n_{t \wedge \tau_n} = X^n_{0 \wedge \tau_n}$ a.s., then there exists a continuous process $X = (X_t)_{t \in \mathbb{R}^+}$ such that for all $t \in \mathbb{R}^+$ and $n \in \mathbb{N}$,

$$X_{t \wedge \tau_n} = X^n_{0 \wedge \tau_n} \text{ a.s.}$$

**Proof.** By Lemma 1.1.5, for a fixed $n \in \mathbb{N}$, $(X^n_{t \wedge \tau_n})_{t \in \mathbb{R}^+}$ and $(X^{n+1}_{t \wedge \tau_n})_{t \in \mathbb{R}^+}$ are indistinguishable. Hence there exists a set $N$ of measure zero such that for all $\omega \notin N$,

$$\tau_n(\omega) \nearrow +\infty \text{ and } X^n_{t \wedge \tau_n}(\omega) = X^{n+1}_{t \wedge \tau_n}(\omega) \text{ for all } n \in \mathbb{N}, \ t \geq 0.$$ 

We define $X = (X_t)_{t \in \mathbb{R}^+}$ by

$$X_t = \begin{cases} X^n_t(\omega) : \omega \notin N, \ t < \tau_n(\omega), \\ 0 : \omega \in N. \end{cases}$$

Given $\Phi \in A^{2}_{loc}$, we construct the stochastic integral $(\int_{0}^{t} \Phi_s dB_s)_{t \in \mathbb{R}^+}$ by the following steps.

i) Let $\Phi \in A^{2}_{loc}$. Then

$$\tau_n = \inf \left\{ t \geq 0 : \int_{0}^{t} \Phi^2_s ds \geq n \right\}$$

is a stopping time such that $\tau_n \nearrow +\infty$ (Example 1.6.3) and $E \left[ \int_{0}^{\tau_n} \Phi^2_s ds \right] < +\infty$.

ii) Consider $M^n = (M^n_t)_{t \in \mathbb{R}^+}$, where $M^n_t = \int_{0}^{t} \Phi_s 1_{[0,\tau_n]}(s) dB_s$. Then for all $n \in \mathbb{N}$, $M^n \in H^2_{C}$ and $M^n_t = \int_{0}^{t \wedge \tau_n} \Phi^2_s ds, \ t \geq 0$.

iii) Since $M^n_{t \wedge \tau_n} = M^{n+1}_{t \wedge \tau_n}$ for all $n \in \mathbb{N}$ and $t \geq 0$, by Lemma 2.4.1 there exists a continuous process $M = (M_t)_{t \in \mathbb{R}^+}$ such that $M_{t \wedge \tau_n} = M^n_{t \wedge \tau_n}$ a.s.

iv) By construction $M \in M^{loc}_{C}$ and $U = (U_t)_{t \in \mathbb{R}^+} \in M^{loc}_{C}$, where

$$U_t = M^2_t - \int_{0}^{t} \Phi^2_s ds.$$
Stochastic Integration

v) Stability by stopping: if \( \tau \) is a stopping time such that \( E \left[ \int_0^\tau \Phi_s^2 ds \right] < +\infty \), then by Corollary 2.3.4 for all \( t \geq 0 \) we have

\[
M_{t \wedge \tau} = M_{t \wedge \tau} = \int_0^{t \wedge \tau} \Phi_s \mathbb{1}_{[0, \tau]}(s) dB_s \text{ a.s.}
\]

Letting \( n \to \infty \),

\[
M_{t \wedge \tau} = \int_0^t \Phi_s \mathbb{1}_{[0, \tau]}(s) dB_s \text{ a.s.}
\]

We summarize this construction in the following theorem.

**Theorem 2.4.2.** Let \( \Phi \in A^2_{loc} \). Then there exists a unique \( M \in M^C_{loc} \) such that for every stopping time \( \tau \) with \( E \left[ \int_0^\tau \Phi_s^2 ds \right] < +\infty \) and for all \( t \geq 0 \),

\[
M_{t \wedge \tau} = \int_0^t \Phi_s \mathbb{1}_{[0, \tau]}(s) dB_s \text{ a.s. and}
\]

\[
\langle M \rangle_t = \int_0^t \Phi_s^2 ds.
\]

We call \( M \) the stochastic integral of \( \Phi \in A^2_{loc} \), which we as before denote by \((\int_0^t \Phi_s dB_s)_t \in \mathbb{R}_+ \). Given \( T > 0 \), we now consider the space of integrands

\[
A^2_{loc}(T) = \{ \Phi : \Phi \text{ progressive and } \int_0^T \Phi_s^2 ds < +\infty \text{ a.s.} \}.
\]

If \( \Phi \in A^2_{loc}(T) \), then \((\Phi_t \mathbb{1}_{[0, T]}(t))_t \in \mathbb{R}_+ \in A^2_{loc} \) and we define the stochastic integral of \( \Phi \) on \([0, T]\) as

\[
\int_0^T \Phi_s dB_s := M_T,
\]

where \( M_t = \int_0^t \Phi_s \mathbb{1}_{[0, T]}(s) dB_s, t \geq 0 \). Please pay attention to the fact, that now \( E[\int_0^T \Phi_s dB_s] \) may not necessarily exist. However, the following results are valid anyway.

**Corollary 2.4.3.** Let \( \Phi, \Psi \in A^2_{loc} \), and define \( M = (M_t)_{t \in \mathbb{R}_+} \) and \( N = (N_t)_{t \in \mathbb{R}_+} \) by \( M_t = \int_0^t \Phi_s dB_s \) and \( N_t = \int_0^t \Psi_s dB_s \). Then for all \( t \geq 0 \),

\[
\langle M, N \rangle_t = \int_0^t \Phi_s \Psi_s ds.
\]

**Corollary 2.4.4.** Let \( \Phi \in A^2_{loc} \), let \( \tau \) be a stopping time and define \( M = (M_t)_{t \in \mathbb{R}_+} \) by \( M_t = \int_0^t \Phi_s dB_s \). Then for all \( t \geq 0 \),

\[
M_{t \wedge \tau} = \int_0^t \Phi_s \mathbb{1}_{[0, \tau]}(s) dB_s \text{ a.s.}
\]

Note that Corollary 2.4.4 holds for general stopping times \( \tau \), it is not required anymore that \( \tau \) satisfies \( E \left[ \int_0^\tau \Phi_s^2 ds \right] < +\infty \).
2.5 The Vectorial Case

Let \( B = (B^1, \ldots, B^d) \) be an \((\mathcal{F}_t)_{t \in \mathbb{R}_+}\)-Brownian motion on \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})\) with values in \( \mathbb{R}^d \). Define

\[
\Lambda^2_\mathbb{R} = \{ \Phi = (\Phi^1, \ldots, \Phi^d) : \Phi^i \in \Lambda^2, \ i = 1, \ldots, d \},
\]

\[
\Lambda^2_{\mathbb{R}, d} = \{ \Phi = (\Phi^1, \ldots, \Phi^d) : \Phi^i \in \Lambda^2_{\mathbb{R}}, \ i = 1, \ldots, d \},
\]

\[
\Lambda^2_{\mathbb{R}, d}(T) = \{ \Phi = (\Phi^1, \ldots, \Phi^d) : \Phi^i \in \Lambda^2_{\mathbb{R}, d}(T), \ i = 1, \ldots, d \}.
\]

For \( \Phi \in \Lambda^2_{\mathbb{R}, d} \) we define \( M = (M_t)_{t \in \mathbb{R}_+} \) by

\[
M_t = \int_0^t \Phi^r_s dB_s = \int_0^t \Phi_s \cdot dB_s = \sum_{k=1}^d \int_0^t \Phi^k_s dB^k_s,
\]

where \( \Phi^r \) denotes the transposed vector. Then \( M \in M^{\mathbb{R}}_{\mathbb{C}} \) and if \( \Phi \in \Lambda^2_\mathbb{R} \), \( M \in \mathbb{H}^2_\mathbb{C} \).

**Lemma 2.5.1.** Let \( \Phi, \Psi \in \Lambda^2_{\mathbb{R}, d} \). Define \( M^i = (M^i_t)_{t \in \mathbb{R}_+} \) and \( M^j = (M^j_t)_{t \in \mathbb{R}_+} \) by \( M^i_t = \int_0^t \Phi^i_s dB^i_s \) and \( M^j_t = \int_0^t \Psi^j_s dB^j_s \), \( i, j = 1, \ldots, d \). Then for every \( i \neq j \), \( \langle M^i, M^j \rangle = 0 \).

**Proposition 2.5.2.** Let \( \Phi, \Psi \in \Lambda^2_{\mathbb{R}, d} \). Define \( M = (M_t)_{t \in \mathbb{R}_+} \) and \( N = (N_t)_{t \in \mathbb{R}_+} \) by \( M_t = \int_0^t \Phi^r_s dB_s \) and \( N_t = \int_0^t \Psi^r_s dB_s \). Then

\[
\langle M, N \rangle_t = \sum_{k=1}^d \int_0^t \Phi^k_s \Psi^k_s ds.
\]

2.6 Itô’s Formula

Let \( B \) be an \( \mathbb{R}^d \)-valued \((\mathcal{F}_t)_{t \in \mathbb{R}_+}\)-Brownian motion on \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})\).

**Definition 2.6.1.** An Itô process \( X = (X_t)_{t \in \mathbb{R}_+} \) is a real-valued process of the form

\[
X_t = X_0 + \int_0^t \alpha_s ds + \int_0^t \Phi_s \cdot dB_s, \ t \geq 0,
\]

or in short,

\[
dX_t = \alpha_t dt + \Phi^r_t dB_t,
\]

where \( X_0 \) is \( \mathcal{F}_0 \)-measurable, \( \Phi \in \Lambda^2_{\mathbb{R}, d} \) and \( \alpha_t \) is a progressive process such that for all \( t \geq 0 \), \( \int_0^t |\alpha_s| ds < +\infty \) a.s.
We call \( dX_t = \Phi_t dB_t + \alpha_t dt \) the *stochastic differential* of \( X \). In this section we try to answer the question about how to compute the stochastic differential of \( f(X) \), given \( f \in \mathcal{C}^2(\mathbb{R}) \). As we will see, the answer to this question is provided by *Itô’s formula*. However, first we need some more preparation. Note that if we define \( M = (M_t)_{t \in \mathbb{R}_+} \) and \( V = (V_t)_{t \in \mathbb{R}_+} \) by \( M_t = \int_0^t \Phi_s \cdot dB_s \) and \( V_t = \int_0^t \alpha_s ds \), we may rewrite (2.9) as

\[
X_t = X_0 + V_t + M_t, \quad t \geq 0,
\]

and this decomposition is unique by Theorem 2.2.1. In a more general setting, we call a stochastic process that can be decomposed as a sum of a local martingale (not necessarily continuous) and of a finite variation process a *semimartingale*.

Assume given two Itô processes \( X \) as in (2.9) and \( Y = (Y_t)_{t \in \mathbb{R}_+} \), where

\[
Y_t = Y_0 + \int_0^t \beta_s ds + \int_0^t \Psi_s \cdot dB_s,
\]

\[
y = Y_0 + W_t + N_t, \quad t \geq 0,
\]

with \( W \in \mathcal{V}_C \) and \( N \in \mathcal{M}_{loc}^{loc} \).

**Proposition 2.6.2.** Let \( X,Y \) be two Itô processes with semimartingale decomposition given by (2.10) and (2.11). Given \( t > 0 \), let \( \pi_n : 0 = t_0 < \cdots < t_n^k = t \) be a sequence of partitions of \([0,t]\) such that \( |\pi_n| \xrightarrow{n \to \infty} 0 \). Then

\[
\sum_{i=0}^{k_n-1} \left( X_{t_{i+1}} - X_{t_i} \right) \left( Y_{t_{i+1}} - Y_{t_i} \right) \longrightarrow \langle M,N \rangle_t
\]

in probability.

**Proof.** See Proposition 5.4.2 of Priouret [15]. \( \square \)

Motivated by Proposition 2.6.2 we define

\[
\langle X,Y \rangle_t = \sum_{k=1}^{d} \int_0^t \phi_k \psi^k ds.
\]

Let \( X \) be an Itô process as in (2.9). Given a “suitable” integrand \( \Psi \), we wish to define the integral of \( \Psi \) with respect to \( X \). We introduce

\[
\mathcal{L}^0(X) = \left\{ \Psi : \Psi \text{ progr. and for all } t, \int_0^t \Psi^2_t \parallel \Phi_s \parallel^2 ds + \int_0^t |\Psi_s \alpha_s| ds < \infty \text{ a.s.} \right\}.
\]
with $\|\phi_s\|^2 = \sum_{i=1}^d (\phi_s^i)^2$. For $\Psi \in \mathcal{L}^0(X)$ define

$$
\int_0^t \Psi_s dX_s := \int_0^t \Psi_s \alpha_s ds + \int_0^t \Psi_s \Phi_s^r dB_s,
$$

where $\int_0^t \psi_s \phi_s^r dB_s = \sum_{i=1}^d \psi_s \phi_s^i dB_s$.

**Remark 2.6.3.** Obviously, if for all $t \geq 0$,

$$
\sup_{s \leq t} |\Psi| < \infty \text{ a.s.}
$$

(in particular if $\Psi_s$ a.s. continuous), then $\Psi \in \mathcal{L}^0(X)$.

The next proposition is the analogon of Lebesgue’s dominated convergence theorem for the stochastic integral.

**Proposition 2.6.4.** Let $X$ be an Itô process as in (2.9) and $(U^n)_{n \in \mathbb{N}}$ be a sequence of uniformly bounded processes (in particular $U^n \in \mathcal{L}^0(X)$ for all $n \in \mathbb{N}$ by Remark 2.6.3). If for all $t \geq 0$,

$$
U^n_t \xrightarrow{n \to \infty} U_t \text{ a.s., then}
$$

$$
\int_0^t U^n_s dX_s \xrightarrow{n \to \infty} \int_0^t U_s dX_s \quad \text{(2.12)}
$$

in probability.

**Proof.** i) First we suppose that for all $t \geq 0$,

$$
E \left[ \int_0^t \|\Phi_s\|^2 ds + \int_0^t |\alpha_s| ds \right] < +\infty.
$$

Note that

$$
E \left[ \left( \int_0^t U^n_s \Phi_s^r dB_s - \int_0^t U_s \Phi_s^r dB_s \right)^2 \right] =
$$

$$
E \left[ \int_0^t (U^n_s - U_s)^2 \|\Phi_s\|^2 ds \right] \xrightarrow{n \to \infty} 0,
$$

and that

$$
E \left[ \int_0^t U^n_s \alpha_s ds - \int_0^t U_s \alpha_s ds \right] \leq E \left[ \int_0^t |U^n_s - U_s| \alpha_s ds \right] \xrightarrow{n \to \infty} 0
$$

by Lebesgue’s Theorem. Hence for all $t \geq 0$,

$$
\int_0^t U^n_s dX_s \xrightarrow{n \to \infty} \int_0^t U_s dX_s \text{ in } L^1.
$$
ii) Otherwise consider the sequence of stopping times \( (\tau_n)_{n \in \mathbb{N}} \), where
\[
\tau_n = \inf \left\{ t \geq 0 : \int_0^t (\|\Phi_s\|^2 + |\alpha_s|) \, ds \geq n \text{ a.s.} \right\}.
\]
Then \( \tau_n \nearrow +\infty \) a.s. (Example 1.6.3) and (2.12) follows by applying the first step to the stopped integral as in the proof of Theorem 1.10.5.

\[\Box\]

**Corollary 2.6.5.** Let \( X \) be an Itô process as in (2.9) and let \( U \) be a continuous adapted process. Let \( t > 0 \) and let \( (\pi_n)_{n \in \mathbb{N}} \) be a sequence of partitions \( \pi_n : 0 = t^n_0 < t^n_1 < \cdots < t^n_{k_n} = t \) of \([0,t]\) such that \( \|\pi_n\| \xrightarrow{n \to \infty} 0 \). Then
\[
\sum_{i=0}^{k_n-1} U^n_{t^n_i} \left( X^n_{t^n_{i+1}} - X^n_{t^n_i} \right) \xrightarrow{n \to \infty} \int_0^t U_s \, dX_s \quad (2.13)
\]
in probability.

**Proof.** i) First we suppose that \( U \) is bounded. Define \( (U^n)_{n \in \mathbb{N}} \) by
\[
U^n_t = \sum_{i=0}^{k_n-1} U^n_{t^n_i} \mathbb{1}_{(t^n_i, t^n_{i+1})}(t), \quad t \geq 0.
\]
Then
\[
\sum_{i=0}^{k_n-1} U^n_{t^n_i} \left( X^n_{t^n_{i+1}} - X^n_{t^n_i} \right) = \int_0^t U^n_s \, dX_s
\]
by the definition of stochastic integral. Since \( U \) is continuous, we have that for every \( s \leq t \),
\[
U^n \xrightarrow{n \to \infty} U_s
\]
and \( (U^n_s)_{n \in \mathbb{N}} \) is uniformly bounded. Hence (2.13) follows by Proposition 2.6.4.

ii) In the general case (where \( U \) is not bounded) consider the stopped process \( U^{\tau_n}, n \in \mathbb{N} \), where
\[
\tau_n = \inf \{ s : |U_s| \geq n \}
\]
and the thesis again follows by the previous step.

\[\Box\]

We now introduce a fundamental result.

**Proposition 2.6.6 (Integration by parts).** Let \( X, Y \) be two Itô processes as in (2.9) and (2.11). Then
\[
X_t Y_t - X_0 Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t,
\]
or in differential notation:
\[
dX_t Y_t = X_t dY_t + Y_t dX_t + d\langle X, Y \rangle_t, \quad t \geq 0.
\]
2.6 Itô’s Formula

Proposition 2.6.6 for τ

Proof. We write \( XY = \frac{1}{4} \{(X + Y)^2 - (X - Y)^2\} \), so that we only need to prove that

\[
X_t^2 - X_0^2 = 2 \int_0^t X_s dX_s + \langle X \rangle_t, \quad t \geq 0.
\]

Let \((\pi_n)_{n \in \mathbb{N}}\) be a sequence of partitions of \([0, t]\), \(\pi_n \colon 0 = t_0^n < \cdots < t_k^n = t\), such that \(|\pi_n| \to 0\). Then we rewrite

\[
X_t^2 - X_0^2 = \sum_{i=0}^{k_n-1} \left( X_{t_{i+1}^n}^2 - X_{t_i^n}^2 \right)
= 2 \sum_{i=0}^{k_n-1} X_{t_i^n} \left( X_{t_{i+1}^n} - X_{t_i^n} \right) + \sum_{i=0}^{k_n-1} \left( X_{t_{i+1}^n} - X_{t_i^n} \right)^2.
\]

Then \((a) \to \int_0^t X_s dX_s\) and \((b) \to \langle X \rangle_t\) in probability by Corollary 2.6.5 and Proposition 2.6.2. Hence for all \(t \geq 0\), \(X_t^2 - X_0^2 = 2 \int_0^t X_s dX_s + \langle X \rangle_t \). \(\Box\)

We are finally ready to prove Itô’s formula.

Theorem 2.6.7 (Itô’s formula). Let \(X = (X_t)_{t \in \mathbb{R}_+}\) be given by \(X_t = (X_1^t, \ldots, X_n^t)\), where \((X_i^t)_{t \in \mathbb{R}_+}\) are Itô processes, \(i = 1, \ldots, n\), and \(f \in \mathcal{C}^2(\mathbb{R}^n)\). Then \(f(X)\) is an Itô process and for \(t \geq 0\) we have

\[
\begin{align*}
\frac{df(X_t)}{dt} &= \sum_{j=1}^{n} \int_0^t \frac{\partial f}{\partial x_j}(X_s) dX_s^j + \frac{1}{2} \sum_{j,k=1}^{n} \int_0^t \frac{\partial^2 f}{\partial x_k \partial x_j}(X_s) d\langle X^j, X^k \rangle_s, \\
\text{or in differential notation:} &
\end{align*}
\]

\[
\begin{align*}
df(X_t) &= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(X_t) dX_t^j + \frac{1}{2} \sum_{j,k=1}^{n} \frac{\partial^2 f}{\partial x_k \partial x_j}(X_t) d\langle X^j, X^k \rangle_t.
\end{align*}
\]

Proof. i) If \(f : x \mapsto x_i, 1 \leq i \leq n\), then the result holds.

ii) If the statement holds for \(f\) and \(g\), then it is also true for \(f + g\) and by Proposition 2.6.6 for \(fg\). Then the result holds when \(f\) is a polynomial.

iii) We can always suppose that \(X\) is bounded by introducing the stopping time \(\tau_n(\omega) = \inf\{t \geq 0 : \|X_t(\omega)\| \geq n\}\). Let thus sup \(\|X_t(\omega)\| \leq M\) and set \(K = \{x \in \mathbb{R}^n : \|x\| \leq M\}\). Since \(f \in \mathcal{C}^2(\mathbb{R}^n)\), then there exists a sequence \((f^n)_{n \in \mathbb{N}}\) of polynomials such that \(f^n, \frac{\partial f^n}{\partial x_i}\) and \(\frac{\partial^2 f^n}{\partial x_i \partial x_j}\) converge uniformly on \(K\) to \(f, \frac{\partial f}{\partial x_i}\) and \(\frac{\partial^2 f}{\partial x_i \partial x_j}\).

iv) It is sufficient to prove Itô’s formula for \((f^n)_{n \in \mathbb{N}}\) and then to pass to the limit by Proposition 2.6.4. \(\Box\)
Corollary 2.6.8. Let \( X = (X_t)_{t \in \mathbb{R}_+} \) be given by \( X_t = (X^1_t, \ldots, X^n_t) \), where \( (X^i_t)_{t \in \mathbb{R}_+} \) are Itô processes, \( i = 1, \ldots, n \), and \( f = f(t, x) \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n) \). Then \( (f(t, X_t))_{t \in \mathbb{R}_+} \) is an Itô process and for \( t \geq 0 \) we have

\[
f(t, X_t) - f(0, X_0) = \int_0^t \frac{\partial f}{\partial s}(s, X_s)ds + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(s, X_s)dX^i_s
\]

or in differential notation:

\[
df(t, X_t) = \frac{\partial f}{\partial t}(t, X_t)dt + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t, X_t)dX^i_t + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(t, X_t)d\langle X^i, X^j \rangle_t.
\]

2.7 The Stochastic Exponential

Let \( \Phi \in \mathcal{A}^2_{\text{loc},d} \) and consider the process \( Z = (Z_t)_{t \in \mathbb{R}_+} \) given by

\[
Z_t = Z_t(\Phi) = \exp \left( \int_0^t \Phi^r_s dB_s - \frac{1}{2} \int_0^t \|\Phi_s\|^2 ds \right), \quad t \geq 0,
\]

where \( B \) is a \( d \)-dimensional Brownian motion. This process is called stochastic exponential or Doléans-Dade exponential and is often denoted by

\[
Z_t = \mathcal{E} \left( \int_0^t \Phi^r_s dB_s \right). \quad t \geq 0.
\]

We now investigate some of its properties that will be very useful later on.

**Proposition 2.7.1.** Let \( \Phi \in \mathcal{A}^2_{\text{loc},d} \) and let \( Z \) be defined as in (2.14).

i) \( Z_t = 1 + \int_0^t Z_s \Phi^r_s dB_s \), i.e. \( dZ_t = Z_t \Phi^r_t dB_t \), \( t \geq 0 \), and \( Z \) is a local martingale.

ii) \( Z \) is a positive supermartingale, \( (Z_t)_{t \in \mathbb{R}_+} \) converges a.s. to \( Z_\infty \) and for all \( t \in [0, +\infty) \), \( E[Z_t] \leq 1 \).

iii) Let \( T > 0 \). If \( E[Z_T] = 1 \), then \( Z \) is a martingale on \([0, T]\).

iv) If \( E[Z_\infty] = 1 \), then \( Z_t = E[Z_\infty | \mathcal{F}_t] \) a.s., \( t \geq 0 \), i.e. \( Z \) is closed by \( Z_\infty \).
Proof. i) Let \( t \geq 0 \) and consider the Itô process
\[
X_t = \int_0^t \Phi_s^t dB_s - \frac{1}{2} \int_0^t \|\Phi_s\|^2 ds.
\]
Then
\[
\langle X \rangle_t = \int_0^t \|\Phi_s\|^2 ds.
\]
We apply Itô’s formula with \( f(x) = e^x \). Then
\[
dZ_t = Z_t dX_t + \frac{1}{2} Z_t d\langle X \rangle_t
\]
\[
= Z_t \Phi^t_t dB_t - \frac{1}{2} Z_t \|\Phi_t\|^2 dt + \frac{1}{2} Z_t \|\Phi_t\|^2 dt
\]
\[
= Z_t \Phi^t_t dB_t.
\]
Hence
\[
Z_t = Z_0 + \int_0^t Z_s \Phi^t_s dB_s = 1 + \int_0^t Z_s \Phi^t_s dB_s,
\]
and \( Z \) is a local martingale.
i) Clearly \( Z_t \geq 0, \ t \geq 0 \), and \( Z \) is a positive supermartingale (Lemma 1.10.2).
In particular \( E[Z_t] \) is decreasing in \( t \) and
\[
E[Z_t] \leq E[Z_0] = 1.
\]
By Theorem 1.8.3 \((Z_t)_{t \in \mathbb{R}_+}\) converges a.s. to a finite limit \( Z_\infty \) and again by Fatou’s lemma we have
\[
E[Z_\infty] = E[\lim Z_t] \leq \lim \inf E[Z_t] \leq E[Z_0] = 1.
\]
iii) Let \( 0 < t < T \). Since \( Z \) is a supermartingale, for all \( A \in \mathcal{F}_t \),
\[
E[\mathbbm{1}_A Z_t] \geq E[\mathbbm{1}_A Z_T]. \tag{2.15}
\]
Now suppose there exists \( A \in \mathcal{F}_t \), such that
\[
E[\mathbbm{1}_A Z_t] > E[\mathbbm{1}_A Z_T].
\]
Then
\[
1 = E[Z_0] \geq E[Z_t] = E[\mathbbm{1}_A Z_t] + E[\mathbbm{1}_{A^c} Z_t]
\]
\[
> E[\mathbbm{1}_A Z_T] + E[\mathbbm{1}_{A^c} Z_T] = E[Z_T] = 1,
\]
and this is impossible! Therefore for all \( A \in \mathcal{F}_t \), \( E[\mathbbm{1}_A Z_t] = E[\mathbbm{1}_A Z_T] \) and \( Z \) is a martingale for \( t \in [0, T] \).
iv) For \( T = +\infty \) it is sufficient to note that (2.15) follows once more by Fatou’s lemma and then to proceed in the same way as before. \( \square \)
As we will see soon, one is often interested in establishing when $E[Z_T] = 1$ and, if possible, when $E[Z_\infty] = 1$. This second case is of course more difficult.

For example, consider $e^{B_t - \frac{1}{2}t}$.

By Proposition 1.3.2 we have

$$B_t - \frac{1}{2}t = -t \left( \frac{1}{2} - \frac{B_t}{t} \right) \to -\infty,$$

hence

$$Z_t \uparrow \infty = 0.$$

This means we have $E[Z_T] = 1$, but $E[Z_\infty] = 0$.

The most common criterion for checking whether $E[Z_T] = 1$, $T > 0$, is the Novikov condition.

**Proposition 2.7.2 (Novikov criterion).** Let $Z$ be defined as in (2.14) and $0 \leq T \leq +\infty$. If

$$E\left[ \exp\left\{ \frac{1}{2} \int_0^T \|\Phi_s\|^2 ds \right\} \right] < +\infty,$$

then $E[Z_T] = 1$.

**Proof.** See Priouret [15], Protter [17].

### 2.8 Girsanov’s Theorem

Consider the Itô process $X = (X_t)_{t \in \mathbb{R}_+}$, where

$$X_t = X_0 + \int_0^t \alpha_s ds + \int_0^t \Phi_s^\tau dB_s.$$

If $(\alpha_t)_{t \in \mathbb{R}_+} \equiv 0$, then $X$ is a (local) martingale. Otherwise (in general) $X$ is not a martingale. A natural question is then if one can change the probability on $(\Omega, \mathcal{F})$ to transform $X$ into a martingale. The answer is provided by Girsanov’s theorem. As we will see, here a key role is played by the Doléans-Dade exponential. Let $\Phi \in A_{\mathbb{P}, \mathbb{Q}, \mathcal{F}}$ and $Z$ be defined as in (2.14), furthermore assume that $E[Z_\infty] = 1$. Then by Proposition 2.7.1 we have that $Z$ is a martingale and $Z_t = E[Z_\infty | \mathcal{F}_t]$ a.s., $t \geq 0$. Consider the probability measure $\mathbb{Q}$ with Radon-Nikodym density$^\dagger$

$^\dagger$ Recall that in this case for all $A \in \mathcal{F}$

$$Q(A) = \int_A d\mathbb{Q} = \int_A Z_\infty d\mathbb{P} = E[\mathbb{1}_A Z_\infty].$$
2.8 Girsanov’s Theorem

\[ \frac{dQ}{dP} = Z_\infty. \]

(Why is this density well-defined?) Then for \( A \in \mathcal{F}_t \)

\[ \int_A Z_\infty \, dP = \int_A Z_t \, dP, \]

i.e.

\[ \frac{d(Q|_{\mathcal{F}_t})}{d(P|_{\mathcal{F}_t})} = Z_t. \]

**Lemma 2.8.1.** A process \((X_t)_{t \in \mathbb{R}_+}\) is a \(Q\)-(local) martingale if and only if \((X_tZ_t)_{t \in \mathbb{R}_+}\) is a \(P\)-(local) martingale.

**Proof.** In the martingale case, the result is straightforward, since \(E^Q[|X_t|] = E[|X_t|Z_t]\), and for \( s < t \) and \( A \in \mathcal{F}_s \) we have

\[ \int_A X_s \, dQ = \int_A X_s Z_s \, dP = \int_A X_t Z_t \, dP = \int_A X_t \, dQ. \]

Let now \( X_tZ_t \) be a \(P\)-local martingale with a reducing sequence of stopping times \( \tau_n, n \in \mathbb{N} \). Since \( Q \sim P \)

\[ Q\left( \lim_{n \to \infty} \tau_n = +\infty \right) = 1 \]

and we have to show that \( X_{t \wedge \tau_n} \) is a \(Q\)-martingale. Let for \( s < t, A \in \mathcal{F}_s \). Then \( A \cap \{ \tau_n > s \} \in \mathcal{F}_{s \wedge \tau_n} \) and

\[ \int_A X_{t \wedge \tau_n} \, dQ = \int_{A \cap \{ \tau_n > s \}} X_{t \wedge \tau_n} \, dQ + \int_{A \cap \{ \tau_n \leq s \}} X_{t \wedge \tau_n} \, dQ \]

\[ = \int_{A \cap \{ \tau_n > s \}} X_{t \wedge \tau_n} Z_{t \wedge \tau_n} \, dP + \int_{A \cap \{ \tau_n \leq s \}} X_{\tau_n} \, dQ \]

\[ = \int_{A \cap \{ \tau_n > s \}} X_{s \wedge \tau_n} Z_{s \wedge \tau_n} \, dP + \int_{A \cap \{ \tau_n \leq s \}} X_{\tau_n} \, dQ \]

\[ = \int_A X_{s \wedge \tau_n} \, dQ. \] (2.17)

\[ \square \]

We now introduce Girsanov’s well-known theorem.

**Theorem 2.8.2.** Let \( Z \) be defined as in (2.14) and assume \( E[Z_\infty] = 1 \). Consider \( Q \sim P \) with \( \frac{dQ}{dP} = Z_\infty \). Then the process \( \tilde{B} = (\tilde{B}_t)_{t \in \mathbb{R}_+}, \) where

\[ \tilde{B}_t = B_t - \int_0^t \Phi_s \, ds, \quad t \geq 0, \]

is an \((\mathcal{F}_t)_{t \in \mathbb{R}_+}\)-Brownian motion under \( Q \).
For the proof of Theorem 2.8.2 we need the following help lemma:

**Lemma 2.8.3.** Let $B$ be an adapted, a.s. continuous $\mathbb{R}^d$-valued process with $B_0 = 0$. Then $B$ is an $\mathcal{F}_t$-Brownian motion if and only if for all $\alpha \in \mathbb{R}^d$

$$Z^{\alpha}_t := \exp \left\{ i \alpha \cdot B_t + \frac{1}{2} \| \alpha \|^2 t \right\}$$

fulfills

$$E \left[ e^{i \alpha \cdot (B_t - B_s)} \mid \mathcal{F}_s \right] = e^{-\frac{1}{2} \| \alpha \|^2 (t-s)}, \quad (2.18)$$

(i.e. $Z^{\alpha}_t$ is a complex-valued $\mathcal{F}_t$-martingale).

**Proof.** Let $B$ be a Brownian motion. Then, since $(B_t - B_s)$ is independent of $\mathcal{F}_s$ and $N_d(0, (t-s)I_d)$ distributed,

$$E \left[ e^{i \alpha \cdot (B_t - B_s)} \mid \mathcal{F}_s \right] = E \left[ e^{i \alpha \cdot (B_t - B_s)} \right] = e^{-\frac{1}{2} \| \alpha \|^2 (t-s)},$$

and hence $Z^{\alpha}_t$ fulfills (2.18).

If for all $s < t$ and $\alpha \in \mathbb{R}^d$ we have

$$E \left[ e^{i \alpha \cdot (B_t - B_s)} \mid \mathcal{F}_s \right] = e^{-\frac{1}{2} \| \alpha \|^2 (t-s)},$$

we also have

$$E \left[ e^{i \alpha \cdot (B_t - B_s)} \right] = e^{-\frac{1}{2} \| \alpha \|^2 (t-s)}.$$

Thus the conditional characteristic function is equal to the characteristic function of a $N_d(0, (t-s)I_d)$ random variable, and hence, $(B_t - B_s)$ is independent of $\mathcal{F}_s$ and $N_d(0, (t-s)I_d)$ distributed. \(\square\)

**Proof (Girsanov Theorem).** By Lemma 2.8.3 we show:

$$M_t := \exp \left\{ i \alpha \cdot \tilde{B}_t + \frac{1}{2} \| \alpha \|^2 \right\}$$

is a complex valued $\mathbb{Q}$-martingale for all $\alpha \in \mathbb{R}^d$. By Remark 1.10.4, since $\sup_{t \leq T} M_t$ is bounded for all $T \geq 0$, it suffices to show that $M$ is a $\mathbb{Q}$-local martingale if and only if $MZ$ is a $\mathbb{P}$-local martingale (Lemma 2.8.1). We have

$$M_t Z_t = \exp \left\{ i \alpha \cdot B_t - \int_0^t i \alpha \cdot \Phi_s dB_s + \frac{1}{2} \| \alpha \|^2 t + \int_0^t \Phi^r_s dB_s - \frac{1}{2} \int_0^t \| \Phi_s \|^2 ds \right\}$$

$$= \exp \left\{ \int_0^t (\Phi_s + i \alpha)^r dB_s - \frac{1}{2} \sum_{k=1}^d \int_0^t (\Phi^k_s + i \alpha^k)^2 dB_s \right\}$$

which is a $\mathbb{P}$-local martingale (Lemma 2.7.1). \(\square\)
We reformulate Definition 1.4.6 (Brownian motion) for a finite horizon \([0, T]\) as follows.

**Definition 2.8.4.** An \((\mathcal{F}_t)_{t \in [0, T]}\)-adapted process \(B = (B_t)_{t \in [0, T]}\) with values in \(\mathbb{R}^d\) on \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, P)\) is a standard \((\mathcal{F}_t)_{t \in [0, T]}\)-Brownian motion on \([0, T]\) if

i) \(B_0 = 0\) a.s.,

ii) for every \(s < t \leq T\), \(B_t - B_s\) is independent of \(\mathcal{F}_s\),

iii) for every \(s < t \leq T\), \(B_t - B_s \sim N(0, (t-s)I_d)\),

iv) \(B\) has continuous sample paths.

We can restate Theorem 2.8.2 in the case of a finite horizon.

**Theorem 2.8.5 (Girsanov’s theorem).** Let \(B_t = (B_t)_{t \in [0, T]}\) be an \((\mathcal{F}_t)_{t \in [0, T]}\)-Brownian motion of dimension \(d\) on \([0, T]\) and \((\Phi_t)_{t \in [0, T]} \in \mathcal{A}^2_{loc,d}(T)\). Let

\[
Z_T = \exp \left( \int_0^T \Phi_t dB_t - \frac{1}{2} \int_0^T \|\Phi_t\|^2 ds \right)
\]

and suppose \(E[Z_T] = 1\). Then \((\tilde{B}_t)_{t \in [0, T]}\), where

\[
\tilde{B}_t = B_t - \int_0^t \Phi_s ds,
\]

is an \((\mathcal{F}_t)_{t \in [0, T]}\)-Brownian motion on \([0, T]\) under \(Q\), where

\[
\frac{dQ}{dP} = Z_T.
\]

We show now how we can use Girsanov’s Theorem to transform an Itô process into a martingale. Consider \(d = 1\) and the Itô process \(X = (X_t)_{t \in \mathbb{R}_+}\) given by

\[
X_t = X_0 + \int_0^t \alpha_s ds + \int_0^t \Psi_s dB_s. \tag{2.19}
\]

Suppose that for all \(t \geq 0\), \(\Psi_t, \alpha_t \neq 0\) a.s., i.e. \(X\) is not a martingale. Let \(\Phi \in \mathcal{A}^2_{loc}(T)\) such that \(E[Z_T] = 1\), where

\[
Z_T = E \left( \int \Phi_s dB_s \right)_{T}.
\]

By Theorem 2.8.5 we have that \((\tilde{B}_t)_{t \in [0, T]}\) given by

\[
\tilde{B}_t = B_t - \int_0^t \Phi_s ds \tag{2.20}
\]

is an \((\mathcal{F}_t)_{t \in [0, T]}\)-Brownian motion on \([0, T]\) under \(Q\), where

\[
\frac{dQ}{dP} = Z_T.
\]
We substitute (2.20) in (2.19) and obtain
\[
dX_t = \alpha_t dt + \Psi_t d\tilde{B}_t \\
= \alpha_t dt + \Psi_t (d\tilde{B}_t + \Phi_t dt) \\
= (\alpha_t + \Psi_t \Phi_t) dt + \Psi_t d\tilde{B}_t.
\]
If \( \Phi_t = -\frac{\alpha_t}{\Psi_t} \in A^2_{loc}(T) \) and \( E[E(\phi \cdot B)_T] = 1 \), then \( X \) becomes a local martingale under \( Q \).

2.9 Brownian Martingales

Let \( B \) be a \( d \)-dimensional Brownian motion and \( (F^B_t)_{t \in \mathbb{R}_+} \) its natural filtration, i.e. if \( F^0_t := \sigma(B_s, s \leq t) \), \( F^0_\infty := \sigma(B_s, s \geq 0) \), and \( N \) is the set of measure zero sets of \( F^0_\infty \) then we have
\[
F^B_t := \sigma(F^0_t, N), \quad F^B_\infty := \sigma(F^0_\infty, N).
\]
Remind: \( (F^B_t)_{t \in \mathbb{R}_+} \) is standard, i.e. right-continuous and complete ((\( F^0_t \))\( _{t \in \mathbb{R}_+} \) is neither right-continuous nor complete).

**Proposition 2.9.1 (Structure of random variables in \( L^2(\Omega, F^B_\infty, \mathbb{P}) \)).**

Let \( X \in L^2(\Omega, F^B_\infty, \mathbb{P}) \). Then there exists a unique \( \Phi \in A^2_d \) such that
\[
X = E[X] + \int_0^\infty \Phi_s dB_s
\]

**Proof.** (For notational simplicity we show the case \( d = 1 \)).

Let \( H := \left\{ X \in L^2(\Omega, F^B_\infty, \mathbb{P}) \mid X = a + \int_0^\infty \Phi_s dB_s, \Phi \in A^2 \right\} \).

Then \( H \) is a closed vector space in \( L^2(\Omega, F^B_\infty, \mathbb{P}) \) (Itô-Isometry). For \( f \in L^2(\mathbb{R}_+, ds) \) let
\[
M_t(f) := \exp \left\{ \int_0^t f(s) dB_s - \frac{1}{2} \int_0^t f^2(s) ds \right\} \\
= \mathcal{E}_t \left( \int f(s) dB_s \right).
\]

We show \( M_\infty(f) \in H \): We have
\[
E[M^2_\infty(f)] = \exp \left\{ \int_0^\infty f^2(s) ds \right\} E[M_t(2f)].
\]

Since \( E[M_t(2f)] \leq 1 \) (Proposition 2.7.1),
\[ E [M_t^2(f)] \leq \exp \left\{ \int_0^t f_s^2 \, ds \right\} \]

for every \( t \geq 0 \). We know (Proposition 2.7.1)

\[ M_t(f) = 1 + \int_0^t M_s(f) f(s) \, dB_s \]

for all \( t \geq 0 \). But since

\[ E \left[ \int_0^\infty M_s^2(f) f^2(s) \, ds \right] \leq \exp \left\{ \int_0^\infty f^2(s) \, ds \right\} \int_0^\infty f^2(s) \, ds < \infty \]

it follows \( M_\infty(f) \in H \).

We show \( H = L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}) \): Let \( Y \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}) \) be such that

\[ E[Y Z] = 0 \]

for all \( Z \in H \). We have to show \( Y = 0 \) a.s. We have for all \( f \in L^2(\mathbb{R}_+, ds) \)

\[ E[Y M_\infty(f)] = 0 \]

and therefore

\[ E \left[ Y \exp \left( \sum_{k=1}^n c_k B_{t_k} \right) \right] = 0 \]

for every \( c_k \in \mathbb{R}, t_k \in \mathbb{R}_+ \) and \( n \in \mathbb{N} \) (with \( f = \sum_{k=1}^n c_k \mathbb{1}_{[0,t_k]} \)). With

\[ g(B_{t_1}, \ldots, B_{t_n}) = E[Y | B_{t_1}, \ldots, B_{t_n}] \]

we get

\[ E \left[ g(B_{t_1}, \ldots, B_{t_n}) \exp \left( \sum_{k=1}^n c_k B_{t_k} \right) \right] = 0, \]

i.e.

\[ \int g(x_1, \ldots, x_n) \exp \left( \sum_{k=1}^n c_k x_k \right) h(x_1, \ldots, x_n) \, dx_1 \ldots dx_n = 0 \]

for every \( c_k \in \mathbb{R} \), where \( h \) is the density of \((B_{t_1}, \ldots, B_{t_n})\). But since the linear span of

\[ \left\{ \exp \left( \sum_{k=1}^n c_k x_k \right) : c_k \in \mathbb{R} \right\} \]

is dense in \( L^1(\mathbb{R}^n, dx_1 \ldots dx_n) \), we get

\[ g h = 0 \text{ a.e.} \]

and since \( h > 0 \), \( g = 0 \text{ a.e.} \), and thus

\[ E[Y | B_{t_1}, \ldots, B_{t_n}] = 0 \text{ a.s.} \]
Thus for all \(A \in \sigma(B_{t_1}, \ldots, B_{t_n})\) we have \(E[\mathbb{1}_A Y] = 0\), from which we infer
\[
E[\mathbb{1}_A Y] = 0 \quad \text{for all } A \in \mathcal{F}^B_{\infty}
\]
and hence
\[
Y = 0 \text{ a.s.}
\]

\[\square\]

**Theorem 2.9.2 (Martingale representation Theorem).**

Let \(M\) be a \(\mathcal{F}^B_t\)-local martingale. Then there exists a unique \(\Phi \in \Lambda^2_{loc,d}\) such that
\[
M_t = M_0 + \int_0^t \Phi_s dB_s \text{ a.s.}
\]
(in particular every Brownian local martingale has a continuous modification).

If \(M\) is a square integrable martingale then for all \(t \geq 0\)
\[
E \left[ \int_0^t \|\Phi_s\|^2 ds \right] < \infty.
\]

**Proof.** (Again we restrict ourself to the case \(d = 1\))

i) Let \(M\) be a right-continuous martingale such that \(M_0 = 0\) and
\[
E \left[ \sup_t M_t^2 \right] < \infty.
\]

Then \(M_t \overset{L^2}{\to} M_\infty \in L^2(\Omega, \mathcal{F}_\infty^B, \mathbb{P})\) (Corollary 1.8.4).
From Proposition 2.9.1 we get \(\Phi \in \Lambda^2\) such that
\[
M_\infty = \int_0^\infty \Phi_s dB_s.
\]
Let \(N_t := \int_0^t \Phi_s dB_s\). Then \(N_\infty = M_\infty\) a.s. and
\[
M_t = E \left[ M_\infty \mid \mathcal{F}^B_t \right] = E \left[ N_\infty \mid \mathcal{F}^B_t \right] = N_t \text{ a.s.}
\]
Uniqueness of \(\Phi_s\) follows from the Itô isometry. Considering \(M^T_t, T \geq 0\), we obtain the result for all square integrable right-continuous martingales.

ii) Let \(M\) be a closed martingale, i.e.
\[
M_t = E \left[ M_\infty \mid \mathcal{F}^B_t \right].
\]
Because \((\mathcal{F}^B_t)_{t \in \mathbb{R}_+}\) is right-continuous one can show that \(M\) has a right-continuous modification. From that and i), one can again infer that every \(\mathcal{F}^B_t\)-local martingale has a continuous modification.
iii) Let $M$ be a continuous $\mathcal{F}_t^B$-local martingale, $M_0 = 0$, and

$$\tau_n := \inf\{t \geq 0 : |M_t| \geq n\}.$$ 

Applying i) to $M^\tau_n$ gives

$$M_{t \wedge \tau_n} = \int_0^{t \wedge \tau_n} \Phi^n_s dB_s.$$ 

From uniqueness in i) it follows

$$\Phi^n_s = \Phi^{n+1}_s$$

for all $s \leq \tau_n$. We can thus take $\Phi_s := \sum_n \Phi^{n+1}_s 1_{(\tau_n, \tau_{n+1})}(s)$. Uniqueness follows by localizing.
Stochastic Differential Equations

We summarize here without proofs some important results concerning stochastic differential equations. The main references for this chapter are Karatzas and Shreve [9], Priouret [15], Øksendal [14] and Protter [17].

Given two functions $\mu, \sigma$, the problem which we want to analyze now, is to find a stochastic process $\Phi$ that solves the stochastic differential equation (SDE)

$$d\Phi_t = \mu(\Phi_t)dt + \sigma(\Phi_t)dB_t,$$

where $B$ is a Brownian motion. We remark that all filtrations considered in the following are assumed to be complete.

3.1 Definition of Stochastic Differential Equations

From now on for $A \in \mathbb{R}^{n \times d}$, we define $\|A\| = \|\text{vec}(A)\|$, where as usual $\|\cdot\|$ denotes the Euclidean norm.

**Definition 3.1.1.** Let $B$ be an $\mathbb{R}^d$-valued $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$-Brownian motion on a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$, $\Phi$ a progressive process with values in $\mathbb{R}^{n \times d}$ such that $\|\Phi\| \in A_{loc}^2$, i.e. for all $t \geq 0$,

$$\int_0^t \sum_{i=1}^n \sum_{k=1}^d (\Phi_{i,k}^s)^2 ds < +\infty,$$

and $\alpha$ a progressive process with values in $\mathbb{R}^n$, such that $\sqrt{\|\alpha\|} \in A_{loc}^2$, i.e. for all $t \geq 0$,

$$\int_0^t \sqrt{\sum \alpha_i^s}^2 ds < +\infty.$$

Then the $\mathbb{R}^n$-valued process $(X_t)_{t \in \mathbb{R}_+} = (X_1^t, \ldots, X_n^t)_{t \in \mathbb{R}_+}$ given by

$$X_i^t = X_0^i + \int_0^t \alpha_i^s ds + \int_0^t \Phi_i^s \cdot dB_s, \quad i = 1, \ldots, n,$$
where $X^i_0$ is an $\mathcal{F}_0$-measurable random variable and $\int_0^t \Phi^i_s dB_s = \sum_{k=1}^d \int_0^t \Phi^i_{s,k} dB^k_s$, is called a vector Itô process. We write

$$X_t = X_0 + \int_0^t \alpha_s ds + \int_0^t \Phi_s dB_s.$$

Let now $\mu$ and $\sigma$ be two measurable maps,

\begin{align*}
\mu : \mathbb{R}_+ \times \mathbb{R}^n &\rightarrow \mathbb{R}^n, \quad (3.1) \\
\sigma : \mathbb{R}_+ \times \mathbb{R}^n &\rightarrow \mathbb{R}^{n \times d}. \quad (3.2)
\end{align*}

We want to analyze solutions of the **stochastic differential equation**

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dB_t. \quad (SDE(\mu, \sigma))$$

**Definition 3.1.2 (Solution).** Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, B_t, \mathbb{P})$ be a given stochastic basis where $B$ is an $\mathbb{R}^d$-valued $\mathcal{F}_t$-Brownian motion. A $\mathbb{R}^n$-valued stochastic process $X$ is called a solution of the SDE($\mu, \sigma$) relative to $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, B_t, \mathbb{P})$ if $X$ is continuous and $\mathcal{F}_t$-adapted and fulfills

$$X_t = X_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s,$$

(here the existence of the integrands is implied in the definition).

**Definition 3.1.3 (Strong and weak solutions).**

i) If for any given stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, B_t, \mathbb{P})$ and $\mathbb{R}^n$-valued $\mathcal{F}_0$-measurable random variable $\eta$ there exists a solution $X$ of SDE($\mu, \sigma$) relative to $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, B_t, \mathbb{P})$ such that $X_0 = \eta$, we say there exists a strong solution of the SDE($\mu, \sigma$).

ii) If for any $\mathbb{R}^n$-valued $\mathcal{F}_0$-measurable random variable $\eta$ there exists a solution $X$ of SDE($\mu, \sigma$) relative to some stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, B_t, \mathbb{P})$ such that $X_0 \overset{d}{=} \eta$, i.e. $X_0$ and $\eta$ are equal in distribution, we say there exists a weak solution of the SDE($\mu, \sigma$).

**Remark 3.1.4.** i) Existence of strong solution implies existence of weak solution.

ii) Putting $\mathcal{F}_t = \mathcal{G}_t$ in Definition 3.1.3 i), where $\mathcal{G}_t$ is the filtration given by the completion of $(\sigma(\eta) \lor \mathcal{F}^B_t)_{t \in \mathbb{R}_+}$, one sees: "Strong solutions exist if and only if $\mathcal{G}_t$-adapted solutions exit".

One could thus substitute $\mathcal{F}_t$ with $\mathcal{G}_t$ in Definition 3.1.3 i). Note that if $\eta \in \mathbb{R}$ deterministic then $\mathcal{G}_t = \mathcal{F}^B_t$. 
3.2 Existence and Uniqueness of Solutions

Definition 3.1.5 (Pathwise and weak uniqueness).

i) We say there is pathwise uniqueness of solutions of $SDE(\mu, \sigma)$ if for any two solutions $X, \tilde{X}$ of $SDE(\mu, \sigma)$ relative to some $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}^+}, B_t, \mathbb{P})$ such that $X_0 = \tilde{X}_0$ a.s. we have

$$X_t = \tilde{X}_t \text{ for all } t \geq 0 \text{ a.s.}$$

ii) We say there is weak uniqueness of solutions of $SDE(\mu, \sigma)$ if for any two solutions $X$ relative to $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}^+}, B_t, \mathbb{P})$ and $\tilde{X}$ relative to $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \in \mathbb{R}^+}, \tilde{B}_t, \tilde{\mathbb{P}})$ such that $X_0 \sim \tilde{X}_0$ the process $X$ and $\tilde{X}$ are equivalent (i.e. equal in distribution).

3.2 Existence and Uniqueness of Solutions

Now we investigate some conditions (also referred to as Itô conditions) under which the solution of an SDE exists.

Definition 3.2.1. 

i) We say that $\mu, \sigma$ have sub-linear growth, if for every $T > 0$ there exists a constant $M(T)$, such that for all $t \leq T, x \in \mathbb{R}^n$,

$$||\mu(t, x)|| + ||\sigma(t, x)|| \leq M(T)(1 + ||x||). \quad (3.3)$$

ii) We say that $\mu, \sigma$ are Lipschitz, if for every $T > 0$ there exists a constant $L(T)$, such that for all $t \leq T, x, y \in \mathbb{R}^n$,

$$||\mu(t, x) - \mu(t, y)|| + ||\sigma(t, x) - \sigma(t, y)|| \leq L(T)||x - y||. \quad (3.4)$$

Theorem 3.2.2 (Existence and uniqueness). Let $\mu : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^{n \times d}$ be two measurable maps satisfying conditions (3.3) and (3.4) of Definition 3.2.1. Then there exists a pathwisely and weakly unique strong solution of the SDE($\mu, \sigma$).

Proof. For the proof of the uniqueness we refer to Karatzas and Shreve [9], Priouret [15], Øksendal [14] and Protter [17].

For the existence we recall:

Fixed-Point Lemma: Let $(E, ||\cdot||)$ be a Banach space and

$$U : E \to E$$

a map such that

$$||U(x) - U(y)|| \leq \rho||x - y||, \quad \rho < 1.$$  

Then there exists one and only one $\bar{x} \in E$ such that $\bar{x} = U(\bar{x})$ and

$$U^n(x_0) \xrightarrow{n \to \infty} \bar{x} \quad \text{for all } x_0 \in E.$$
Let now $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, B_t, \mathbb{P})$, $\mathcal{F}_0$-measurable $\eta$, and $T > 0$ be given and consider the Banach space

$$H := L^2_m([0, T] \times \Omega, \mathbb{P}, dt \otimes \mathbb{P})$$

with norm

$$\|\Phi\|_H = \left( E \left[ \int_0^T \|\Phi_t\|^2 dt \right] \right)^{\frac{1}{2}}, \Phi \in H.$$ Define for $\Phi \in H$

$$S(\Phi)_t := \eta + \int_0^t \mu(s, \Phi_s)ds + \int_0^t \sigma(s, \Phi_s)dB_s, \ t \leq T.$$ Then $S$ is a map from $H$ to $H$ (exercise) and for $\Phi, \Psi \in H$, $t \leq T$,

$$E \left[ \|S(\Phi)_t - S(\Psi)_t\|^2 \right] \leq 2E \left[ \| \int_0^t (\sigma(s, \Phi_s) - \sigma(s, \Psi_s))dB_s \|^2 \right]$$

$$+ 2E \left[ \| \int_0^t (\mu(s, \Phi_s) - \mu(s, \Psi_s))ds \|^2 \right]$$

$$\leq 2E \left[ \int_0^t \| (\sigma(s, \Phi_s) - \sigma(s, \Psi_s)) \|^2 ds \right]$$

$$+ 2TE \left[ \int_0^t \| (\mu(s, \Phi_s) - \mu(s, \Psi_s)) \|^2 ds \right]$$

$$\leq C(T) \int_0^t E \left[ \| \Phi_s - \Psi_s \|^2 \right] ds$$

where we used the Lipschitz assumption in the last inequality with $C(T) = 2(1 + T)L(T)$.

Now consider the to $\| \cdot \|_H$ equivalent norm on $H$

$$\|\Phi\|_{H,c} := \left( E \left[ \int_0^T e^{-ct} \|\Phi_t\|^2 dt \right] \right)^{\frac{1}{2}}, c > 0.$$ Then

$$\|S(\Phi) - S(\Psi)\|_{H,c}^2 = \int_0^T E \left[ \|S(\Phi)_t - S(\Psi)_t\|^2 \right] e^{-ct} dt$$

$$\leq C(T) \int_0^T \left\{ \int_0^t E[\|\Phi_s - \Psi_s\|^2]ds \right\} e^{-ct} dt$$

$$= C(T) \int_0^T E[\|\Phi_s - \Psi_s\|^2] \left\{ \int_s^T e^{-ct} dt \right\} ds$$

$$\leq \frac{C(T)}{c} \int_0^T E[\|\Phi_s - \Psi_s\|^2]e^{-cs} ds.$$
3.3 Markov Property of Solutions

Let be $c > 0$ such that $C(T) < 1$. Thus $S$ fulfills the assumptions of the Fixed-Point Lemma and there exists one and only one $\Phi \in H$ such that

$$S(\Phi) = \Phi.$$ 

A continuous representative $X$ of $\Phi$ (which exists because $S(\Phi)$ is continuous) is thus a solution of SDE($\mu$, $\sigma$) relative to $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, B_t, \mathbb{P})$. □

Remark 3.2.3. Under the assumptions of Theorem 3.2.2 we denote with $X^\eta$ the solution of SDE($\mu$, $\sigma$) relative to $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, B_t, \mathbb{P})$ that starts in $\eta$, i.e. $X^\eta_0 = \eta$. From pathwise uniqueness we infer that $X^\eta_t$ must be $\mathcal{F}_t^B$-adapted for all $x \in \mathbb{R}^n$.

3.3 Markov Property of Solutions

Let now $\mu, \sigma$ fulfill the assumptions of Definition 3.2.1. Further, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, B_t, \mathbb{P})$ be given and consider the $\mathcal{F}_t^{(u)}$-Brownian motion $B_t^{(u)} := B_{u+t} - B_u$ where $\mathcal{F}_t^{(u)} := \mathcal{F}_{u+t}$, and let $X_{u+h}^{u,\eta}$ be the unique solution of

$$Y_h = \eta + \int_0^h \mu(u + s, Y_s)ds + \int_0^h \sigma(u + s, Y_s)dB_s^{(u)} \quad (3.5)$$

where $\eta$ is $\mathcal{F}_0^{(u)} = \mathcal{F}_u$-measurable. It is not difficult to show that

$$Y_h = \eta + \int_u^{u+h} \mu(s, Y_{s-u})ds + \int_u^{u+h} \sigma(s, Y_{s-u})dB_s$$

so $X_{t}^{u,\eta}$ is the unique solution of

$$X_t^{u,\eta} = \eta + \int_u^t \mu(s, X_{s-u}^{u,\eta})ds + \int_u^t \sigma(s, X_{s-u}^{u,\eta})dB_s, \quad t \geq u,$$

($X_t^{u,\eta}$ solution of stochastic differential equation that starts in $\eta$ in time $u$). Since for $t \geq v \geq u \geq 0$

$$X_t^{u,\eta} = \eta + \int_u^v \mu(s, X_{s-u}^{u,\eta})ds + \int_u^v \sigma(s, X_{s-u}^{u,\eta})dB_s$$

$$+ \int_v^t \mu(s, X_{s-u}^{u,\eta})ds + \int_v^t \sigma(s, X_{s-u}^{u,\eta})dB_s$$

$$= X_v^{u,\eta} + \int_v^t \mu(s, X_{s-v}^{u,\eta})ds + \int_v^t \sigma(s, X_{s-v}^{u,\eta})dB_s,$$
we obtain by uniqueness
\[ X_t^{u,\eta} = X_t^{v,X_{v}^{u,\eta}}. \]
In particular for \( u = 0 \):
\[ X_t^\eta = X_t^{v,X_{v}^\eta}. \]
From this we can infer that Itô solutions are Markov processes (actually, they are diffusions and therefore often referred to as Itô diffusions):

**Proposition 3.3.1.** Let \( \mu, \sigma \) be as in Definition 3.2.1 and \( X_t^{u,\eta} \) the unique solution as above. Then for all \( u \leq t \leq t+h, \mathcal{F}_t \)-measurable \( \eta \), and \( f: \mathbb{R}^n \to \mathbb{R} \) Borel measurable and bounded
\[
E \left[ f \left( X_{t+h}^{u,\eta} \right) \bigg| \mathcal{F}_t \right] = E \left[ f \left( X_{t+h}^{\eta} \right) \right] \bigg|_{x=X_{t}^{u,\eta}}.
\]
In particular, if \( X \) is time homogeneous (i.e. \( \mu \) and \( \sigma \) do not depend on \( t \)):
\[
E \left[ f \left( X_{t+h}^{\eta} \right) \bigg| \mathcal{F}_t \right] = E \left[ f \left( X_{t+h}^{\eta} \right) \right] \bigg|_{x=X_{t}^{\eta}}.
\]
**Proof.** Set
\[
F(x, u, t, \omega) := X_{t}^{u,x}(\omega).
\]
By Remark 3.2.3 and (3.5) is \( F(x, u, \cdot, \cdot) \mathcal{B}_t^{(u)} \)-adapted and thus independent of \( \mathcal{F}_u \). Further, one can show that for \( \eta \mathcal{F}_u \)-measurable
\[
X_{t}^{u,\eta}(\omega) = F(\eta, u, t, \omega).
\]
Thus
\[
E \left[ f \left( X_{t+h}^{u,\eta} \right) \bigg| \mathcal{F}_t \right] = E \left[ f \left( F(\eta, u, t + h, \omega) \right) \bigg| \mathcal{F}_t \right] \\
= E \left[ f \left( F(X_{t+h}^{\eta,\cdot}, t + h, \omega) \right) \bigg| \mathcal{F}_t \right] \\
= E \left[ f \left( F(x, t, t + h, \omega) \right) \right] \bigg|_{x=X_{t}^{u,\eta}} \\
= E \left[ f \left( X_{t+h}^{\cdot,\cdot} \right) \right] \bigg|_{x=X_{t}^{u,\eta}}.
\]
\( \square \)

**Remark 3.3.2.** One can show that Itô diffusions are even **strong Markov processes**, i.e. for any \( \mathcal{F}_t \)-stopping time \( \tau \)
\[
E \left[ f \left( X_{\tau+h}^{u,\eta} \right) \bigg| \mathcal{F}_\tau \right] = E \left[ f \left( X_{\tau+h}^{\cdot,\cdot} \right) \bigg| \mathcal{F}_\tau \right] \bigg|_{x=X_{\tau}^{u,\eta}}.
\]
In particular, if \( X \) is time homogeneous
\[
E \left[ f \left( X_{\tau+h}^{\eta} \right) \bigg| \mathcal{F}_\tau \right] = E \left[ f \left( X_{\tau+h}^{\cdot} \right) \bigg| \mathcal{F}_\tau \right] \bigg|_{x=X_{\tau}^{\cdot}}.
\]
3.4 Examples: Geometric Brownian Motion and Ornstein-Uhlenbeck Process

We now consider some well-known one-dimensional \((n = 1)\) linear stochastic differential equations. Let \(B\) be a real-valued \((\mathcal{F}_t)_{t \in \mathbb{R}_+}\)-Brownian motion on \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})\). Consider the linear stochastic differential equation given by

\[
\begin{aligned}
&\begin{cases}
 \text{d}X_t = X_t(\mu \text{d}t + \sigma \text{d}B_t), & \mu, \sigma \in \mathbb{R}, \\
 X_0 = \eta, & \eta \in \mathbb{R}_+.
\end{cases} \\
&\quad (3.6)
\end{aligned}
\]

By Theorem 3.2.2 there exists a unique solution. If \(\mu = 0\), we already know (see Proposition 2.7.1) that the solution of

\[
\begin{aligned}
&\begin{cases}
 \text{d}X_t = \sigma X_t \text{d}B_t, \\
 X_0 = \eta,
\end{cases} \\
&\quad \text{is given by } X_t = \eta e^{\sigma B_t - \frac{1}{2} \sigma^2 t}. \text{ If } \sigma = 0, \text{ then the solution of}
\end{aligned}
\]

\[
\begin{aligned}
&\begin{cases}
 \text{d}X_t = \mu X_t \text{d}t, \\
 X_0 = \eta,
\end{cases} \\
&\quad \text{is given by } X_t = \eta e^{\mu t}. \text{ We now consider the general case where } \mu, \sigma \neq 0, \text{ and look for a solution of the form}
\end{aligned}
\]

\[
X_t = C_t e^{\mu t},
\]

where \(C\) is an Itô process. By Itô’s formula we have

\[
\begin{aligned}
\text{d}X_t &= e^{\mu t} \text{d}C_t + \mu C_t e^{\mu t} \text{d}t \\
&= e^{\mu t} \sigma C_t \text{d}B_t + \mu C_t e^{\mu t} \text{d}t, \\
&\quad \text{X solves (3.6)}
\end{aligned}
\]

hence \(\text{d}C_t = \sigma C_t \text{d}B_t\) and

\[
X_t = \eta e^{(\mu - \frac{\sigma^2}{2})t + \sigma B_t}, \quad t \geq 0. \quad (3.7)
\]

As a useful exercise the reader may verify that \(X\) given in (3.7) is a solution of (3.6) by using Itô formula. We call \(X\) geometric Brownian motion.

Next, consider the stochastic differential equation

\[
\begin{aligned}
&\begin{cases}
 \text{d}X_t = \lambda(m - X_t) \text{d}t + \sigma \text{d}B_t, & \lambda, m, \sigma \in \mathbb{R}, \\
 X_0 = \eta \in \mathbb{R}
\end{cases} \\
&\quad (3.8)
\end{aligned}
\]

Again there exists a unique solution \(X\) by Theorem 3.2.2. Consider \(Y_t := e^{\lambda t} X_t\). Then
\[ Y_t - Y_0 = \int_0^t \lambda e^{\lambda s} X_s \, ds + \int_0^t e^{\lambda s} dX_s \]
\[ = \int_0^t \lambda Y_s \, ds + \int_0^t (e^{\lambda s} \lambda m - \lambda Y_s) \, ds + \int_0^t e^{\lambda s} \sigma dB_s \]
\[ = \lambda m \int_0^t e^{\lambda s} \, ds + \sigma \int_0^t e^{\lambda s} dB_s \]
\[ = \lambda m \left( \frac{1}{\lambda} (e^{\lambda t} - 1) \right) + \sigma \int_0^t e^{\lambda s} dB_s. \]

It follows that
\[ X_t = e^{-\lambda t} Y_t \]
\[ = \eta e^{-\lambda t} + m - me^{-\lambda t} + \sigma \int_0^t e^{-\lambda(t-s)} dB_s \]
\[ = m + (\eta - m) e^{-\lambda t} + \sigma \int_0^t e^{-\lambda(t-s)} dB_s \]

is indeed the solution of (3.8) (exercise) which is called **Ornstein-Uhlenbeck process**. For \( \lambda > 0 \) the paths of this process are **mean-reverting** to the level \( m \) with mean reversion rate \( \lambda \).
Part II

Continuous Time Finance
4

Financial Markets in Continuous Time

In this chapter we show how to use the results of the first chapters to model financial markets in continuous time. The main references for this part are Karatzas and Shreve [10], Øksendal [14], Musiela and Rutkowski [13], Lamberton and Lapeyre [11] and Björk [1].

We consider an idealized financial market that satisfies the following hypotheses concerning market activities.

- Trading takes place continuously in time.
- Borrowing and lending is possible at the same interest rate with no amount restriction.
- The market is frictionless and there are neither transaction costs nor taxes.
- Short sales are allowed.
- Stocks do not pay dividends.
- The market is liquid.
- The information is shared equally among the traders and any change of it is seen instantaneously by all agents on the market.

4.1 The Market Model

All models we consider here are defined on a finite horizon \([0, T]\), \(T > 0\). Consider a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})\) with a \(d\)-dimensional Brownian motion \(B = (B_t)_{t \in [0, T]}\). If not said otherwise, for the rest of the text we will always assume \(\mathcal{F}_t = \mathcal{F}_B^t\) and \(\mathcal{F} = \mathcal{F}_B^T\). Furthermore we assume that in our market there exist \((n + 1)\) assets represented by an \((n + 1)\)-dimensional vector Itô process \(X = (X_t)_{t \in [0, T]}\), where \(X_t = (X^0_t, X^1_t, \ldots, X^n_t)\). The dynamics of \(X^0\) are given by

\[
dX^0_t = r_t X^0_t dt, \quad X^0_0 = 1, \quad (4.1)
\]

where \(X^0_t\) represents the possibility of depositing money in (or borrowing money from) a bank at a (random) interest rate \(r = (r_t)_{t \in [0, T]}\) at time \(t\).
Therefore $X^0$ is called the *money market account*. The solution of (4.1) is of the form

$$X^0_t = X^0_0 e^{\int_0^t r_s ds} = e^{\int_0^t r_s ds}, \quad t \in [0, T]. \quad (4.2)$$

From now on, for the sake of simplicity we assume that $r$ is a bounded progressive process. For $i = 1, \ldots, n$, $X^i$ is given by

$$\begin{cases}
    dX^i_t = \mu^i_t dt + \sigma^i_t \cdot dB^i_t \\
    X^i_0 = x^i, \quad x^i \in \mathbb{R},
\end{cases} \quad (4.3)$$

or in short notation,

$$dX_t = \mu_t dt + \sigma_t dB_t,$$

where $\mu = (\mu_t)_{t \in [0,T]}$, $\mu_t = (\mu^1_t, \ldots, \mu^n_t)$ is an $\mathbb{R}^n$-valued progressive process, such that $\sqrt{\|\mu\|} \in A^2_{\text{loc}}(T)$, i.e.

$$\int_0^T \sqrt{\sum_{i=1}^n (\mu^i_s)^2} ds < +\infty \quad \text{a.s.}$$

and $\sigma = (\sigma_t)_{t \in [0,T]}$ is an $\mathbb{R}^{n \times d}$-valued progressive process, such that $\|\sigma\| \in A^2_{\text{loc}}(T)$, i.e.

$$\int_0^T \sum_{i=1}^n \sum_{k=1}^d (\sigma^i_{t,k})^2 dt < +\infty \quad \text{a.s.}$$

We refer to $X^1, \ldots, X^n$ as risky assets and define $\hat{X} = (X^1, \ldots, X^n)$. The process $\mu^i$ is called the *drift* and $\sigma^i$ the *volatility* of the asset price $X^i$.

**Remark 4.1.1.** Often one assumes that asset prices are *strictly positive* *Itô* processes $X^i, i = 0, \ldots, n$. By Itô’s formula we then obtain with $Y^i_t = \log(X^i_t)$

$$dY^i_t = \frac{1}{X^i_t} dX^i_t - \frac{1}{2 (X^i_t)^2} (dX^i)_t = \left( \frac{\mu^i_t}{X^i_t} - \frac{1}{2} (\sigma^i_t)^2 \right) dt + \frac{\sigma^i_t}{X^i_t} dB^i_t$$

$$= \left( \hat{\mu}^i_t - \frac{1}{2} (\hat{\sigma}^i_t)^2 \right) dt + \hat{\sigma}^i_t dB^i_t$$

where $\hat{\mu}^i_t := \frac{\mu^i_t}{X^i_t}$ and $\hat{\sigma}^i_t := \frac{\sigma^i_t}{X^i_t}$ are in $A^2_{\text{loc},d}$. Then

$$X^i_t = e^{Y^i_t} = x^i_0 \exp \left\{ \int_0^t \left( \hat{\mu}^i_s - \frac{1}{2} (\hat{\sigma}^i_s)^2 \right) ds + \int_0^t \hat{\sigma}^i_s dB^i_s \right\} \quad (4.4)$$

which solves (exercise)

$$dX^i_t = X^i_t \hat{\mu}^i_t dt + X^i_t \hat{\sigma}^i_t dB^i_t.$$
Thus, most often one sees the following market model in the literature:

Money market account: \(dX^0_t = r_t X^0_t dt; \quad X^0_0 = 1\)

Risky assets: \(dX^i_t = X^i_t \mu^i_t dt + X^i_t \sigma^i_t dB_t; \quad X^i_0 = x_i \in \mathbb{R}_+\)

where \(\mu\) and \(\sigma\) are progressive processes with \(\sqrt{\|\mu\|, \|\sigma\|} \in A^2_{loc}(T)\) and the \(X^i, i = 0, \ldots, n\), can be represented as in (4.4). In particular, for \(X^0\) we get (4.2). In such exponential models one usually refers to \(\mu^i\) respectively \(\sigma^i\) (and not to \(X^i \mu^i\) respectively \(X^i \sigma^i\)) as drift respectively volatility processes.

**Example 4.1.2.** One of the first models for financial markets in continuous time is the Black-Scholes model, due to Fischer Black, Myron Scholes and Robert Merton (1973) and Samuelson (1995). They assumed \(n = d = 1\) and modelled the risky asset price \(X^1_t = (X^1_t)_{t \in [0,T]}\) as a geometric Brownian motion

\[
\begin{cases}
    dX^1_t = X^1_t (\mu dt + \sigma dB_t), & t \in [0,T], \\
    X^1_0 = x_0, & x_0 \in \mathbb{R}_+,
\end{cases}
\]  

(4.5)

with constants \(\mu \in \mathbb{R}\) and \(\sigma \in \mathbb{R}_+\). By Section 3.4 of Chapter 3, we have that the solution of (4.5) is given by

\[
X^1_t = x_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma B_t}, \quad t \in [0,T].
\]

In this model the interest rate is constant, i.e. \(X^0_t = e^{rt}, \quad t \in [0,T]\).

An agent in the market invests in risky assets \(X^i, i \in \{1, \ldots, n\}\), and puts money in or borrows money from the bank. We represent his investment strategy as follows.

**Definition 4.1.3.** A trading strategy or portfolio is an \((n + 1)\)-dimensional progressive process \(\theta = (\theta_t)_{t \in [0,T]}\), with \(\theta_t = (\theta^0_t, \theta^1_t, \ldots, \theta^n_t)\), where \(\theta^i_t\) represents the number of assets \(i\) hold at time \(t\) and \(\tilde{\theta} = (\theta^1_t, \ldots, \theta^n_t)_{t \in [0,T]}\) denotes the investment in the risky assets without the bank account. The value at time \(t\) of a portfolio \(\theta\) is then defined by the value process

\[
V^\theta_t = V_t := \theta_t \cdot X_t = \sum_{i=0}^n \theta^i_t X^i_t.
\]

If \(\theta^i_t \geq 0\) we say that the agent is long asset \(i\) or that he has a long position in asset \(i\), which means he has bought asset \(i\). Conversely, if \(\theta^i_t < 0\), we say that the agent is short asset \(i\) or that he has a short position in asset \(i\), which means he has sold asset \(i\). Note that we are not making any assumptions on the sign of \(\theta^i_t\), i.e. in our model short sales are allowed.

**Definition 4.1.4.** Consider a portfolio \(\theta = (\theta_t)_{t \in [0,T]}\), with \(\theta_t = (\theta^0_t, \theta^1_t, \ldots, \theta^n_t)\), such that
We say that $\theta$ is self-financing, if the value process $V^\theta = V = (V_t)_{t \in [0,T]}$ satisfies

$$V_t = V_0 + \int_0^t \theta_s \cdot dX_s, \quad t \in [0,T],$$

or in short notation

$$dV_t = \theta_t \cdot dX_t.$$  

Note that (4.7) and (4.8) can be written in terms of the Brownian motion $B$ as follows:

$$V_t = V_0 + \sum_{i=0}^n \int_0^t \theta^i_s \cdot dX^i_s$$

$$= V_0 + \int_0^t \theta^0_s X^0_s \, ds + \sum_{i=1}^n \int_0^t \mu^i_s \theta^i_s \, ds + \sum_{k=1}^d \int_0^t \left( \sum_{i=1}^n \theta^i_s \sigma^{i,k}_s \right) \, dB^k_s$$  

or

$$dV_t = \theta_t \cdot dX_t = \left( \theta^0_t X^0_t + \sum_{i=1}^n \mu^i_t \theta^i_t \right) \, dt + \sum_{k=1}^d \left( \sum_{i=1}^n \theta^i_t \sigma^{i,k}_t \right) \, dB^k_t.$$  

Hence condition (4.6) is necessary to obtain that the stochastic integrals in (4.9) and (4.10) are well-defined.

**Example 4.1.5.** In the case of the Black-Scholes market, $\theta = (\theta_t)_{t \in [0,T]}$, with $\theta_t = (\theta^0_t, \theta^1_t)$, satisfies condition (4.6) if and only if

$$\int_0^T |\theta^0_s e^{\mu_s r} + \mu \theta^1_s X^1_s| \, ds + \int_0^T \sigma^2(\theta^1_s)^2 (X^1_s)^2 \, ds < +\infty \quad \text{a.s.}$$

**Remark 4.1.6.** The notion of self-financing strategies is motivated by the analog definition in discrete-time finance. Consider the value process

$$V_t = \theta_t \cdot X_t$$  

associated to a strategy $\theta$, not necessarily self-financing. Using the integration by parts formula, (4.11) would lead to

$$dV_t = \theta_t \cdot dX_t + X_t \cdot d\theta_t + d\langle X, \theta \rangle_t,$$

if $\theta$ was also an Itô process.

Actually, the definition of self-financing strategies originates from the corresponding discrete-time model. If we consider a setting, where the investments are done at discrete times $t = t_k$, then the portfolio value variation $\Delta V_{t_k} = V_{t_k} - V_{t_{k-1}}$ is given by
\[ \Delta V_{t_k} = \theta_{t_k} \cdot \Delta X_{t_k}, \quad (4.12) \]

\[ \Delta X_{t_k} = X_{t_k} - X_{t_{k-1}}, \quad \text{if the strategy is self-financing.} \]

If we consider our continuous-time model as a limit of the discrete-time case when \( \Delta t_k = t_k - t_{k-1} \to 0 \), then by Corollary 2.6.5, if \( \theta \) is continuous, it follows that (4.10) is the limit of (4.12). This is the reason why also for general strategies \( \theta \), not necessarily continuous, in view of (4.12) we say \( \theta \) is self-financing if (4.10) holds. The financial interpretation of a self-financing strategy is that all changes in the portfolio value are due to gains or losses of the asset values, as opposed to injections or withdrawals of additional cash.

Often it is useful (for example for the purpose of comparing prices at different points in time) to express prices in other units than currencies.

**Definition 4.1.7.** A numéraire is a price process \( X = (X_t)_{t \in \mathbb{R}^+} \) a.s. strictly positive for all \( t \in [0,T] \). Using \( X \) as numéraire means to use \( X \) as unity of price.

In the following, we will use \( X^0 \) as numéraire which results in the discounted market, where all asset prices are discounted by \( X^0 \), i.e.

\[ \bar{X}^i_t := \frac{X^i_t}{X^0_t}, \quad t \in [0,T]. \]

\( \bar{X}^i_t \) can be regarded as the amount of money we need to deposit today (at \( t = 0 \)) in the bank, in order to have \( X^i_t \) at time \( t \).

We now compute the dynamics of the discounted asset prices. Note first that

\[ d \left( \frac{1}{X^0_t} \right) = -X^0_t dt. \]

Then we obtain

\[ d \bar{X}^i_t = \frac{1}{X^0_t} dX^i_t + X^i_t d \left( \frac{1}{X^0_t} \right) \]

\[ = \frac{1}{X^0_t} dX^i_t + X^i_t \frac{1}{X^0_t} dX^0_t + d \langle X^i, \frac{1}{X^0} \rangle_t \]

\[ = \frac{1}{X^0_t} (dX^i_t - r_i X^i_t dt) \]

\[ = \frac{1}{X^0_t} (\mu^i_t dt + \sigma^i_t dB_t) - r_i \frac{X^i_t}{X^0_t} dt \]

\[ = \frac{1}{X^0_t} (\mu^i_t - r_i X^i_t) dt + \frac{\sigma^i_t}{X^0_t} dB_t. \quad (4.13) \]

**Example 4.1.8.** In the exponential model of Remark 4.1.1 we obtain

\[ d \bar{X}^i_t = \frac{1}{X^0_t} dX^i_t + X^i_t d \left( \frac{1}{X^0_t} \right) \]

\[ = \frac{X^i_t}{X^0_t} (\mu^i_t dt + \sigma^i_t dB_t - r_i dt) \]

\[ = \frac{X^i_t}{X^0_t} (\mu^i_t dt + \sigma^i_t dB_t - r_i dt). \tag{4.14} \]
= \bar{X}_t^i \left( (\mu^i_t - r_t) dt + \sigma^i_t dB_t \right),

i.e.

\bar{X}_t^i = x_i \exp \left\{ \int_0^t \left( \mu^i_s - r_s - \frac{(\sigma^i_s)^2}{2} \right) ds + \int_0^t \sigma^i_s dB_s \right\}, \quad t \in [0, T].

Note that discounting changes only the drift in the dynamics of \( X^i \), the volatility remains unchanged!

In particular, for the Black-Scholes model we obtain

\[ dX^1_t = X^1_t (\mu - r) dt + \sigma^1_t dB_t, \]

i.e.

\[ X^1_t = xe^{(\mu - r - \frac{\sigma^2}{2})t + \sigma B_t}, \quad t \in [0, T]. \]

Given a self-financing portfolio \( \theta = (\theta_t)_{t \in [0, T]} \), where \( \theta_t = (\theta^0_t, \ldots, \theta^n_t) \), the associated discounted portfolio value is then given by

\[ V^\theta_t = V_t := \frac{V_t}{X^\theta_t} = \theta^0_t + \sum_{i=1}^n \theta^i_t \frac{X^i_t}{X^\theta_t} = \theta_t \cdot \bar{X}_t, \quad t \in [0, T], \]

where \( \bar{X}_t = \left( 1, \frac{X^1_t}{X^\theta_t}, \ldots, \frac{X^n_t}{X^\theta_t} \right) \).

**Proposition 4.1.9.** A portfolio \( \theta = (\theta_t)_{t \in [0, T]} \) is self-financing in the original market \( X \) if and only if it is self-financing in the discounted market \( \bar{X} \).

**Proof.** Let first \( \theta \) be self-financing with respect to the price process \( X \) (original market). By the integration by parts formula we obtain for the discounted portfolio

\[
d\bar{V}_t = d \left( \frac{V_t}{X^\theta_t} \right) = \frac{1}{X^\theta_t} dV_t + V_t d \left( \frac{1}{X^\theta_t} \right)
\]

\[
= \frac{\theta_t}{X^\theta_t} \cdot dX_t - \frac{V_t}{X^\theta_t} r_t dt
\]

\[
= \frac{1}{X^\theta_t} (\theta_t \cdot dX_t - \theta_t \cdot X_t r_t dt)
\]

\[
= \theta_t \cdot \frac{1}{X^\theta_t} (dX_t - r_t X_t dt)
\]

\[
= \theta_t \cdot d\bar{X}_t, \quad \text{see (4.13)}
\]

i.e.
4.2 Arbitrage and Equivalent Local Martingale Measures

\[ V_t = V_0 + \int_0^t \theta_s \cdot d\overline{X}_s = V_0 + \sum_{i=1}^n \int_0^t \theta^i_s d\overline{X}^i_s. \]  

(4.16)

Thus \( \theta \) yields also well defined integrals in the discounted market and from (4.15) and (4.16) we obtain that \( \theta_t \) is also self-financing in the discounted market. If \( \theta \) is self-financing with respect to the price process \( X \) (discounted market) we apply the analogue argument to \( V_t = X^0_t V_t \).

Note that in (4.16) the dynamics \( d\overline{X}^0_t \) of \( \overline{X}^0 \) does not contribute to the dynamics \( dV_t \) of \( V \) (since \( \overline{X}^0 \equiv 1 \) is constant). From this fact we obtain the following useful lemma:

**Lemma 4.1.10.** A self-financing strategy \( \theta \) is completely determined by the initial value \( V^\theta_0 \) of the portfolio and the investments \( \hat{\theta}_t = (\theta^1_t, \ldots, \theta^n_t)_{t \in [0,T]} \) in the risky assets. Conversely, given \( V_0, \theta \) there exists a unique \( \theta^0 \) such that the strategy \( \theta = (\theta^0_t, \ldots, \theta^n_t)_{t \in [0,T]} \) is self-financing:

\[ \theta^0_t := V_t - \sum_{i=1}^n \theta^i_t \overline{X}^i_t = V_0 + \sum_{i=1}^n \left( \int_0^t \theta^i_s d\overline{X}^i_s - \theta^i_t \overline{X}^i_t \right), \quad t \in [0,T]. \]

4.2 Arbitrage and Equivalent Local Martingale Measures

The proof of the following result can be found in Dellacherie and Meyer [5].

**Proposition 4.2.1.** Let \( B = (B_t)_{t \in [0,T]} \) be a \( d \)-dimensional Brownian motion on \( (\Omega, \mathcal{F}, (\mathcal{F}^B_t)_{t \in [0,T]}, \mathbb{P}) \) and assume \( F \) is an \( \mathcal{F}^B_T \)-measurable random variable. Then there exists \( \Phi \in A^2_{loc,d}(T) \), such that

\[ F = \int_0^T \Phi_s \cdot dB_s. \]

This implies that for any constant \( z \in \mathbb{R} \), there exists \( \Psi \in A^2_{loc,d}(T) \) such that

\[ F = z + \int_0^T \Psi_s \cdot dB_s. \]

This proposition implies that if we have a simple financial market model with \( n = d, \overline{X}^0_t = 1, X^i_t = B^i_t, i = 1, \ldots, d \), then this means that we can generate any \( \mathcal{F}_T \)-measurable final value \( V_T = F \) with any initial fortune \( V_0 = z \), as long as we can freely choose \( \theta = \Psi \) in \( A^2_{loc,d}(T) \). This, of course, is economically not meaningful. Hence we need to impose some additional restrictions on the choice of self-financing strategies. One of the most common is introduced in the next definition.
Definition 4.2.2. A self-financing portfolio $\theta$ is called admissible or tame, if the corresponding value process $V = V^\theta$ is almost surely lower bounded in $(t, \omega)$, i.e. if there exists a constant $K > 0$, such that

$$V^\theta(t, \omega) \geq -K, \quad \text{for a.e. } (t, \omega) \in [0, T] \times \Omega.$$  \hfill (4.17)

Condition (4.17) may be interpreted as the limit of debt that a potential creditor would tolerate.

Lemma 4.2.3. A portfolio $\theta$ is admissible in the original market given by $X$ if and only if it is admissible in the discounted market $\overline{X}$.

Proof. By Proposition 4.1.9, it remains to show that $V^\theta$ is lower bounded a.e. if and only if $\overline{V}^\theta$ is lower bounded a.e.. But this is clear since the risk free rate $r_t$ is assumed to be bounded. \hfill $\square$

We are now ready to introduce the definition of arbitrage.

Definition 4.2.4. An admissible portfolio $\theta = (\theta_t)_{t\in[0,T]}$ is called an arbitrage (strategy) or arbitrage possibility in the market $(X_t)_{t\in[0,T]}$, if the corresponding value process $V^\theta = V = (V_t)_{t\in[0,T]}$ satisfies

1) $V_0 = 0$,
2) $V_T \geq 0$ a.s. and $\mathbb{P}(V_T > 0) > 0$.

We say that the market is arbitrage-free, if no such arbitrage strategy exists.

In other words, a portfolio $\theta$ is an arbitrage strategy if it produces a profit without any risk of losing money.

Remark 4.2.5. Note that by positivity of the discount factor $X^0$ and Lemma 4.2.3, a portfolio $\theta$ is an arbitrage in the market $X$ if and only if it is an arbitrage in the discounted market $\overline{X}$.

Definition 4.2.6 (Equivalent (local) martingale measure). In a financial market model $X = (X_t)_{t\in[0,T]}$, where $X_t = (X^0_t, X^1_t, \ldots, X^n_t)$, with $n$ risky assets and a bank account defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^B)_{t\in[0,T]}, \mathbb{P})$, an equivalent (local) martingale measure $Q \sim \mathbb{P}$ is a probability measure that is equivalent to $\mathbb{P}$, such that all discounted asset prices $\overline{X}^i$, $i = 1, \ldots, n$ are (local) martingales under $Q$.

As in the discrete-time case, to check if a market model is arbitrage-free can be done by investigating the existence of an equivalent martingale measure.

Proposition 4.2.7. On $(\Omega, \mathcal{F}_T, (\mathcal{F}_t^B)_{t\in[0,T]}, \mathbb{P})$ any Radon-Nykodym density $Z := dQ/d\mathbb{P}$ of an equivalent probability measure $Q \sim \mathbb{P}$ is of the form

$$Z = \mathcal{E} \left( \int H \cdot dB \right)_T := \exp \left\{ \int_0^T H_s \cdot dB_s - \frac{1}{2} \int_0^T \|H_s\|^2 ds \right\}$$
for $H = (H^1, \ldots, H^d) \in A^2_{loc,d}(T)$.

In particular, if $Q$ is an equivalent local martingale measure then $H$ is such that

$$\mu_t - r_t X_t + \sigma_t H_t = 0 \quad \text{a.e. in } (t, \omega),$$

and the (local) martingale dynamics of $X^i$ is given by

$$dX^i_t = \sigma^i_t X^0_t dB^Q_t,$$

where $B^Q_t := B_t - \int_0^t H_s ds$ is a $Q$-Brownian motion.

**Proof.** Consider the martingale $(Z_t)_{t \in [0,T]}$, where $Z_t = E[Z|\mathcal{F}_t]$. Then by Theorem 2.9.2 (martingale representation theorem) there exists $\tilde{H} \in A^2_{loc,d}(T)$, such that

$$Z_t = 1 + \int_0^t \tilde{H}_s \cdot dB_s, \quad t \in [0,T].$$

Since $Z_t > 0$ a.s. for all $t \in [0,T]$, $X_t = \log Z_t$ is well-defined, and we obtain as in Remark 4.1.1

$$Z_t = E[Z] E\left(\int H_s \cdot dB_s\right), \quad t \in [0,T],$$

where $H_t := \frac{\tilde{H}_t}{Z_t}$. Now let $\frac{dQ}{dP} = E\left(\int H_s \cdot dB_s\right)_T$ be the density of an equivalent local martingale measure. By Girsanov’s theorem, $B^Q_t = B_t - \int_0^t H_s ds$ is a $Q$-Brownian motion. Then by using (4.14), the dynamics of the discounted asset price $\bar{X}$ under $Q$ are given by

$$d\bar{X}^i_t = (\mu^i_t - r_t X^i_t) \frac{1}{X^0_t} dt + \frac{\sigma^i_t}{X^0_t} dB^Q_t$$

$$= (\mu^i_t - r_t X^i_t + \sigma^i_t \cdot H_t) \frac{1}{X^0_t} dt + \frac{\sigma^i_t}{X^0_t} dB^Q_t.$$

Since $\bar{X}$ is a $Q$-local martingale, $H$ must be defined by

$$\mu^i_t - r_t X^i_t + \sigma^i_t \cdot H_t = 0 \quad \text{a.e. in } (t, \omega),$$

i.e.

$$d\bar{X}^i_t = \frac{\sigma^i_t}{X^0_t} \cdot dB^Q_t. \quad (4.18)$$

**Remark 4.2.8.** Proposition 4.2.7 can also be formulated for an infinite horizon $[0, \infty)$, assuming the integrability conditions hold for all $t \geq 0$. 

\[ \Box \]
Lemma 4.2.9. Suppose that $X$ is an $n$-dimensional $\mathbb{Q}$-local martingale and that $\theta$ is an admissible strategy. Then the discounted portfolio process $\overline{V} = \frac{V}{X_0}$ is also a $\mathbb{Q}$-local martingale.

Proof. By (4.15) and (4.18) we have that
$$d\overline{V}_t = \theta_t \cdot \frac{\sigma_t}{X_0^t} \cdot dB^Q_t = \sum_{k=1}^d \left( \sum_{i=1}^n \frac{\theta_i^k}{X_i^t} \sigma_{i,k}^t \right) dB_t^Q,$$
hence the discounted portfolio value is a $\mathbb{Q}$-local martingale. $\square$

Proposition 4.2.10. Suppose that there exists an equivalent (local) martingale measure for the discounted asset price $X = (X_t)_{t \in [0,T]}$. Then the market $X = (X_t)_{t \in [0,T]}$ is arbitrage-free, i.e. there exist no arbitrage possibilities.

Proof. Suppose $\theta$ is an arbitrage strategy in the market $X = (X_t)_{t \in [0,T]}$. Then the corresponding discounted value process $\overline{V} = \overline{V}^\theta$ is a lower bounded $\mathbb{Q}$-local martingale. Hence it is a supermartingale under $\mathbb{Q}$ (Lemma 1.10.2), and therefore
$$E[\overline{V}_T] \leq \overline{V}_0 = 0. \quad (4.19)$$
Since $\mathbb{Q} \sim \mathbb{P}$ and $V_T \geq 0$ $\mathbb{P}$-a.s. with $\mathbb{P}(V_T > 0) > 0$, then $\overline{V}_T \geq 0$ $\mathbb{Q}$-a.s. and $\mathbb{Q}(\overline{V}_T > 0) > 0$. Hence
$$E[\overline{V}_T] > 0,$$
but this contradicts (4.19). Therefore such a strategy $\theta$ cannot exist in this market $X$. $\square$

Remark 4.2.11. Note that contrary to discrete time models the opposite direction is NOT valid in continuous time models, i.e.

existence of equivalent local martingale measures $\Downarrow$ \# \hspace{1cm} (4.20)
the market is arbitrage free

In order to establish equivalence in (4.20) in our model (or more general, locally bounded semimartingale models) one needs to introduce a stronger concept of "no arbitrage", the so called "no free lunch with vanishing risk" (NFLVR). For general locally unbounded semimartingale models one even needs to further generalize the concept of "local martingale measures" to the concept of "sigma martingale measures".

Example 4.2.12 (The Black-Scholes model is arbitrage-free). Consider once again the Black-Scholes market, where $X_0^t = e^{rt}$ and $X_1^t = x_0 e^{\left( \mu - \frac{\sigma^2}{2} \right) t + \sigma B_t}$, $t \in [0,T]$. From Example 4.1.8 we already know that for all $t \in [0,T]$,
4.2 Arbitrage and Equivalent Local Martingale Measures

\[ \bar{X}_t^1 = \frac{X_t^1}{X_0} = x_0 e^{(\mu - r - \frac{\sigma^2}{2}) t + \sigma B_t}, \]
i.e.
\[ d\bar{X}_t^1 = \bar{X}_t^1 \left[ (\mu - r) dt + \sigma dB_t \right]. \]

Consider the strictly positive random variable \( Z = \exp \left( \lambda B_T - \frac{\lambda^2}{2} T \right), \lambda \in \mathbb{R}. \) Then by the Novikov condition (2.16) or by Proposition 1.5.2 iii) we have that \( E[Z] = 1, \) i.e. an equivalent probability measure \( Q \) is well-defined by the density \( \frac{dQ}{dP} = Z. \) By Girsanov’s theorem (see Theorem 2.8.5), we obtain that \( \tilde{B} = (\tilde{B}_t)_{t \in [0,T]} \) given by \( \tilde{B}_t = B_t - \lambda t \)

is a \( Q \)-Brownian motion. The dynamics of \( \bar{X}_t^1 \) under \( Q \) are then given by

\[ d\bar{X}_t^1 = \bar{X}_t^1 \left[ (\mu - r + \sigma \lambda) dt + \sigma d\tilde{B}_t \right]. \]

We see now, that if we impose that

\[ \mu - r + \sigma \lambda = 0, \]
i.e. we choose \( \lambda = -\frac{\mu - r}{\sigma}, \) then \( Q \) is an equivalent martingale measure in this market. Then by Proposition 4.2.10 the market is arbitrage-free.

Note that the dynamics of \( X^1 \) under \( Q \) are given by

\[ dX^1_t = rX^1_t dt + \sigma X^1_t d\tilde{B}_t, \quad t \in [0, T]. \]

More generally, we have in our market model:

**Theorem 4.2.13.** Let \( X = (X^0_t, X^1_t, \ldots, X^n_t)_{t \in [0,T]} \) be a financial market with \( n \) risky assets and a bank account as defined in (4.1) and (4.3).

i) Suppose that there exists a progressive \( \mathbb{R}^d \)-valued process \( \lambda = (\lambda_t)_{t \in [0,T]}, \) such that

\[ \mu_t - r_t X_t = \sigma_t \lambda_t, \]

for almost all \((t, \omega), \) \( \int_0^T \|\lambda_s\|^2 ds < +\infty, \)

and

\[ E \left[ \exp \left( \frac{1}{2} \int_0^T \|\lambda_s\|^2 ds \right) \right] < +\infty. \]

Then the market \((X_t)_{t \in [0,T]} \) is arbitrage-free.
ii) Conversely, if the market is arbitrage-free, then there exists a progressive $\mathbb{R}^d$-valued process $\lambda = (\lambda_t)_{t \in [0, T]}$, such that

$$\mu_t - r_t X_t = \sigma_t \lambda_t$$

(4.21)

for almost all $(t, \omega) \in [0, T] \times \Omega$. (However, $\mathbb{E} \left( \int -\lambda_s \cdot dB_s \right)_T$ is not necessarily the density of an equivalent measure $Q \sim P$, which then would be an equivalent local martingale measure.)

Proof. i) Consider the measure $Q$ defined by its density

$$\frac{dQ}{dP} = \mathbb{E} \left( \int -\lambda_s \cdot dB_s \right)_T.$$

By the Novikov condition (Proposition 2.7.2) we obtain that $Q \sim P$, and as in the proof of Proposition 4.2.7 the dynamics of $X$ under $Q$ are then given by

$$dX_t = \frac{1}{X_0} (\mu_t - r_t X_t - \sigma_t \lambda_t) dt + \sigma_t \frac{1}{X_0} dB_t,$$

or for all $i = 1, \ldots, n$,

$$dX^i_t = \frac{1}{X_0} \left( \mu^i_t - \sum_{k=1}^d \sigma^i_k \lambda^k_t - r_t X^i_t \right) dt + \sigma^i_t \frac{1}{X_0} dB^i_t,$$

where for $t \in [0, T]$ we have used the notation

$$\sigma_t \lambda_t := \begin{pmatrix}
\sigma^{11}_t & \ldots & \sigma^{1d}_t \\
\vdots & \ddots & \vdots \\
\sigma^{n1}_t & \ldots & \sigma^{nd}_t \\
\lambda^1_t & \vdots & \lambda^d_t
\end{pmatrix},$$

and $B_t = B_0 + \int_0^t \lambda_s ds$, $t \in [0, T]$ is a $Q$-Brownian motion by Girsanov’s theorem 2.8.5. Since we have assumed that $\mu_t - r_t X_t - \sigma_t \lambda_t = 0$ for almost all $(t, \omega)$, we have that $Q$ is a (local) martingale measure for $X$ and the market is then arbitrage-free by Proposition 4.2.10.

ii) Suppose that the market has no arbitrage. By Remark 4.2.5 we can without loss of generality assume $r_t \equiv 0$, i.e. $X = \overline{X}$. For $t \in [0, T], \omega \in \Omega$, let

$$F_t := \{ \omega : (4.21) \text{ has no solution} \} = \{ \omega : \mu(t, \omega) \notin \text{Span } \sigma(t, \omega) \},$$

where $\text{Span } \sigma(t, \omega)$ denotes the linear span generated by the columns of $\sigma(t, \omega)$. Since $\text{Span } \sigma(t, \omega) = \text{ker}(\sigma^{tr}(t, \omega))^{\perp}$, there exists $v(t, \omega) \in \ker(\sigma(t, \omega)^{tr})$, $t \in [0, T]$, i.e.

$$\sigma(t, \omega)^{tr} v(t, \omega) = 0,$$
such that \( v(t, \omega) \), \( \mu(t, \omega) \) are not orthogonal, i.e.

\[
v(t, \omega) \cdot \mu(t, \omega) \neq 0.
\]

We then define a strategy \( \theta = (\theta^0, \theta^1, \ldots, \theta^n) \), where for all \( t \in [0, T] \),

\[
\theta^i(t, \omega) = \begin{cases} 
\text{sign} \left( v(t, \omega) \cdot \mu(t, \omega) \right) v^i(t, \omega), & \omega \in F_t, \\
0, & \text{otherwise}
\end{cases}
\]

for \( i = 1, \ldots, n \), and \( \hat{\theta} = (\theta^1, \ldots, \theta^n) \) is completed with \( \theta^0 \) such that the resulting strategy is self-financing with \( V^0 = 0 \) (Lemma 4.1.10). Then we have

\[
V_T = V_T = 0 + \int_0^T \theta_s \cdot dX_s
\]

\[
= \int_0^T 1_{F_t}(\omega) \cdot \text{sign} \left( v(s, \omega) \cdot \mu(s, \omega) \right) \left( v(s, \omega) \cdot \mu(s, \omega) \right) ds
\]

\[
+ \int_0^T \sum_{k=1}^d \left( \sum_{i=1}^n \theta^i(s, \omega) \sigma^{i,k}(s, \omega) \right) dB^k_s
\]

\[
= \int_0^T 1_{F_t} |v(s, \omega) \cdot \mu(s, \omega)| ds
\]

\[
+ \int_0^T 1_{F_t} \text{sign} \left( v(s, \omega) \cdot \mu(s, \omega) \right) \sigma^{sT}(s, \omega) v(s, \omega) dB_s
\]

\[
= \int_0^T 1_{F_t} |v(s, \omega) \cdot \mu(s, \omega)| ds \geq 0.
\]

Since the market is arbitrage-free, we must have

\[
\mathbb{P}(F_t) = 0
\]

for almost all \( t \in [0, T] \), i.e. there exists a solution \( \lambda = (\lambda(t, \omega))_{t \in [0, T]} \), such that \( \mu(t, \omega) - r(t, \omega)X(t, \omega) - \sigma(t, \omega)\lambda(t, \omega) = 0 \) for a.a. \((t, \omega) \in [0, T] \times \Omega\). To prove that \( \lambda \) is progressively measurable is more delicate and we refer to Lemma 4.4 and Corollary 4.5 of Karatzas and Shreve [10] for a proof. \( \square \)

**Remark 4.2.14.** Suppose that the dynamics of \( X = (X_t)_{t \in [0, T]} \) are given by the exponential model from Remark 4.1.1:

\[
\begin{align*}
\begin{cases}
\text{d}X_i^t = X_i^t(\mu_i^t dt + \sigma_i^t \cdot dB_t), & t \in [0, T], \\
X_i^0 = x_i^0, & x_i^0 \in \mathbb{R}_+
\end{cases}
\end{align*}
\]

for \( i = 1, \ldots, n \), with \( X^0 \) as in (4.1). Here \( \mu \in \mathcal{A}^2_{loc,n}(T) \) and \( \sigma \) is an \( \mathbb{R}^{n \times d} \)-valued progressive process such that \( \|\sigma\| \in \mathcal{A}^2_{loc}(T) \). Then the progressive process \( \lambda \) in (4.21) solves
\[ \sigma_t \lambda_t = \mu_t - r_t \mathbb{1}, \]

where \( \mathbb{1} = (1, \ldots, 1) \in \mathbb{R}^n \). If the left-inverse \( \Lambda_t \in \mathbb{R}^{d \times n} \) of \( \sigma_t \) exists for a.a. \((t, \omega) \in [0, T] \times \Omega \) (see Theorem 4.4.7 of Section 4.3) then,

\[ \lambda_t = \Lambda_t (\mu_t - r_t \mathbb{1}). \]

We call the process \( \lambda \) the market price of risk.

### 4.3 Risk-neutral Pricing of Contingent Claims

During the whole of this section let \( X = (X^0_t, X^1_t, \ldots, X^n_t)_{t \in [0, T]} \) be a financial market with \( n \) risky assets and a bank account as defined in (4.1) and (4.3). Throughout the rest of this chapter we will always assume that the set

\[ M(\mathbb{P}) := \{ Q \sim \mathbb{P} \mid Q \text{ is an equivalent martingale measure} \} \]

is non-empty, i.e. the market is arbitrage-free.

In addition to investing in the \( n+1 \) primitive or primary assets in the market we now assume that there is the possibility of buying or selling additional financial contracts whose value depends on the value of an underlying (asset), i.e. on the value of the primary assets. These financial products are called derivatives, or more generally contingent claims. Here we focus on European (in comparison to American) type contingent claims, i.e. financial contracts that stipulate a cash flow \( H(\omega) \) (payoff) from the seller to the buyer at a fixed time \( T > 0 \), called expiration date or maturity. Some examples of derivatives contracts are:

i) The seller of a forward (contract) agrees to sell another agent the financial asset \( X^i \) at a fixed delivery date at a fixed delivery price \( K \). Since forward contracts can be used to lock in a future price in advance, this kind of contract is used e.g. by companies that need a certain amount of a commodity at a future date, or in the future need to pay a certain amount of money in a foreign currency. Furthermore, forward contracts can also be used for speculation, since both parties enter into the contract at an initial price of zero as we will show later on. The payoff is given by \( H = (X^i - K) \).

ii) Futures (contracts) are very similar to forward contracts in the sense that they are agreements between two parties to buy or sell a certain asset at a prespecified time \( T \) in the future. In comparison to forwards, futures contracts are typically traded on exchanges and have standardized features. Besides that, futures incorporate regular payment streams (“daily settlement”), whereas there is only one exchange of payments at time \( T \) in the case of a forward contract. These features ensure that futures markets are in general more liquid than forward markets. We will come back to this in Section 6.3.
iii) A European call (option) on asset $X^i$, with maturity $T$ and strike (price) $K > 0$ gives the buyer or holder the right but not the obligation to buy the $i$-th asset at time $T$ at a prespecified price $K$. The payoff at time $T$ is given by $H = (X^i_T - K)^+$. 

iv) A European put (option) on asset $X^i$, with maturity $T$ and strike (price) $K > 0$ gives the buyer or holder the right but not the obligation to sell the $i$-th asset at time $T$ at a prespecified price $K$. The payoff at time $T$ is given by $H = (K - X^i_T)^+$. 

**Definition 4.3.1.** We represent a European contingent claim with maturity $T$ by an $\mathcal{F}^B_T$-measurable random variable $H \geq 0$. $H$ is also referred to as payoff of the claim. For notational convenience, in the sequel we will sometimes say contingent claim or claim, meaning European contingent claim. 

**Proposition 4.3.2.** Let $H$ be a European contingent claim such that there exists a $Q \in \mathcal{M}(\mathbb{P})$ with $H \in L^1(\mathcal{F}^B_T, Q)$. Then the risk-neutral pricing formula (with respect to $Q$)

$$\pi_t(\omega) := X^0_t E^Q \left[ \frac{H}{X^0_T} \bigg| \mathcal{F}^B_t \right], \quad t \in [0, T],$$ 

gives an arbitrage-free price (process) for the claim $H$ in the sense that $\pi_T = H$ a.s. and the enlarged market 

$$(X^0, X^1, \ldots, X^n, \pi)$$

is arbitrage-free. 

**Proof.** First note that since $H \in L^1(\mathcal{F}^B_T, Q)$ and the risk free rate $r$ is assumed to be bounded $\frac{H}{X^0_T} \in L^1(\mathcal{F}^B_T, Q)$ and $\pi$ is well defined. Further, the discounted price process

$$\pi_t X^0_t = E^Q \left[ \frac{H}{X^0_T} \bigg| \mathcal{F}^B_t \right]$$

is an $(\mathcal{F}^B_t)_{t \in [0, T]}$-martingale under $Q$. Thus, $Q$ is an equivalent local martingale measure for $(X^0, X^1, \ldots, X^n, \pi)$ ($Q$ was already an equivalent local martingale measure for $X$), and the enlarged market is arbitrage-free. 

**Remark 4.3.3.** Contrary to discrete time models, the risk-neutral price processes do not, in general, compose all arbitrage-free price processes in the sense of Proposition 4.3.2. One could also have price processes $\pi_t$ with $\pi_T = H$ a.s. such that $\frac{\pi_t}{X^0_T}$ is a local martingale but not a martingale under $Q \in \mathcal{M}(\mathbb{P})$. These price processes might not be representable as risk-neutral prices.

The set

$$\Pi_t = \left\{ E^Q \left[ \frac{X^0_t}{X^0_T} H \bigg| \mathcal{F}^B_t \right] \right\}, \quad Q \in \mathcal{M}(\mathbb{P}) \text{ such that } H \in L^1(Q)$$
of risk-neutral prices at time $t$ (which, as easily seen, is an interval for $t = 0$) can be bounded from above and below as follows. Let $L^0_+(\mathcal{F}^B_t)$ be the set of positive $\mathcal{F}^B_t$-measurable random variables. From the viewpoint of the seller of the claim $H$, at time $t$ surely any price $y_t \in L^0_+(\mathcal{F}^B_t)$ is acceptable such that there exists an admissible strategy $\theta$ on $[t, T]$ with portfolio value $V^\theta_{t, y_t} = y_t$ at time $t$ and

$$V^\theta_{T, y_T} = y_t + \int_t^T \theta_s dX_s \geq H \quad \text{a.s..}$$

We define the superreplication price of $H$ at time $t$ by

$$q_t(H) := \text{essinf} \left\{ \left\{ y_t \in L^0_+(\mathcal{F}^B_t) \mid \exists \text{ admiss. } \theta \text{ such that } V^\theta_{t, y_t} = y_t \text{ and } V^\theta_{T, y_T} \geq H \text{ a.s.} \right\} \cup \{\infty\} \right\}.$$

Thus $q_t(H)$ is the lowest price at time $t$ such that the seller can hedge against all possible future claims of the buyer.

Let $L^\infty_+(\mathcal{F}^B_t)$ be the set of positive bounded $\mathcal{F}^B_t$-measurable random variables. From the viewpoint of the buyer, at time $t$ surely any price $y_t \in L^\infty_+(\mathcal{F}^B_t)$ is acceptable such that there exists an admissible strategy $\theta$ with $V^\theta_{t, -y_t} = -y_t$ and

$$V^\theta_{T, -y_T} = -y_t + \int_t^T \theta_s dX_s + H \geq 0 \quad \text{a.s.}$$

$$\iff V^\theta_{T, -y_T} \geq -H \quad \text{a.s.,}$$

(note that if also $-\theta$ is admissible this is equivalent to $V^-_{T, -y_T} \leq H$ a.s.). We define the subreplication price of $H$ at time $t$ by

$$p_t(H) := \text{esssup} \left\{ y_t \in L^\infty_+(\mathcal{F}^B_t) \mid \exists \text{ admiss. } \theta \text{ such that } V^\theta_{t, -y_t} = y_t \text{ and } V^\theta_{T, -y_T} \geq -H \text{ a.s.} \right\}.$$

Thus $p_t(H)$ is the highest price at time $t$ such that the buyer can completely cover this price by hedging in the market and receiving $H$ at $T$. We then get

**Theorem 4.3.4.** Let $X = (X^0, X^1, \ldots, X^n)$ be a financial market with $n$ risky assets and a bank account as defined in (4.1) and (4.3). Let $H$ be a claim and $Q \in \mathcal{M}(\mathbb{P})$ with $H \in L^1(\mathcal{F}^B_T, Q)$. Then

$$0 \leq p_t(H) \leq E^Q \left[ \frac{X^0_t}{X^n_T} \left| \mathcal{F}^B_t \right| H \right] \leq q_t(H) \quad \text{a.s..}$$

**Proof.** Let $y_t \in L^\infty_+(\mathcal{F}^B_t)$ such that there exists an admissible portfolio $\theta$ with

$$V^\theta_{T, -y_T} = -y_t + \int_t^T \theta_s \cdot dX_s \geq -H \quad \text{a.s..} \quad (4.22)$$
4.3 Risk-neutral Pricing of Contingent Claims

Then

\[ V_T^{\theta,-y_t} = -\frac{y_t}{X_T^0} + \int_t^T \theta_s dX_s \geq -\frac{H}{X_T^0} \quad \text{a.s.} \]

Since \( V_T^{\theta,-y_t} \) is a lower bounded local martingale and hence a supermartingale on \([t,T]\) under \(Q\) we get

\[ -\frac{y_t}{X_t^0} \geq E^Q \left[ V_T^{\theta,-y_t} \mid \mathcal{F}_T^B \right] \geq E^Q \left[ -\frac{H}{X_T^0} \mid \mathcal{F}_T^B \right] \quad \text{a.s.,} \]

and hence

\[ p_t(H) \leq E^Q \left[ \frac{X_t^0}{X_T} H \mid \mathcal{F}_T^B \right] \quad \text{a.s.} \]

Since \( H \geq 0 \), obviously \( y_t = 0 \) and \( \theta = 0 \) fulfill (4.22) and therefore

\[ p_t(H) \geq 0 \quad \text{a.s.} \]

If there exists \( y_t \in L^0_+(\mathcal{F}_T^B) \) and an admissible \( \theta \) such that

\[ V_T^{\theta,y_t} = y_t + \int_t^T \theta_s dX_s \geq H \quad \text{a.s.,} \]

then as before we obtain

\[ y_t \geq E^Q \left[ \frac{X_t^0}{X_T} H \mid \mathcal{F}_T^B \right] \quad \text{a.s., i.e.} \]

\[ q_t(H) \geq E^Q \left[ \frac{X_t^0}{X_T} H \mid \mathcal{F}_T^B \right] \quad \text{a.s.} \]

If no such \( y_t \in L^0_+(\mathcal{F}_T^B) \) exists then \( q_t(H) = \infty > E^Q \left[ \frac{X_t^0}{X_T} H \mid \mathcal{F}_T^B \right] \). \( \Box \)

If \( p_t(H) = q_t(H) \) then \( \Pi_t \) collapses to one price, and the seller and buyer can agree on one price with which both can completely hedge against any loss risk from selling (buying) the claim \( H \).

**Definition 4.3.5.** Let \( H \) be a claim and \( Q \in \mathcal{M}(\mathbb{P}) \) with \( H \in L^1(\mathcal{F}_T^B, Q) \). If \( p_t(H) = q_t(H) \) a.s. for all \( t \in [0,T] \) we call

\[ E^Q \left[ \frac{X_t^0}{X_T} H \mid \mathcal{F}_T^B \right] \]

the (unique) risk-neutral price at time \( t \) of \( H \).

**Remark 4.3.6.** Note that this is well defined since by Theorem 4.3.4 if \( p_t(H) = q_t(H) \) a.s.

\[ E^Q \left[ \frac{X_t^0}{X_T} H \mid \mathcal{F}_T^B \right] = E^{\tilde{Q}} \left[ \frac{X_t^0}{X_T} H \mid \mathcal{F}_T^B \right] \quad \text{a.s.} \]

for \( Q, \tilde{Q} \in \mathcal{M}(\mathbb{P}) \) with \( H \in L^1(\mathcal{F}_T^B, Q) \cap L^1(\mathcal{F}_T^B, \tilde{Q}) \).

When do we have a unique risk-neutral price for a given claim \( H \)? This will be subject of the next section.
4.4 Pricing and Hedging in Complete Markets

As before, let $X = (X^0, X^1, \ldots, X^n)$ be a financial market with $n$ risky assets and a bank account as defined in (4.1) and (4.3) and assume that the set of equivalent local martingale measures $M(\mathbb{P})$ is not empty, i.e. the market is arbitrage-free.

**Definition 4.4.1.** A European contingent claim $H$ is called attainable (or replicable, or hedgable) if there exists $Q \in M(\mathbb{P})$ with $H \in L^1(F^Q_B, \mathbb{Q})$ and an admissible strategy $\theta$ such that

$$V_T^\theta = \theta_T X_T = V_0^\theta + \int_0^T \theta_t dX_t = H \quad \text{a.s.}$$

or equivalently

$$V_T^\theta = \theta_T X_T = V_0^\theta + \sum_{i=1}^n \int_0^T \theta^i_t dX^i_t = \frac{H}{X_T^Q} \quad \text{a.s.,}$$

and

$$V_t^\theta = V_0^\theta + \int_0^t \theta_s dX_s$$

is a $Q$-martingale (and not only a $Q$-local martingale). Such a strategy $\theta$ is called replicating strategy or hedge (with respect to $Q$) for $H$.

**Remark 4.4.2.** Note that a claim $H$ is attainable in the original market $X$ if and only if $H X^0_T$ is attainable in the discounted market $X$.

**Lemma 4.4.3.** Let $H$ be an attainable claim. Then the hedge portfolio value is unique, i.e. for $\theta$ respectively $\tilde{\theta}$ replicating strategy with respect to $Q \in M(\mathbb{P})$ respectively $\tilde{Q} \in M(\mathbb{P})$ of $H$

$$V_t^\theta = V_t^{\tilde{\theta}} \quad \text{a.s., for all } t \in [0, T].$$

Further the corresponding discounted risk-neutral price processes are given by the discounted hedge portfolio:

$$E^Q \left[ \frac{H}{X_T^Q} \mid \mathcal{F}^B_t \right] = E^{\tilde{Q}} \left[ \frac{H}{X_T^Q} \mid \mathcal{F}^B_t \right] = V_t^\theta \quad \text{a.s.}$$

In particular, for a bounded, attainable claim $H$ there is a unique (discounted) risk-neutral price process:

$$E^Q \left[ \frac{H}{X_T^Q} \mid \mathcal{F}^B_t \right] = V_t^\theta \quad \text{a.s. for all } Q \in M(\mathbb{P}).$$
Proof. We show the first result for \( t = 0 \); the extension to \( t \in [0, T] \) follows the same idea. Assume without loss of generality \( X^0 \equiv 1 \) (Remark 4.4.2) and \( V^\theta_0 > V^\tilde{\theta}_0 \). Then
\[
Y_t := V^\tilde{\theta}_t - V^\theta_t
\]
is a supermartingale under \( Q \) (why?) with \( Y_0 < 0 \) and \( Y_T = 0 \) a.s. Then
\[
0 > Y_0 \geq E^Q[Y_T] = 0
\]
which gives a contradiction. Thus \( V^\theta_0 = V^\tilde{\theta}_0 \). But then
\[
E^Q[H|\mathcal{F}_t] = E^Q[V^\theta_t|\mathcal{F}_t] = V^\theta_t = V^\tilde{\theta}_t = E^\tilde{Q}[V^\tilde{\theta}_t|\mathcal{F}_t] = E^\tilde{Q}[H|\mathcal{F}_t],
\]
(4.23) because \( V^\theta \) respectively \( V^\tilde{\theta} \) are martingales under \( Q \) respectively \( \tilde{Q} \).

When is there a unique risk-neutral price for all attainable claims?

Definition 4.4.4. An arbitrage-free market is called complete if every European contingent claim \( H \) such that there exists \( Q \in \mathcal{M}(P) \) with \( H \in L^1(\mathcal{F}^B_T, Q) \) is attainable.

Proposition 4.4.5. Let \( H \) be a European contingent claim such that there exists \( Q \in \mathcal{M}(P) \) with \( H \in L^1(\mathcal{F}^B_T, Q) \). If the market \( X \) is complete there exists a unique risk-neutral price process for \( H \), i.e.
\[
p_t(H) = E^Q[H|\mathcal{F}^B_t] = q_t(H) \quad \text{a.s.}
\]
for all \( Q \in \mathcal{M}(P) \) with \( H \in L^1(\mathcal{F}^B_T, Q) \).

Further the unique (discounted) risk-neutral price process is given by the hedge portfolio:
\[
E^Q[H|\mathcal{F}^B_t] = V^\theta_t \quad \text{a.s.}
\]
where \( Q \in \mathcal{M}(P) \) with \( H \in L^1(\mathcal{F}^B_T, Q) \) and \( \theta \) is a replicating strategy for \( H \).

Proof. Without loss of generality \( X^0 \equiv 1 \) (Remark 4.4.2). Assume first that \( H \) is bounded. Then a hedge portfolio \( V^\theta \) is also bounded and hence \( V^{-\theta} = -V^\theta \) is also admissible. Since \( V^\theta_T = H \) and \( V^{-\theta}_T = -H \) a.s.
\[
q_t(H) \leq V^\theta_t \leq p_t(H) \quad \text{a.s.}
\]
By Theorem 4.3.4 and Lemma 4.4.3 the result follows.

If \( H \) is unbounded with hedge portfolio \( V^\theta \) with respect to \( Q \in \mathcal{M}(P) \) consider the sequence of bounded claims
\[ H^n(\omega) := H(\omega) \land n \]

with hedge portfolio \( V^n \), \( n \in \mathbb{N} \). Since \( H^n \leq H \) and \( V^{-n} \) is admissible
\[ V_t^n \leq p_t(H) \quad \text{a.s. for all } n \in \mathbb{N}. \]

But since by Lemma 4.4.3 and monotone convergence
\[ V_t^n = E_Q\left[ H^n | \mathcal{F}_t^B \right] \xrightarrow{n \to \infty} E_Q\left[ H | \mathcal{F}_t^B \right] = V_t^\theta \]

we get
\[ V_t^\theta \leq p_t(H) \quad \text{a.s.}. \]

Also, because \( V_T^\theta = H \) a.s., obviously
\[ q_t(H) \leq V_t^\theta \quad \text{a.s.}, \]

and together with Theorem 4.3.4 there exists a unique risk-neutral price for \( H \).
But this means that for any \( \tilde{Q} \in \mathcal{M}(\mathbb{P}) \) with \( H \in L^1(\mathcal{F}_T^B, \tilde{Q}) \)
\[ E_{\tilde{Q}}\left[ H | \mathcal{F}_t^B \right] = E_Q\left[ H | \mathcal{F}_t^B \right] = V_t^\theta \quad \text{a.s.}. \]

\( \square \)

We now investigate under which conditions an arbitrage-free market is complete. We start with a preliminary result.

**Lemma 4.4.6.** Let \( \tilde{Q} \sim \mathbb{P} \) be an equivalent probability measure given by
\[ \frac{d\tilde{Q}}{d\mathbb{P}} = \mathcal{E} \left( \int u_s dB_s \right)_T, \]

with \( \tilde{Q} \)-Brownian motion \( \tilde{B}_t := B_t - \int_0^t u_s ds \). Then every \( H \in L^1(\mathcal{F}_T^B, \tilde{Q}) \)
admits the representation
\[ H = E_{\tilde{Q}}[H] + \int_0^T \Phi_t \cdot d\tilde{B}_t, \quad t \in [0, T], \quad (4.24) \]

where \( \Phi \) is a progressive (with respect to \( (\mathcal{F}_t^B)_{t \in [0, T]} \)) process with values in \( \mathbb{R}^d \), such that
\[ \int_0^T \|\Phi_s\|^2 ds < \infty, \quad (4.25) \]

and \( \int \Phi_s \cdot d\tilde{B}_s \) is a \( \tilde{Q} \)-martingale.

Note that the sigma algebra \( \mathcal{F}_t^\theta = \sigma(\tilde{B}_s, s \leq t) \) generated by \( \tilde{B}_t \) is contained in \( \mathcal{F}_t^B \), but not necessarily equal to \( \mathcal{F}_t^B \).
Proof. This Lemma is not a direct consequence of Theorem 2.9.2, since $\Phi$ in (4.24) is $(\mathcal{F}^B_t)_{t \in [0,T]}$-adapted, however may not be $(\mathcal{F}^\tilde{B}_t)_{t \in [0,T]}$-adapted, since $\mathcal{F}^B_t \subseteq \mathcal{F}^\tilde{B}_t$. We define $Z_T = \frac{dQ}{dP}$ and note that if $H \in L^1(\mathcal{F}^B_T, Q)$, then $HZ_T \in L^1(\mathcal{F}^B_T, P)$, since

$$E[HZ_T] = E^Q[H].$$

By Theorem 2.9.2 there exists $\Psi \in \Lambda^2_{\text{loc,d}}(T)$, such that

$$E[HZ_T|\mathcal{F}^B_t] = E[HZ_T] + \int_0^t \Psi_s \cdot dB_s,$$

in particular, for $t = T$ we have

$$HZ_T = E[HZ_T] + \int_0^T \Psi_s \cdot dB_s.$$ 

Given $A \in \mathcal{F}^B_t$, we have

$$\int_A HZ_T dP = \int_A H dQ = \int_A E^Q[H|\mathcal{F}^B_t] dQ = \int_A E^Q[H|\mathcal{F}^B_t] Z_T dP = \int_A E^Q[H|\mathcal{F}^B_t] E[Z_T|\mathcal{F}^B_t] dP,$$

i.e. we obtain the well-known Bayes rule for the conditional expectation,

$$E^Q[H|\mathcal{F}^B_t] = \frac{E[HZ_T|\mathcal{F}^B_t]}{E[Z_T|\mathcal{F}^B_t]} = \frac{N_t}{Z_t},$$

where $Z_t = E[Z_T|\mathcal{F}^B_t]$ and we have defined $N_t = E[HZ_T|\mathcal{F}^B_t]$. By the integration by parts formula we obtain

$$d \left( \frac{N_t}{Z_t} \right) = \frac{1}{Z_t} dN_t + N_t d \left( \frac{1}{Z_t} \right) + \frac{1}{Z_t^2} dB_t + \frac{1}{Z_t} \left( \frac{u_t}{Z_t} \cdot dB_t + \frac{\|u_t\|^2}{Z_t} dt \right) - \Psi_t \cdot \frac{u_t}{Z_t} dt$$

$$= \frac{1}{Z_t} (\Psi_t - u_t N_t) \cdot dB_t - \frac{u_t}{Z_t} \cdot (\Psi_t - u_t N_t) dt$$

$$= \frac{1}{Z_t} (\Psi_t - u_t N_t) \cdot [dB_t - u_t dt]$$

$$= \frac{1}{Z_t} (\Psi_t - u_t N_t) \cdot \tilde{d}B_t,$$  \hspace{1cm} (4.26)
which yields \( \Phi = (\Phi_t)_{t \in [0,T]} \in A^2_{loc,d}(T) \), since \((Z_t)_{t \in [0,T]}, (N_t)_{t \in [0,T]} \) are continuous and \( \Psi, \eta \in A^2_{loc,d}(T) \), which means the stochastic integral in (4.26) is well-defined. Hence

\[ H = E^Q [H] + \int_0^T \Phi_t \cdot d\tilde{B}_t, \]

and (4.25) holds. Further, \( \int_0^T \Phi_t \cdot d\tilde{B}_t = z_t = E^Q [H | \mathcal{F}_T] \) is a \( Q \)-martingale. \( \square \)

With the help of the preceding lemma, we will now be able to prove the next important theorem, which characterizes market completeness in terms of the rank of the volatility matrix \( \sigma(t, \omega) \).

**Theorem 4.4.7.** The following are equivalent:

i) The market \( X = (X_t)_{t \in [0,T]} \) is complete.

ii) For a.e. \((t, \omega) \in [0,T] \times \Omega \) the volatility matrix \( \sigma(t, \omega) \) has a left-inverse, i.e. there exists a progressive process \( \Lambda(t, \omega) \) with values in \( \mathbb{R}^{d \times n} \), such that

\[ \Lambda(t, \omega) \sigma(t, \omega) = I_d \text{ for a.e. } (t, \omega) \in [0,T] \times \Omega. \]  

(4.27)

iii) There exists a unique equivalent local martingale measure; i.e.

\[ \mathcal{M}(P) = \{ Q \}. \]

**Proof.** ii) \Rightarrow i) Suppose that (4.27) holds. Given a contingent claim \( H \in L^1(\mathcal{F}_T^B, Q) \), we want to prove that there exists an admissible portfolio \( \theta = (\theta_t)_{t \in [0,T]} \), where \( \theta_t = (\theta^1_t, \ldots, \theta^n_t) \) such that with \( V_0^\theta = z \)

\[ H = z + \int_0^T \theta_s \cdot dX_s. \]

Equivalently, by using (4.18) where we derived the dynamics of \( X \) under \( Q \),

\[ \frac{H}{X_T^0} = z + \int_0^T \theta_s \cdot dX_s = z + \int_0^T \sum_{k=1}^d \left( \sum_{i=1}^n \theta^i_k \sigma^i,k \right) \frac{d\tilde{B}_k}{X^i_0} \left( \frac{\sigma^{i,k}}{X^i_0} \right) d\tilde{B}_k, \]

(4.28)

where \( \tilde{B}_k = B_k - u_k dt \). By Lemma 4.4.6 we have that there exists an \((\mathcal{F}_t^B)_{t \in [0,T]}\)-progressiv \( \mathbb{R}^d \)-valued process \( \Phi \), such that

\[ \int_0^T \| \Phi_s \|^2 ds < \infty \]

and

\[ \mathbb{E}^Q \left[ \frac{H}{X_T^0} \bigg| \mathcal{F}_T^B \right] = \mathbb{E}^Q \left[ \frac{H}{X_T^0} \right] + \int_0^T \Phi_s \cdot d\tilde{B}_s, \]

i.e. for \( t = T \) we have
\[ \frac{H}{\mathcal{X}^u_T} = E^Q \left[ \frac{H}{\mathcal{X}^u_T} \right] + \int_0^T \Phi_s \cdot d\tilde{B}_s. \] (4.29)

Comparing (4.28) and (4.29) we define \( z = E^Q \left[ \frac{H}{\mathcal{X}^u_T} \right] \), and choose \( \hat{\theta} = (\theta^1, \ldots, \theta^n) \) (investment in the risky assets, without the bank account) such that for all \( t \in [0, T] \)

\[ \sum_{i=1}^n \frac{\theta_i^t \sigma_i^t}{X^u_t} = \Phi_t^k, \quad k = 1, \ldots, d, \]

or in vectorial notation

\[ \frac{1}{X^u_t} \hat{\theta}_t^{tr} \sigma_t = \Phi_t^{tr}. \]

Since \( \sigma_t \) has a left-inverse by assumption, we obtain

\[ \hat{\theta}_t^{tr} = X^u_t \Phi_t^{tr} \Lambda_t, \quad t \in [0, T]. \]

Then we can complete the portfolio choosing \( \theta_0^t \) such that \( \theta_t = (\theta_0^t, \hat{\theta}_t) = (\theta_0^t, \theta_1^t, \ldots, \theta_n^t) \) is self-financing with initial capital \( V_0 = z = E^Q \left[ \frac{H}{\mathcal{X}^u_T} \right] \) (Lemma 4.1.10). The associated portfolio value is then

\[ \nabla_t = z + \int_0^t \theta_s \cdot d\mathcal{X}_s, \quad t \in [0, T], \]

and by construction

\[ \nabla_t = E^Q \left[ \frac{H}{\mathcal{X}^u_T} \bigg| \mathcal{F}_t \right], \]

which means that \( \nabla \) is a \( \mathcal{Q} \)-martingale. Moreover \( \nabla_t \) is bounded from below since \( H \geq 0 \). Since we have shown that \( \theta \) is an admissible strategy in the discounted market \( \mathcal{X} \), by Remark 4.2.3 the proof of this direction is complete.

i) \( \Rightarrow \) ii)) Now assume that the market is complete, i.e. every contingent claim is attainable. Then, again by Remark 4.4.2, the discounted market \( \mathcal{X} \) is also complete, and therefore we can assume that \( X^0 \equiv 1 \). Let \( \Phi \) be an \( (\mathcal{F}_t^B)_{t \in [0, T]} \)-adapted process, such that \( E^Q \left[ \int_0^T \| \Phi_s \|^2 ds \right] < \infty \). Then \( H = \int_0^T \Phi_s \cdot d\tilde{B}_s \in L^2(\mathcal{F}_t^B, \mathcal{Q}) \) by the definition of the stochastic integral (under \( \mathcal{Q} \)), i.e. we have the analogon to equation (4.29) with \( z = E[H] = 0 \), which is obvious by the properties of the stochastic integral \( \int_0^T \Phi_s \cdot d\tilde{B}_s \). Note that \( \Phi \) is also \( (\mathcal{F}_t^B)_{t \in [0, T]} \)-adapted because \( \mathcal{F}_t^B \subseteq \mathcal{F}_t^B \) for all \( t \in [0, T] \). Since we have assumed that the market is complete and using the same argumentation as in (4.28), there exists an admissible portfolio \( (\theta_0^t, \ldots, \theta_n^t) \) such that

\[ H = 0 + \int_0^T \hat{\theta}_t^{tr} \sigma_t \cdot d\tilde{B}_t, \]
96 4 Financial Markets in Continuous Time

where \( \hat{\theta}_t = (\theta^1_t, \ldots, \theta^n_t) \). Conditioning with respect to \( \mathcal{F}^B_t \) gives us

\[
\int_0^t \hat{\theta}_s^\tau \sigma_s \cdot d\tilde{B}_s
\]

for every \( t \in [0, T] \), hence

\[
\Phi^\tau(t, \omega) = \hat{\theta}^\tau(t, \omega) \sigma(t, \omega)
\]

for a.a. \((t, \omega) \in [0, T] \times \Omega \). Hence \( \Phi(t, \omega) \in \mathbb{R}^d \) belongs to the linear span of the rows \( \sigma_i(t, \omega), i = 1, \ldots, n, \) of the matrix \( \sigma(t, \omega) \in \mathbb{R}^{n \times d} \). Since \( \Phi(t, \omega) \) was arbitrary, then we conclude that the span of the rows of \( \sigma(t, \omega) \) must be \( \mathbb{R}^d \) for a.a. \((t, \omega) \). Therefore rank \( \sigma(t, \omega) = d \), hence there exists \( A(t, \omega) \in \mathbb{R}^{d \times n} \), such that

\[
A(t, \omega)\sigma(t, \omega) = I_d.
\]

To prove that \( A(t, \omega) \) is also progressive, we refer to Karatzas and Shreve [9]. 

ii) \( \iff \) iii)) can be derived from the Girsanov structure of equivalent local martingale measures presented in Proposition 4.2.7.

We summarize our findings regarding market completeness and the role of the volatility matrix \( \sigma \) as follows:

- i) \( n = d \): the market is complete \( \iff \sigma(t, \omega) \) is invertible for a.a. \((t, \omega)\).
- ii) \( n > d \): the market is complete \( \iff \sigma(t, \omega) \) is left-invertible for a.a. \((t, \omega)\).
- iii) \( n < d \): the market cannot be complete.

That the market cannot be complete when \( n < d \) also makes sense intuitively, since in this case the number of assets available for trading is less than the number of sources of randomness \((B^1, \ldots, B^d)\) that affect the market.

In sum, we thus get the following procedure to price and hedge a claim \( H \in L^1(\mathbb{Q}) \) in a complete market with (unique) risk-neutral pricing measure \( \mathbb{Q} \) and \( \mathbb{Q} \)-Brownian motion \( \tilde{B} \):

i) Compute the (unique) risk-neutral price of \( H \) at \( t = 0 \) by

\[
E^\mathbb{Q} \left[ \frac{H}{X_T^\mathbb{Q}} \right].
\]

ii) Find a \((\mathcal{F}^B_t)_{t \in [0, T]}\)-progressive process \( \Phi \) such that \( \int \Phi \cdot d\tilde{B} \) is a \( \mathbb{Q} \)-martingale and

\[
\frac{H}{X_T^\mathbb{Q}} = E^\mathbb{Q} \left[ \frac{H}{X_T^\mathbb{Q}} \right] + \int_0^T \Phi_s d\tilde{B}_s.
\]

iii) Compute the hedge strategy \( \theta = (\theta^0, \hat{\theta}^\tau) \) with \( \hat{\theta} = (\theta^1, \ldots, \theta^n) \) by

\[
\hat{\theta}^\tau := X_t^0 \Phi^\tau_t A_t,
\]

\[
\theta^0 = \nabla^\theta_t - \hat{\theta}_t \cdot X_t,
\]

where \( A_t \) is the left-inverse of \( \sigma_t \) and the discounted hedge portfolio value \( \nabla^\theta_t \) is given by

\[
\nabla^\theta_t = E^\mathbb{Q} \left[ \frac{H}{X_T^\mathbb{Q}} \bigg| \mathcal{F}^B_t \right].
\]
The essential step is thus to find $\Phi$ in $ii)$, which in general can be formulated in terms of the so called Malliavin calculus. In the case of an Itô diffusion model, however, there is a simpler way that we will present in the next section.
4.5 Pricing and Hedging in Itô Diffusions Models via Partial Differential Equations

Consider the \( n \)-dimensional Itô diffusion (SDE) given by
\[
dX_t = \mu(X_t)dt + \sigma(X_t)dB_t.
\]
(4.30)

We assume

i) the coefficients \( \mu : \mathbb{R}^n \to \mathbb{R}^n \), \( \sigma : \mathbb{R}^n \to \mathbb{R}^{n \times d} \) are continuous and satisfy the linear growth condition in Definition 3.2.1.

ii) there exists a weakly unique weak solution of (4.30) denoted by \( (X_t^x)_{t \in [0,T]} \) when it starts in \( x \in \mathbb{R}^n \) \( (X_0^x = x) \).

**Remark 4.5.1.** By Theorem 3.2.2, assumptions i) and ii) above are fulfilled when \( \mu \) and \( \sigma \) fulfill the Itô assumptions in Definition 3.2.1.

Now, consider the partial differential equation (PDE)
\[
-\frac{\partial v}{\partial t}(t,x) = (L - k)v(t,x), \quad (t,x) \in [0,T) \times \mathbb{R}^n,
\]
(4.31)

where \( L - k \) denotes the differential operator
\[
(L - k)v := \sum_{i=1}^n \mu_i(x) \frac{\partial v}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n (\sigma \sigma^T)_{ij}(x) \frac{\partial^2 v}{\partial x_i \partial x_j} - k(x)v,
\]
and \( f : \mathbb{R}^n \to \mathbb{R}, k : \mathbb{R}^n \to \mathbb{R}_+ \) are continuous functions such that
\[
|f(x)| \leq L(1 + \|x\|^{2\lambda}) \quad \text{or} \quad f(x) \geq 0; \quad x \in \mathbb{R}^n,
\]
for some \( L > 0, \lambda \geq 1 \).

**Theorem 4.5.2 (Backward Feynman-Kac formula).** Under the preceding assumptions, suppose \( v(t,x) \in C^{1,2}([0,T) \times \mathbb{R}^n) \) is a solution of the PDE (4.31) such that
\[
\max_{0 \leq t \leq T} |v(t,x)| \leq M(1 + \|x\|^{2\mu}), \quad x \in \mathbb{R}^n,
\]
(4.32)

for some \( M > 0, \mu \geq 1 \). Then \( v(t,x) \) is unique and admits the stochastic representation
\[
v(t,x) = E \left[ f(X_T^x) \exp \left\{ - \int_0^{T-t} k(X_s^x)ds \right\} \right]
\]
on \([0,T] \times \mathbb{R}^n\).
Remark 4.5.3. a) A set of sufficient conditions for the existence of a solution $v \in C^{1,2}([0, T] \times \mathbb{R}^n)$ as in Theorem 4.5.2 is:
   i) Uniform ellipticity: There exists a $\delta > 0$ such that
   $$x^T \sigma \sigma^T (y)x \geq \delta \|x\|^2$$
   for all $x, y \in \mathbb{R}^n$.
   ii) $\sigma, \mu, k$ are bounded.
   iii) $\sigma, \mu, k$ are uniformly Hölder continuous.
   iv) $|f(x)| \leq L(1 + \|x\|^2 \lambda)$; $L \geq 0$, $\lambda \geq 1$.

b) Theorem 4.5.2 can also be formulated for a more general time-inhomogeneous Itô diffusion setting.

c) For less restrictive (exponential growth) or more restrictive (boundedness) growth conditions on $v(t, x)$ than the polynomial growth given in (4.32) one can formulate more or less restrictive conditions on the functions $\mu, \sigma, f,$ and $k$ such that Theorem 4.5.2 still holds.

In the following proof of Theorem 4.5.2 we assume $\sigma$ to be bounded and the solution $v(t, x)$ to be bounded with bounded derivatives for simplicity. For the general proof and more information on Remark 4.5.3 we refer to Karatzas and Shreve [9], Theorem 5.7.6, and references therein.

Proof (Theorem 4.5.2).

Define
$$w(s, x) := v(t + s, x)$$
$$Z_t := \exp \left\{-\int_0^t k(X^*_s)ds \right\}.$$  

Itô’s formula gives for $0 \leq s \leq T - t$
$$\begin{align*}
  \text{d} (w(s, X^*_s)Z_s) &= -w(s, X^*_s)k(X^*_s)Z_s \text{d}s + Z_s \text{d}w(s, X^*_s) \\
  &= Z_s \left\{ \frac{\partial}{\partial s} w(s, X^*_s) + (L - k)w(s, X^*_s) \right\} \text{d}s \\
  &+ Z_s \nabla_x w(s, X^*_s) \sigma(X^*_s) \text{d}B_s,
\end{align*}$$

where $\nabla_x w(s, x) = (\partial_{x_1} w(s, x), \ldots, \partial_{x_n} w(s, x))$ is the gradient of $w$. Since
$$\frac{\partial}{\partial s} w + (L - k)w = 0 \text{ on } [0, T - t) \times \mathbb{R}^n$$

and $Z_s \nabla_x w \sigma \in \Lambda^2_T$ by boundedness assumptions, we get by taking expectation
$$E \left[ w(T - t, X^*_t)Z_{T-t} \right] = E \left[ w(0, X^*_0)Z_0 \right]$$
$$\iff \quad E \left[ v(T, X^*_t)Z_{T-t} \right] = E \left[ v(t, x) \cdot 1 \right]$$
$$\iff \quad E \left[ f(X^*_t) \exp \left\{-\int_0^{T-t} k(X^*_s)ds \right\} \right] = v(t, x).$$

$\blacksquare$
Now let our financial market model under a given risk-neutral pricing measure $Q$ be given by the Itô diffusion

$$
\begin{align*}
\text{d}X_t^0 &= r(\hat{X}_t^x)X_t^0 \text{d}t, \quad X_0^0 = 1; \\
\text{d}\hat{X}_t^x &= r(\hat{X}_t^x)\hat{X}_t^x \text{d}t + \sigma(\hat{X}_t^x)\text{d}\tilde{B}_t, \quad \hat{X}_0^x = x \in \mathbb{R}^n; \\
\end{align*}
$$

where $\tilde{B}$ is $\mathbb{R}^d$-valued $Q$-Brownian motion, $r: \mathbb{R}^n \to \mathbb{R}_+$ bounded, $\sigma: \mathbb{R}^n \to \mathbb{R}^{n \times d}$ fulfills the Itô assumptions of Definition 3.2.1, and $\hat{X} = (X^1, \ldots, X^n)$ denotes the risky assets. Consider a derivative contract with payoff $\varphi(\hat{X}_T^x)$ for some function $\varphi: \mathbb{R}^n \to \mathbb{R}_+$ such that $\mathbb{E}^Q[\varphi(\hat{X}_T^x)] < \infty$. The risk neutral price process was then given by

$$
\pi_t = \mathbb{E}^Q \left[ \varphi(\hat{X}_T^x) \exp \left\{ - \int_t^T r(\hat{X}_s^x) \text{d}s \right\} \right]_{\mathcal{F}_t},
$$

where in the last equality we used the time homogeneous Markov property of $\hat{X}$ (Proposition 3.3.1). Assume now that with $\mu = r$, $k = r$, and $f = \varphi$. Theorem 4.5.2 holds. Then the risk-neutral price is given by

$$
\pi_t = v(t, \hat{X}_t^x), \quad (4.33)
$$

where $v(t, x) \in \mathcal{C}^{1,2}([0, T] \times \mathbb{R}^n)$ uniquely solves the PDE (4.31). Further, as in the proof of Theorem 4.5.2 with

$$
Z_t := \exp \left\{ \int_0^t -r(\hat{X}_s^x) \text{d}s \right\} = \frac{1}{X_t^0}
$$

we see by Itô’s formula that the discounted price process is given by

$$
\pi_t Z_t = v(t, \hat{X}_t^x)Z_t
$$

$$
= v(0, \hat{X}_0^x) + \int_0^t Z_s \nabla_x v(s, \hat{X}_s^x) \sigma(\hat{X}_s^x) \text{d}\tilde{B}_s
$$

$$
= v(0, x) + \int_0^t \nabla_x v(s, \hat{X}_s^x) \text{d}\hat{X}_s.
$$

Since $\pi_t Z_t \geq 0$ is a $Q$-martingale, it follows that

$$
\theta_i^* := \frac{\partial v}{\partial x_i}(s, \hat{X}_s^x) = \frac{\partial}{\partial y_i} \mathbb{E} \left[ \varphi(\hat{X}_{T-t}^y)Z_{T-t} \right]_{y=\hat{X}_s^x}, \quad i = 1, \ldots, n, \quad (4.34)
$$

determines the hedge portfolio of the claim (with respect to $Q$). We call $\frac{\partial v}{\partial \sigma}(s, \hat{X}_s^x)$ Delta or Delta hedge.
We now focus our attention to the Black-Scholes model as introduced in Example 4.1.2, i.e. $n = d = 1$ and

bank account: \[
\begin{align*}
\frac{dX^0_t}{dt} &= rX^0_t; \quad r \in \mathbb{R}_+ \\
X^0_0 &= 1
\end{align*}
\]

risky asset: \[
\begin{align*}
\frac{dX^1_t}{dt} &= X^1_t(\mu dt + \sigma dB_t); \quad \mu, \sigma \in \mathbb{R}_+ \\
X^1_0 &= x; \quad x \in \mathbb{R}_+
\end{align*}
\]

We recall that for $t \in [0, T]$

\[
X^0_t = e^{rt}, \\
X^1_t = xe^{(\mu - \frac{\sigma^2}{2})t + \sigma B_t}.
\]

From Example 4.2.12 we know the market is complete with unique equivalent martingale measure $Q$ given by

\[
\frac{dQ}{dP} = e^{-\frac{\mu - r}{2} B_T - \frac{1}{2} (\frac{\mu - r}{\sigma})^2 T}.
\]

We have

\[
\frac{dX^1_t}{dt} = rX^1_t dt + \sigma X^1_t d\tilde{B}_t,
\]

i.e.

\[
X^1_t = xe^{(r - \frac{\sigma^2}{2})t + \sigma \tilde{B}_t}
\]

with $Q$-Brownian motion $\tilde{B}_t = B_t + \frac{\mu - r}{\sigma} t$.

We first develop the famous Black & Scholes formula to price and hedge call and put options before we consider further aspects in the Black-Scholes model including the pricing of exotic derivatives in the following subsections.
5.1 The Black-Scholes formula

Consider a derivative contract with pay off \( \varphi(X^1_T) \in L^1(Q) \) for some function \( \varphi: \mathbb{R}_+ \rightarrow \mathbb{R}_+ \). Then the discounted risk-neutral price process is given by

\[
E^Q \left[ \varphi(X^1_T) e^{-rT} | \mathcal{F}_t \right] = e^{-rT} E^Q \left[ \varphi(X^1_{T-t}) \right] \bigg|_{y=X^1_t} = e^{-rT} E^Q \left[ \varphi \left( ye^{(r-x^2/2)(T-t) + \sigma \tilde{B}(T-t)} \right) \right] \bigg|_{y=X^1_t}.
\]

where in the first equality we have used the time homogeneous Markov property of \( X^1_t \) (Proposition 3.3.1). Note that the superscript 1 in \( X^1_t \) refers to the risky asset and not a time point. Now let \( \varphi(X) = (X - K)^+ \) be the pay off of a call option with strike \( K \). Then (5.1) can be written as

\[
e^{-rT} E^Q \left[ (X^1_{T-t} - K) \mathbb{1}_{\{X^1_{T-t} \geq K\}} \right] \bigg|_{y=X^1_t} = e^{-rT} E^Q \left[ X^1_{T-t} \mathbb{1}_{\{X^1_{T-t} \geq K\}} \right] y=X^1_t.
\]

We have

\[
Q \left( X^1_{T-t} \geq K \right) = Q \left( \sigma \tilde{B}_{T-t} \geq \log \frac{K}{y} + \left( \frac{\sigma^2}{2} - r \right)(T-t) \right)
\]

\[
= Q \left( Z \geq + \frac{\log \frac{K}{y} + \left( \frac{\sigma^2}{2} - r \right)(T-t)}{\sigma \sqrt{T-t}} \right) := z_2
\]

\[
= Q \left( Z \leq + \frac{\log \frac{K}{y} + \left( r - \frac{\sigma^2}{2} \right)(T-t)}{\sigma \sqrt{T-t}} \right)
\]

where \( Z \) is a standard normal random variable. Further,

\[
E^Q \left[ X^1_{T-t} \mathbb{1}_{\{X^1_{T-t} \geq K\}} \right] = ye^{r(T-t)} \int_{z_2}^{\infty} e^{\sigma \sqrt{T-t} z - \frac{\sigma^2}{2}(T-t) + \frac{z^2}{2}} \frac{1}{\sqrt{2\pi}} dz
\]

\[
= ye^{r(T-t)} \int_{z_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z - \frac{r(\sigma^2/2) - x^2)}{2}}{2}} dz
\]
5.1 The Black-Scholes formula

\[
= ye^{r(T-t)} \int_{x_2 - \sigma \sqrt{T-t}}^{\infty} e^{-\frac{z^2}{2}} \, dz
\]

\[
= ye^{r(T-t)} \int_{-\infty}^{-x_2 + \sigma \sqrt{T-t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \, dz
\]

\[
= ye^{r(T-t)} Q\left(Z \leq \frac{\log \left(\frac{x}{K}\right) + \left(r \pm \frac{\sigma^2}{2}\right) (T-t)}{\sigma \sqrt{T-t}}\right).
\]

Putting everything together we obtain the first part of the following Theorem.

Theorem 5.1.1 (Black-Scholes formula).

i) The risk-neutral price process of a call option with maturity \(T\) and strike \(K\) written on the underlying \(X^1\) in the Black-Scholes model is given by

\[
C_t(X^1_t, K, T) = X^1_t N\left(x_1(t, X^1_t)\right) - Ke^{-r(T-t)} N\left(x_2(t, X^1_t)\right)
\]

where \(N(\cdot)\) denotes the distribution function of a standard normal random variable and

\[
x_{1,2}(t, x) := \frac{\log \left(\frac{x}{K}\right) + \left(r \pm \frac{\sigma^2}{2}\right) (T-t)}{\sigma \sqrt{T-t}}.
\]

ii) The hedge strategy of the call option is given by the Delta hedge

\[
\theta^1_t = N(x_1(t, X^1_t))
\]

and the corresponding investment in the bank account

\[
\theta^0_t = \nabla_t - \theta^1_t X^1_t = -Ke^{-rT} N(x_2(t, X^1_t)).
\]

Remark 5.1.2. Note that \(\theta^1_t \geq 0, \theta^0_t \leq 0\) for all \(t \in [0, T]\), i.e. in order to hedge the call option one is always long in the underlying and borrows money from the bank account. The situation is opposite for a put option (see later in this section).

Proof. It remains to show ii) of Theorem 5.1.1. We remark that all assumptions of Theorem 4.5.2 are satisfied for the call option in the Black-Scholes model (i.e. \(f(x) = (x - K)^+\)). From (4.33) we thus get that the call price process is also given in terms of the solution of a partial differential equation:

\[
C_t = X^1_t N\left(x_1(t, X^1_t)\right) - Ke^{-r(T-t)} N\left(x_2(t, X^1_t)\right) = v(t, X^1_t)
\]

where \(v(t, x) \in C^{1,2}([0, T] \times \mathbb{R})\) solves the Black-Scholes partial differential equation

\[
\begin{aligned}
\frac{\partial v}{\partial t}(t, x) + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 v}{\partial x^2}(t, x) + rx \frac{\partial v}{\partial x}(t, x) - rv(t, x) &= 0 \\
v(T, x) &= (x - K)^+
\end{aligned}
\]
Further, we then know from Section 4.5 that the hedge strategy is determined by the Delta:

$$\theta_t^1 = \frac{\partial v}{\partial x}(t, X_t^1)$$

which we now compute explicitly for the call option. We have

$$v(t, x) = xN(x_1(t, x)) - Ke^{-r(T-t)}N(x_2(t, x))$$

$$= x \int_{-\infty}^{\log \frac{x}{K} + \left(\frac{r + \sigma^2}{2}\right)(T-t)} \frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}} ds - Ke^{-r(T-t)} \int_{-\infty}^{\log \frac{x}{K} + \left(\frac{r - \sigma^2}{2}\right)(T-t)} \frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}} ds.$$

We obtain

$$\frac{\partial v}{\partial x}(t, x) = N(x_1(t, x))$$

$$+ x \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{x \sigma \sqrt{T-t}} \exp\left(\frac{-1}{2} \left(\log \frac{x}{K} + \left(\frac{r + \sigma^2}{2}\right)(T-t)\right)^2 \sigma^2(T-t)\right)$$

$$- \frac{Ke^{-r(T-t)}}{x \sqrt{2\pi} \sigma \sqrt{T-t}} \exp\left(\frac{-1}{2} \left(\log \frac{x}{K} + \left(\frac{r - \sigma^2}{2}\right)(T-t)\right)^2 \sigma^2(T-t)\right).$$

Note that

$$- \frac{1}{2} \left(\log \frac{x}{K} + \left(\frac{r - \sigma^2}{2}\right)(T-t)\right)^2 \sigma^2(T-t) - r(T-t)$$

$$= - \frac{1}{2\sigma^2(T-t)} \left[\left(\log \frac{x}{K}\right)^2 + 2 \left(\frac{r - \sigma^2}{2}\right)(T-t) \log \frac{x}{K}\right]$$

$$- \frac{1}{2\sigma^2} \left(\frac{r - \sigma^2}{2}\right)^2 (T-t) - r(T-t)$$

$$- \frac{1}{2\sigma^2} \left(r^2 + \frac{\sigma^4}{4} - r\sigma^2 + 2\sigma^2 r(T-t)\right)$$

$$= - \frac{1}{2\sigma^2(T-t)} \left[\left(\log \frac{x}{K} + \left(\frac{r + \sigma^2}{2}\right)(T-t)\right)^2 - 2\sigma^2(T-t) \log \frac{x}{K}\right],$$

hence

$$\exp(-r(T-t)) \exp\left(\frac{-1}{2} \left(\log \frac{x}{K} + \left(\frac{r - \sigma^2}{2}\right)(T-t)\right)^2 \sigma^2(T-t)\right).$$
5.1 The Black-Scholes formula

\[ v(t, x) = \exp \left( \log \frac{x}{K} \right) \exp \left( - \frac{1}{2\sigma^2(T-t)} \left( \log \frac{x}{K} + \left( r + \frac{\sigma^2}{2} \right)(T-t) \right)^2 \right) \]

\[ = \frac{x}{K} n(x), \]

and putting everything together we have

\[ \frac{\partial v(t, x)}{\partial x} = N \left( x_1(t, x) \right) + \frac{1}{\sigma \sqrt{2\pi(T-t)}} \cdot \left( \frac{n(x) - K \cdot x}{x} \right) \]

\[ = N \left( x_1(t, x) \right), \]

i.e.

\[ \theta_1^t = N \left( x_1(t, X_1^t) \right) \]

and

\[ \theta_0^t = \nabla_t - \theta_1^t \nabla_{X_1^t} = -Ke^{-rT}N \left( x_2(t, X_1^t) \right). \]

This is Theorem 5.1.1 ii). \( \square \)

The Black-Scholes price process of a put option with pay off \((K - X_1^t)^+\) can now easily be obtained via the Put-Call-Parity (see Section 6.2). Since

\[ (X_1^t - K)^+ - (K - X_1^t)^+ = X_1^t - K \]

we get by taking \(Q\)-conditional expectations that the put price process \(P_t\) is given by

\[ P_t = C_t - F_t \]

where \(C_t\) is the call price from Theorem 5.1.1 and

\[ F_t = e^{rt}E^Q \left[ X_1^t - \frac{K}{e^{rT}} \cdot \frac{\theta_1^t}{\sigma \sqrt{T-t}} \right] = X_1^t - Ke^{-r(T-t)}. \]

We immediately obtain the Black & Scholes formula for the put option

\[ P_t = X_1^t \left[ N(x_1(t, X_1^t)) - 1 \right] + Ke^{-r(T-t)} \left[ 1 - N(x_2(t, X_1^t)) \right] \]

\[ = -X_1^t N(-x_1(t, X_1^t)) + Ke^{-r(T-t)}N(-x_2(t, X_1^t)), \]

where \(x_1(t), x_2(t)\) are given by

\[ x_{1,2}(t) = \frac{\log \left( \frac{X_1^t}{K} \right) + \left( r \pm \frac{\sigma^2}{2} \right)(T-t)}{\sigma \sqrt{T-t}}. \]

Further, analogously to the hedge of a call option we compute the hedge strategy of a put option by computing the Delta:
\[
\begin{align*}
\theta_t^1 &= \frac{\partial}{\partial x} C_t - \frac{\partial}{\partial x} F_t = N(x_1(t, X_1^t)) - 1, \\
\theta_t^0 &= V_t - \theta_t^1 X_t = K e^{-rT} \left[ 1 - N(x_2(t, X_1^t)) \right].
\end{align*}
\]

We thus see that in order to hedge a put option in the Black & Scholes market one is always short in the underlying and puts money into the bank account.

### 5.2 Sensitivity Analysis

Consider the Black-Scholes model and let \(v(t, x)\) be the pricing function at time \(t\) for a portfolio based on the underlying asset \(X_1^t\) (for example a portfolio based on the asset \(X_1^t\) and derivatives on \(X_1^t\)). For practical purposes it is interesting to study the sensitivity of the price with respect to variations in the asset price \(X_1^t\) and with respect to the parameters of the model. In the first case, we wish to measure our risk exposure, in the second one we want to check our sensitivity to misspecification of model parameters. We introduce the so-called greeks:

- i) Delta: \(\frac{\partial v}{\partial x}(t, x)\)
- ii) Gamma: \(\frac{\partial^2 v}{\partial x^2}(t, x)\)
- iii) Rho: \(\frac{\partial v}{\partial r}(t, x)\)
- iv) Theta: \(\frac{\partial v}{\partial t}(t, x)\)
- v) Vega: \(\frac{\partial v}{\partial \sigma}(t, x)\)

If the portfolio value does not change with respect to a small change in one of the parameters it is said to be neutral with respect to this parameter. Formally this means that the corresponding greek is zero. We have already seen, that if \(v(t, x)\) is the replicating portfolio of a derivative \(H = \varphi(X_1^T)\), then the Delta provides us with the strategy component with respect to \(X_1^t\). We show now how to construct a delta- and gamma-neutral portfolio for \(v(t, x)\). Note that for the underlying stock, the delta and gamma are obviously

\[
\Delta_x = 1 \quad \text{and} \quad \Gamma_x = 0,
\]

hence we cannot use only the stock to change the gamma of a portfolio. Therefore we also consider a call option with price \(C(t, x)\). We denote by \(y_1\) the number of units of the call option in the portfolio at time \(t\) and by \(y_2\) the number of units of stock. Then the value of the whole portfolio at time \(t\) including \(v(t, x)\), the call and the stock is

\[
V = v(t, x) + y_1 C(t, x) + y_2 x.
\]

\(\ast\) Recall that the pricing function actually depends on the four parameters \(t, r, \sigma\) and \(x\), e.g. Theorem 5.1.1 i).
We now impose
\[
\frac{\partial V}{\partial x} = \frac{\partial^2 V}{\partial x^2} = 0, \text{ i.e.}
\]
\[
\Delta_v + y_1 \Delta_C + y_2 = 0,
\]
\[
\Gamma_v + y_1 \Gamma_C = 0.
\]

Hence in order to obtain a delta- and gamma-neutral portfolio, we first choose \(y_1\) so that the portfolio is gamma-neutral and then choose \(y_2\) to make the portfolio delta-neutral,
\[
y_1 = -\frac{\Gamma_v}{\Gamma_C},
\]
\[
y_2 = \Delta_C \frac{\Gamma_v}{\Gamma_C} - \Delta_v.
\]

5.3 Historical and Implied Volatility

According to the Black-Scholes model, the pricing formula for a given contingent claim involves only the knowledge of the interest rate \(r\) (that we can observe directly on the market) and the volatility parameter of the stock price return. The latter is not directly observable and therefore must be estimated by means of statistical methods. A natural approach is to estimate the standard deviation based on historical time series of the underlying stock return. Although estimating the stock price volatility by this method in general is a rather simple procedure, the variance is often observed to be non-stationary, hence the sampling error cannot be reduced by using a longer series of historical observations or more frequent data. Furthermore, since the volatility is usually instable through time, estimating volatility by means of historical data is not helpful when aiming at predicting future volatility. Therefore, alternatively, it has also become very common to derive the volatility of an asset by examining the prices at which options on that asset are traded. If we consider the call option price at time \(t = 0\),
\[
C_0 = X_0^1 N\left(x_1(X_0^1, 0, T, \sigma)\right) - Ke^{-rT}N\left(x_2(X_0^1, 0, T, \sigma)\right),
\]
then the only unknown parameter is \(\sigma\), since \(X_0^1, C_0\) are the current prices of the asset and the call option and are available on the market. The implied volatility is the value of \(\sigma\) that substituted in (5.2) results in a model price equal to the current option price \(C_0\). It of course depends on the strike price \(K\) and on the remaining time \((T - t)\) to maturity. Considered as a function of \(K\), the implied volatility sometimes resembles a \(U\)-shape. This feature called volatility smile shows that the Black-Scholes model does not completely grasp the empirical structure of the option prices, since by assumption the volatility should actually be constant. Several alternative financial market models have
therefore been proposed, among them so-called stochastic volatility models, where also the volatility itself is described by a (continuous) stochastic process. A typical example is
\[ dX_t^1 = \mu(t, X_t^1)dt + \sigma_t X_t^1 dB_t, \]
\[ d\sigma_t = \alpha(t, \sigma_t)dt + \beta(t, \sigma_t)dB_t^1, \]
where \( B_t, B_t^1 \) are standard Brownian motions on \((\Omega, \mathcal{F}, \mathbb{P})\) with
\[ d\langle B, B^1 \rangle_t = \rho dt, \quad \rho \in \mathbb{R}. \]

5.4 Exotic Options

In the previous sections we have seen how to price put and call options in the Black-Scholes market. These kinds of contingent claims are sometimes called vanilla or plain vanilla options. Their payoffs depend only on the final value of the underlying asset. This section is now dedicated to the pricing and hedging of contingent claims whose payoff depends on the whole path of the underlying asset. These kinds of contracts are called path-dependent or exotic derivatives. As a typical example of an exotic option we will here focus on the lookback call option that has an explicit pricing formula based on the reflection principle for Brownian motion. However, first we will need some more preparation. The main reference for this section is Shreve [19].

Consider the Brownian motion with drift \( \hat{B} = (\hat{B}_t)_{t \in [0,T]} \) defined on \((\Omega, \mathcal{F}, (\mathcal{F}^B_t)_{t \in [0,T]}, \mathbb{P})\),
\[ \hat{B}_t = B_t + \alpha t, \quad 0 \leq t \leq T, \quad \alpha \in \mathbb{R}. \]

We define the maximum-to-date or running maximum \( \hat{M} = (\hat{M}_t)_{t \in [0,T]} \), where
\[ \hat{M}_t = \sup_{0 \leq s \leq t} \hat{B}_s. \]

We first note that \( B \) and \( \hat{B} \) generate the same filtration. Furthermore \( \hat{M}_T \geq 0 \) since \( \hat{B}_0 = B_0 = 0 \) and \( \hat{M}_T \geq \hat{B}_T \). Hence the pair \((\hat{M}_T, \hat{B}_T)\) takes values in the set
\[ \{(m, \omega) : \omega \leq m, \ m \geq 0\}. \]

We are now interested in computing the joint distribution of \((\hat{M}_T, \hat{B}_T)\). To start with we consider the case when \( \alpha = 0 \) (i.e. \( \hat{B}_t = B_t \)) and make the following preliminary observations. We fix a positive level \( m \) and a positive time \( t \). If we look at a specific path of the Brownian motion, maybe the path reached level \( m \) some time prior to \( t \) and at time \( t \) has again fallen below level \( m \), or the path is above level \( m \) at time \( t \). In any case, for each Brownian motion path that reaches level \( m \) prior to time \( t \) but is at a level \( \omega \) below to
m at time $t$, there exists a “reflected path” that is at level $2m - \omega$ at time $t$. If $\tau_m(\omega) = \inf\{t : \hat{B}_t = m\}$ is the stopping time that describes the first passage time through level $m$, then we obtain the reflected path by exchanging the up and down moves of $\hat{B}_t$ from time $\tau_m$ onwards. In particular we obtain the well-known reflection equality

$$P(\tau_m \leq t, \hat{B}_t \leq \omega) = P(\hat{B}_t \geq 2m - \omega)$$

(5.3)

for every $m > 0$ and $\omega \leq m$. Note that $P(\tau_m \leq t) = P(\hat{M}_t \geq m)$, i.e. the maximum of $\hat{B}$ over $[0, t]$ is at least $m$, if $\hat{B}$ over $[0, t]$ has reached level $m$.

**Lemma 5.4.1.** Suppose $\alpha = 0$. For every $t \in [0, T]$ the joint density of $(\hat{M}_t, \hat{B}_t)$ under $P$ is given by

$$f(m, \omega) = \frac{2(2m - \omega)}{t \sqrt{2\pi t}} e^{-\frac{(2m - \omega)^2}{2t}}, \quad \omega \leq m, \quad m > 0,$$

and 0 otherwise.

**Proof.** To compute the joint density we consider

$$P(\hat{M}_t \geq m, \hat{B}_t \leq \omega) = \int_m^\infty \int_{-\infty}^\omega f(x, y) \, dx \, dy.$$

Since by (5.3)

$$P(\hat{M}_t \geq m, \hat{B}_t \leq \omega) = P(\hat{B}_t \geq 2m - \omega),$$

we have

$$\int_m^\infty \int_{-\infty}^\omega f(x, y) \, dx \, dy = \int_0^\infty \int_{2m-\omega}^{+\infty} e^{-\frac{s^2}{2t}} \, ds,$$

hence

$$f(m, \omega) = -\frac{\partial^2 \psi}{\partial \omega \partial m}(m, \omega).$$

We obtain

$$\frac{\partial \psi}{\partial m}(m, \omega) = -\frac{2e^{-\frac{(2m-\omega)^2}{2t}}}{\sqrt{2\pi t}},$$

$$\frac{\partial \psi}{\partial \omega \partial m}(m, \omega) = -\frac{2}{\sqrt{2\pi t}} e^{-\frac{(2m-\omega)^2}{2t}} \cdot \frac{2m - \omega}{t},$$
and finally
\[ f(m, \omega) = \frac{2(2m - \omega)}{t^{\frac{1}{2}}} e^{-\frac{(2m-\omega)^2}{2t}} \]
for \( \omega \leq m, m > 0 \).

We now extend these results to the case when \( \alpha \neq 0 \).

**Proposition 5.4.2.** Suppose \( \alpha \neq 0 \). For every \( t \in [0, T] \) the joint density of \( (\hat{M}_t, \hat{B}_t) \) under \( \mathbb{P} \) is
\[ f(m, \omega) = \frac{2(2m - \omega)}{t^{\frac{1}{2}}} e^{-\frac{(2m-\omega)^2}{2t}} e^{\alpha \omega - \frac{\alpha^2 t}{2}}, \quad \omega \leq m, m \geq 0 \quad (5.4) \]
and 0 otherwise.

**Proof.** The basic idea of the proof is that we can go back to the case \( \alpha = 0 \) by changing the underlying probability measure. Set \( t > 0 \) and consider
\[ Z_t = \mathcal{E}(-\alpha B)_t = \exp \left( -\alpha B_t - \frac{\alpha^2 t}{2} \right) = \exp \left( -\alpha \hat{B}_t + \frac{\alpha^2 t}{2} \right). \]
Then by Girsanov’s theorem, \( \hat{B} = (\hat{B}_s)_{s \in [0, t]} \) is a \( \hat{Q} \)-Brownian motion under the equivalent measure \( \hat{Q} \) with density
\[ \frac{d\hat{Q}}{d\mathbb{P}} = Z_t. \]
Hence by Lemma 5.4.1 we have
\[ \hat{Q} \left( \hat{M}_t \geq m, \hat{B}_t \leq \omega \right) = \hat{Q} \left( \hat{B}_t \geq 2m - \omega \right) \]
and the joint density of \( (\hat{M}_t, \hat{B}_t) \) under \( \hat{Q} \) is
\[ f^\hat{Q}(m, \omega) = \frac{2(2m - \omega)}{t^{\frac{1}{2}}} e^{-\frac{(2m-\omega)^2}{2t}}, \quad \omega \leq m, m > 0 \]
and 0 otherwise. Then the joint distribution function of \( \hat{M}_t, \hat{B}_t \) under \( \mathbb{P} \) is given by
\[ F(m, \omega) = \mathbb{P} \left( \hat{M}_t \leq m, \hat{B}_t \leq \omega \right) = E \left[ \mathbf{1}_{\{\hat{M}_t \leq m, \hat{B}_t \leq \omega\}} \right] \]
\[ = E^{\hat{Q}} \left[ \frac{d\mathbb{P}}{d\hat{Q}} \mathbf{1}_{\{\hat{M}_t \leq m, \hat{B}_t \leq \omega\}} \right] \]
\[ = E^{\hat{Q}} \left[ e^{\alpha \hat{B}_t - \frac{\alpha^2 t}{2}} \mathbf{1}_{\{\hat{M}_t \leq m, \hat{B}_t \leq \omega\}} \right] \]
\[ = \int_{-\infty}^{m} \int_{-\infty}^{\omega} e^{\alpha x - \frac{\alpha^2 t}{2}} f^\hat{Q}(x, y) dx dy. \]
and the joint density is given by
\[ f(m, \omega) = \frac{\partial^2}{\partial m \partial \omega} F(m, \omega) = e^{\alpha \omega - \frac{\alpha^2}{2} t} f^\hat{Q}(m, \omega). \] (5.5)

For \( \omega \leq m, m > 0 \) (5.5) coincides with (5.4), otherwise \( f(m, \omega) \) is zero because \( f^\hat{Q} \) is zero.

\[ \square \]

**Corollary 5.4.3.** For \( t = T \), the distribution function and the density of the random variable \( \hat{M}_T \) are given by
\[ P(\hat{M}_T \leq m) = N\left( \frac{m - \alpha T}{\sqrt{T}} \right) - e^{2 \alpha m}N\left( \frac{-m - \alpha T}{\sqrt{T}} \right), \quad m \geq 0, \] (5.6)
\[ f_\hat{M}(m) = \frac{2}{\sqrt{2\pi T}} e^{-\frac{(m-\alpha T)^2}{2T}} - 2 \alpha e^{2 \alpha m}N\left( \frac{-m - \alpha T}{\sqrt{T}} \right), \quad m \geq 0, \] (5.7)
and 0 otherwise.

**Proof.** The distribution function of \( \hat{M}_T \) is given by
\[ P(\hat{M}_T \leq m) = \int_{-\infty}^{+\infty} \int_{-\infty}^{x} f(x,y) \, dy \, dx, \]
where \( f(x,y) \) is given by (5.4). For the sake of simplicity we only prove (5.6) in the case where \( \alpha = 0 \). Let \( m \geq 0 \), then we have
\[ P(\hat{M}_T \leq m) = \int_{0}^{m} \int_{-\infty}^{x} \frac{2(2x - y)}{T \sqrt{2\pi T}} e^{-\frac{(2x-y)^2}{2T}} \, dy \, dx \]
\[ = \int_{0}^{m} \int_{-x}^{+\infty} \frac{2z}{T \sqrt{2\pi T}} e^{-\frac{z^2}{2T}} \, dz \, dx \]
\[ = \int_{0}^{m} \frac{2}{\sqrt{2\pi T}} \left[ e^{-\frac{x^2}{2T}} \right]_{-\infty}^{+\infty} \, dx \]
\[ = \int_{0}^{m} \frac{2}{\sqrt{2\pi T}} e^{-\frac{x^2}{2T}} \, dx \]
\[ = \int_{0}^{m/\sqrt{T}} \frac{2}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du \]
\[ = 2N(m/\sqrt{T}) - 2N(0) \]
\[ = 2N(m/\sqrt{T}) - 1 \]
\[ = N(m/\sqrt{T}) - (1 - N(m/\sqrt{T})) \]
\[ = N(m/\sqrt{T}) - N(-m/\sqrt{T}). \]

To prove (5.7) we differentiate (5.6) and obtain
\[ f_\hat{M}(m) = e^{-\frac{(m-\alpha T)^2}{2T}} - 2 \alpha e^{2 \alpha m}N\left( \frac{-m - \alpha T}{\sqrt{T}} \right) + \frac{e^{2 \alpha m} e^{-\frac{(m+\alpha T)^2}{2T}}}{\sqrt{2\pi T}}. \]
\[
\frac{2}{\sqrt{2\pi T}}e^{-(m-\alpha T)^2/2} - 2\alpha e^{2\alpha m}N\left(-\frac{m-\alpha T}{\sqrt{T}}\right).
\]

We now apply the previous results to the pricing and hedging of barrier options, whose value depends on if the underlying asset price crosses a given barrier. We distinguish between knock-in options, that have value zero unless the underlying asset price hits the barrier, and knock-out options, that assume value zero if the underlying asset price crosses the barrier. For example, an up-and-out option looses all its value if the asset price rises above a given barrier, and vice-versa a down-and-out option has value zero if the asset price falls below the barrier. Usually the actual payoff at expiration of a barrier option is simply a put or a call option, as we will see below. There also exist more complex barrier options, that for example require the asset not only to cross a barrier but to spend a certain amount of time across the barrier in order to knock in or out.

We will now price and hedge an up-and-out call as an example of an exotic derivative in the context of the Black-Scholes market. Under the equivalent martingale measure \( Q \), we have

\[
X^1_t = x_0e^{(r-\frac{\sigma^2}{2})t+\sigma \hat{B}_t}, \quad t \in [0,T],
\]

with \( X^1_0 = x_0 \), and we assume \( 0 < x_0 < B \). If we now set

\[
\hat{B}_t = \tilde{B}_t + \frac{1}{\sigma} \left( r - \frac{\sigma^2}{2} \right) t,
\]

we have

\[
X^1_t = x_0e^{\sigma \hat{B}_t}, \quad t \in [0,T].
\]

With this notation the payoff of an up-and-out call can be expressed as

\[
H = \begin{cases} 
(X^1_T - K)^+, & x_0e^{\sigma \hat{M}_T} \leq B, \\
0, & x_0e^{\sigma \hat{M}_T} > B.
\end{cases}
\]

To find the replicating portfolio value we rewrite this as

\[
H = (X^1_T - K)^+ \mathbb{1}_{\{x_0e^{\sigma \hat{M}_T} \leq B\}}
\]

\[
= (X^1_T - K) \mathbb{1}_{\{X^1_T \geq K, \quad x_0e^{\sigma \hat{M}_T} \leq B\}}
\]

\[
= (X^1_T - K) \mathbb{1}_{\{x_0e^{\sigma \hat{B}_T} \geq K, \quad x_0e^{\sigma \hat{M}_T} \leq B\}}
\]

\[
= (X^1_T - K) \mathbb{1}_{\left\{B_T \geq \frac{1}{\sigma} \log \left( \frac{K}{x_0} \right), \quad \hat{M}_T \leq \frac{1}{\sigma} \log \left( \frac{B}{x_0} \right) \right\}}
\]

\[
= (X^1_T - K) \mathbb{1}_{\{B_T \geq k, \quad \hat{M}_T \leq b\}}.
\]
and now compute the price at time $t = 0$. We have

$$V_0 = E^Q \left[ \frac{H}{X_T} \right] = E^Q \left[ \frac{(X_T^1 - K)}{e^{rT}} 1_{\{\hat{B}_T \geq k, \hat{M}_T \leq b\}} \right]$$

$$= \int_A (x_0 e^{\sigma \omega} - K) e^{rT} f(m, \omega) dm d\omega,$$

with $A = \{(m, \omega) : m \leq b, \omega \geq k\}$ and $f(m, \omega)$ is the joint density of $(\hat{M}_T, \hat{B}_T)$ given by Proposition 5.4.2. To compute this integral we distinguish between $k < 0$ or $k \geq 0$. Note that $k < 0$ if $K < x_0$, i.e. if the call is in the money at time $t = 0$. If we set $\alpha = \frac{1}{\sigma} \left( r - \frac{\sigma^2}{2} \right)$, we get

$$V_0 = \int_k^b \int_0^\omega e^{-\alpha T} (x_0 e^{\sigma \omega} - K) \frac{2(2m - \omega)}{T \sqrt{2\pi T}} e^{\alpha \omega - \frac{1}{2} \alpha^2 T - \frac{(2m - \omega)^2}{2T}} dm d\omega$$

$$= \int_k^b e^{-\alpha T} (x_0 e^{\sigma \omega} - K) e^{\alpha \omega - \frac{1}{2} \alpha^2 T} \left( \int_0^\omega \frac{2(2m - \omega)}{T} e^{-\frac{(2m - \omega)^2}{2T}} dm \right) d\omega$$

Recall: $\int_y^\infty e^{-\frac{x^2}{2T}} dz = \left[ e^{-\frac{z^2}{2T}} \right]_y^\infty$

$$= \int_k^b e^{-\alpha T} (x_0 e^{\sigma \omega} - K) e^{\alpha \omega - \frac{1}{2} \alpha^2 T - \frac{2\omega^2}{2T}} d\omega$$

$$= \frac{e^{-\alpha T}}{\sqrt{2\pi T}} \int_k^b (x_0 e^{\sigma \omega} - K) e^{\alpha \omega - \frac{1}{2} \alpha^2 T - \frac{2\omega^2}{2T}} d\omega$$

$$= \frac{e^{-\alpha T}}{\sqrt{2\pi T}} \int_k^b \left( \int_0^\omega \exp \left( (\sigma + \alpha) \omega - \frac{1}{2} \alpha^2 T - \frac{\omega^2}{2T} \right) d\omega \right)$$

$$= \int_k^b \left( \int_0^\omega \exp \left( (\sigma + \alpha) \omega - \frac{1}{2} \alpha^2 T - \frac{2b - \omega^2}{2T} \right) d\omega \right)$$

$$= \int_k^b e^{\alpha \omega - \frac{1}{2} \alpha^2 T - \frac{2b}{2T}} d\omega - \int_k^b e^{\alpha \omega - \frac{1}{2} \alpha^2 T - \frac{(2b - \omega)^2}{2T}} d\omega. \quad (5.8)$$

Each of these integrals is of the form

$$\int_k^b e^{\beta + \gamma \omega - \frac{1}{2} \alpha^2 T} d\omega = e^{\frac{1}{2} \gamma^2 T + \beta} \int_k^b e^{-\frac{1}{2} \alpha^2 T} e^{-\frac{\gamma^2}{2} T} dy$$

$$= e^{\frac{1}{2} \gamma^2 T + \beta} \left( \mathcal{N} \left( \frac{1}{\sqrt{T}} (b - \gamma T) \right) - \mathcal{N} \left( \frac{1}{\sqrt{T}} (k - \gamma T) \right) \right), \quad (5.9)$$
where $N(x)$ is as usual the standard normal distribution function. We now define

$$\delta_\pm(\tau, s) = \frac{1}{\sigma \sqrt{\tau}} \left[ \log s + \left( r \pm \frac{\sigma^2}{2} \right) \right],$$

and compute the terms $A_1$ and $A_2$ appearing in (5.8), thereby using (5.9). We start with $A_1$. We have that

$$\int b^k \exp \left\{ (\sigma + \alpha) \omega - \frac{\alpha^2}{2} T - \frac{\omega^2}{2T} \right\} d\omega =$$

$$= \exp \left\{ -\frac{\alpha^2 T}{2} + \frac{(\sigma + \alpha)^2}{2} T \right\} \cdot \left( N \left( \frac{b - (\alpha + \sigma)T}{\sqrt{T}} \right) - N \left( \frac{k - (\alpha + \sigma)T}{\sqrt{T}} \right) \right),$$

with

$$\exp \left\{ -\frac{T}{2\sigma^2} \left( r - \frac{\sigma^2}{2} \right)^2 \right\} = \exp \left\{ \frac{T}{2\sigma^2} \right\} = \exp(rT)$$

and

$$N \left( \frac{b - (\alpha + \sigma)T}{\sqrt{T}} \right) - N \left( \frac{k - (\alpha + \sigma)T}{\sqrt{T}} \right) =$$

$$= N \left( \frac{1}{T} \log \frac{B}{x_0} - \frac{1}{2}(r + \frac{\sigma^2}{2}) T \right) - N \left( \frac{1}{T} \log \frac{K}{x_0} - \frac{1}{2}(r + \frac{\sigma^2}{2}) T \right)$$

$$= N \left( \delta_+ \left( T, \frac{x_0}{B} \right) - N \left( \delta_+ \left( T, \frac{x_0}{K} \right) \right) \right) = 1 - N \left( \delta_+ \left( T, \frac{x_0}{B} \right) \right) = N \left( \delta_+ \left( T, \frac{x_0}{K} \right) \right).$$

Furthermore,

$$\int b^k \exp \left\{ (\sigma + \alpha) \omega - \frac{\alpha^2}{2} T - \frac{2b - \omega^2}{2T} \right\} d\omega =$$

$$= \int b^k \exp \left\{ (\sigma + \alpha + \frac{2b}{T}) \omega - \left( \frac{\alpha^2}{2} T + \frac{2b^2}{T} \right) - \frac{\omega^2}{2T} \right\} d\omega =$$

$$= \exp \left\{ -\left( \frac{\alpha^2 T}{2} + \frac{2b^2}{T} \right) + \left( \sigma + \alpha + \frac{2b}{T} \right)^2 \right\} \cdot \left( N \left( \frac{b - (\alpha + \sigma + 2b/T)T}{\sqrt{T}} \right) - N \left( \frac{k - (\alpha + \sigma + 2b/T)T}{\sqrt{T}} \right) \right),$$

with
\[
\exp \left\{ -\left(\frac{\alpha^2 T}{2} + \frac{2b^2}{T}\right) + \left(\sigma + \alpha + \frac{2b}{T}\right)^2 \frac{T}{2}\right\} = \\
= \exp \left\{ \frac{\sigma^2 T}{2} + \alpha \sigma T + 2ab + 2\sigma b\right\} \\
= \exp \left\{ \frac{\sigma^2 T}{2} + \left( r - \frac{\sigma^2}{2}\right) T + \frac{2}{\sigma} \left( r - \frac{\sigma^2}{2}\right) b + 2\sigma b\right\} \\
= \exp \left\{ rT + \frac{2rb}{\sigma} + \sigma b\right\} \\
= \exp \{rT\} \exp \left\{ \frac{2r}{\sigma} \log \left( \frac{B}{x_0}\right) + \log \left( \frac{B}{x_0}\right)\right\} \\
= \exp \{rT\} \left( \frac{x_0}{B}\right)^{-\frac{2r}{\sigma^2}-1}
\]

and
\[
\mathcal{N} \left( \frac{b - (\alpha + \sigma + 2b/2T)}{\sqrt{T}} \right) - \mathcal{N} \left( \frac{k - (\alpha + \sigma + 2b/2T)}{\sqrt{T}} \right) = \\
\mathcal{N} \left( \frac{-T(\alpha + \sigma) - b}{\sqrt{T}} \right) - \mathcal{N} \left( \frac{k - 2b - (\alpha + \sigma)T}{\sqrt{T}} \right) \\
= \mathcal{N} \left( -\delta_+ \left( T, \frac{B}{x_0} \right) \right) - \mathcal{N} \left( \frac{\frac{1}{\sigma} \log \left( \frac{k}{x_0}\right) \frac{x_0^2}{B^2} - (\alpha + \sigma)T}{\sqrt{T}} \right) \\
= 1 - \mathcal{N} \left( \delta_+ \left( T, \frac{B}{x_0} \right) \right) - \mathcal{N} \left( \frac{-\frac{1}{\sigma} \log \left( \frac{k}{x_0}\right) \frac{x_0^2}{B^2} - (\alpha + \sigma)T}{\sqrt{T}} \right) \\
= 1 - \mathcal{N} \left( \delta_+ \left( T, \frac{B}{x_0} \right) \right) - 1 + \mathcal{N} \left( \delta_+ \left( T, \frac{B^2}{x_0 K} \right) \right) \\
= \mathcal{N} \left( \delta_+ \left( T, \frac{B^2}{x_0 K} \right) \right) - \mathcal{N} \left( \delta_+ \left( T, \frac{B}{x_0} \right) \right). 
\]

Putting everything together we obtain that
\[
A_1 = e^{-rT} x_0 \left\{ e^{rT} \left[ \mathcal{N} \left( \delta_+ \left( T, \frac{x_0}{K} \right) \right) - \mathcal{N} \left( \delta_+ \left( T, \frac{x_0}{B} \right) \right) \right] \right\} \\
- e^{rT} \left( \frac{x_0}{B}\right)^{-\frac{2r}{\sigma^2}-1} \left[ \mathcal{N} \left( \delta_+ \left( T, \frac{B^2}{x_0 K} \right) \right) - \mathcal{N} \left( \delta_+ \left( T, \frac{B}{x_0} \right) \right) \right] \} \\
= x_0 \left[ \mathcal{N} \left( \delta_+ \left( T, \frac{x_0}{K} \right) \right) - \mathcal{N} \left( \delta_+ \left( T, \frac{x_0}{B} \right) \right) \right] \\
- B \left( \frac{x_0}{B}\right)^{-\frac{2r}{\sigma^2}} \left[ \mathcal{N} \left( \delta_+ \left( T, \frac{B^2}{x_0 K} \right) \right) - \mathcal{N} \left( \delta_+ \left( T, \frac{B}{x_0} \right) \right) \right].
\]
For $A_2$ one has to perform similar computations. Finally we obtain that the price at time 0 of the up-and-out call is equal to

$$V_0 = x_0 \left[ N \left( \delta_+ \left( T, \frac{x_0}{K} \right) \right) - N \left( \delta_+ \left( T, \frac{x_0}{B} \right) \right) \right]$$

$$- e^{-rT} K \left[ N \left( \delta_- \left( T, \frac{x_0}{K} \right) \right) - N \left( \delta_- \left( T, \frac{x_0}{B} \right) \right) \right]$$

$$- B \left( \frac{x_0}{B} \right)^{-\frac{2x}{\sigma^2}} \left[ N \left( \delta_+ \left( T, \frac{B^2}{x_0 K} \right) \right) - N \left( \delta_+ \left( T, \frac{B}{x_0} \right) \right) \right]$$

$$+ e^{-rT} K \left( \frac{x_0}{B} \right)^{-\frac{2x}{\sigma^2} + 1} \left[ N \left( \delta_- \left( T, \frac{B^2}{x_0 K} \right) \right) - N \left( \delta_- \left( T, \frac{B}{x_0} \right) \right) \right].$$

The value of $H$ at time $t > 0$ is given by the following theorem.

**Theorem 5.4.4.** Let $v(t, x)$ denote the price at time $t$ of the up-and-out call under the assumption that the call has not been knocked out prior to time $t$ and that $X^1_t = x$. Then $v(t, x)$ satisfies the Black-Scholes-Merton PDE

$$\frac{\partial v}{\partial t}(t, x) + r x \frac{\partial v}{\partial x}(t, x) + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 v}{\partial x^2}(t, x) = rv(t, x) \quad (5.10)$$

on the rectangle $\{(t, x) : 0 \leq t < T, 0 \leq x \leq B\}$, with boundary conditions

$$v(t, 0) = 0, \quad 0 \leq t \leq T; \quad (5.11)$$

$$v(t, B) = 0, \quad 0 \leq t < T; \quad (5.12)$$

$$v(T, x) = (x - K)^+, \quad 0 \leq x \leq B. \quad (5.13)$$

In particular note that $v(t, x)$ is not continuous in $(t, x) = (T, B)$.

**Remark 5.4.5.** Some comments about conditions (5.11), (5.12) and (5.13):

1. About (5.11): if the asset price begins at zero, it stays there and the option expires out of the money.
2. About (5.12): as soon as the asset price reaches the barrier $B$ from below, it immediately rises above it, i.e. it oscillates falling and rising across the level $B$ infinitely many times after hitting it. This is a consequence of the fact that the price process has quadratic variation not zero. Hence the price of the option is zero when the asset price hits $B$ because it will be knocked out immediately afterwards anyway.
3. The only exception is when the asset price reaches $B$ at time of expiration $T$, since there is no time left for the knock-out. This is expressed in (5.13).

Here we cannot directly apply the results of the Feynman-Kac Theorem because the price at time $t$ of the option is

$$V_t = e^{-r(T-t)} E^Q \left[ H | \mathcal{F}_t^B \right] \quad (5.14)$$
while \( v(t, X^1_t) \) is the price of the option under the assumption that the knock out has not happened prior to \( t \). In particular if the asset price rises above the barrier and returns below it by time \( t \), then \( V_t \) will be zero because the option has knocked out, but \( v(t, x) \) is strictly positive for all \( 0 \leq t < T \) and \( 0 < x < B \). The replicating portfolio value is path-dependent and remembers that the option has knocked out, while \( v(t, X^1_t) \) gives the price of the option under the assumption that the knock out has not happened before \( t \), even if this assumption is incorrect. To solve this problem we can introduce the stopping time

\[
\rho = \inf \{ t \mid X^1_t \geq B \}.
\]

If the asset price does not reach \( B \) before expiration, we put \( \rho = +\infty \). Hence we have

**Lemma 5.4.6.** Let \( V_t \) be given by (5.14). Then

\[
V_t = v(t, X^1_t)1_{\{0 \leq t \leq \rho\}}, \quad t \in [0, T],
\]

and the stopped process

\[
e^{-r(t \wedge \rho)} v(t \wedge \rho, X^1_{t \wedge \rho}), \quad t \in [0, T]
\]

is a \( Q \)-martingale.

**Proof.** (5.15) holds because \( v(t, X^1_t) \) is the price of the up-and-out call before knock-out, i.e. for every \( t \leq \rho \). The stopped process in (5.16) coincides with the stopped martingale \((V^\rho_t)_{t \in [0, T]}\), hence it is also a martingale. \( \square \)

We are now able to prove Theorem 5.4.4.

**Proof (Theorem 5.4.4.).** Since by Lemma 5.4.6 the portfolio value process \( V_t \) coincides with \( v(t, X^1_t) \) up to time \( \rho \), by Theorem 4.5.2 we obtain that for \( 0 \leq t \leq \rho \) (i.e. before the knock-out)

\[
v(t, X^1_t) = V_t = X^0_t E^Q \left[ \frac{H}{X^2_T} \right]_{\rho^B_t}
\]

must satisfy equation (5.10) with terminal conditions (5.11), (5.12) and (5.13) for \( 0 \leq t < T, 0 \leq x \leq B \). \( \square \)

If an agent begins with a short position in the up-and-out call and with initial capital \( v_0 = v(0, X^1_0) \), then with

\[
\theta^1_t = \frac{\partial v}{\partial x}(t, X^1_t)
\]

he can replicate the option value \( v(t, X^1_t) \) up to the time \( \rho \) of knock-out, or up to the expiration time \( T \), whichever occurs first. However, in practice this method does not provide a good hedging strategy, since we have just seen
that \( v(t, x) \) is discontinuous in \((T, B)\), e.g. for \( t \) near \( T \) and \( x \) just below \( B \) \( v(t, x) \) has both large negative delta and gamma. This means that the agent will have to assume a large short position in the risky asset and at the same time he will be forced to make large adjustments in the position whenever the asset price moves. Therefore in practical applications it is often common to price and hedge the up-and-out call assuming the barrier were at a level slightly higher than the actual level \( B \).
6

Forward and Future Contracts

6.1 Forward Contracts

As usual assume we are given an arbitrage-free financial market as introduced in Section 4.1. In this section we will show how to derive the price of a forward contract. The forward contract obliges the holder to purchase an asset (e.g. a commodity) at a non random delivery price $q$ at specified delivery date $T$. We represent the value of this asset with an $\mathcal{F}_T^B$-measurable random variable $X \geq 0$ and $X \in L^1(\mathbb{Q})$. At time $T$ the value of this contract is given by

$$H = X - q,$$

and by Proposition 4.4.5 the value process of the replicating strategy is

$$V_t(q) = X_t^0 E^Q \left[ \frac{X - q}{X_T^0} \bigg| \mathcal{F}_t^B \right], \quad t \in [0, T]. \quad (6.1)$$

Note that since $r$ is bounded, we have no integrability problems regarding $\frac{1}{X_T^0}$. If we wish to relax this assumption on $r$, then (6.1) holds if $\frac{X}{X_T^0}, \frac{1}{X_T^0} \in L^1(\mathbb{Q})$. If the forward contract is written on one of the $n$ risky assets, say on $X_1^t$, then (6.1) becomes

$$V_t(q) = X_t^1 E^Q \left[ \frac{X_T^1 - q}{X_T^0} \bigg| \mathcal{F}_t^B \right]$$

$$= X_t^1 - q E^Q \left[ \frac{1}{X_T^0} \bigg| \mathcal{F}_t^B \right],$$

i.e.

$$V_t(q) = X_t^1 - q X_t^0 E^Q \left[ \frac{1}{X_T^0} \bigg| \mathcal{F}_t^B \right]. \quad (6.2)$$

To hedge himself against the forward contract, the seller buys one unit of the risky asset and takes a loan equal to $q E^Q \left[ \frac{1}{X_T^0} \bigg| \mathcal{F}_t^B \right]$ with the bank. At time $T$
he delivers the asset, receives the payment $q$ and pays back the debt with the
bank that is then equal to $q$. We now introduce the concept of forward price.

**Definition 6.1.1.** Let $X$ be an $\mathcal{F}_B^T$-measurable random variable such that $X \in L^1(Q)$. The forward price $(f_t)_{t \in [0,T]}$ for $X$ at time $t$ is given by

$$f_t = \frac{E^Q\left[\frac{X}{X^T}\mid \mathcal{F}_t\right]}{E^Q\left[\frac{1}{X^T}\mid \mathcal{F}_t\right]}.$$

We see immediately that if we substitute $f_t$ for $q$ in (6.2), then we obtain

$$V_t(f_t) = 0,$$

i.e. the value of the forward contract (at time $t$) to buy $X$ at time $T$ for the
price $f_t$ is zero. This means that the forward price is determined in such a
way that the value of the forward contract is equal to zero at the time $t$ when
the contract is made. This models the fact that the forward price should be
a fair price both for the holder, as well as for the seller of the contract at the
time the contract is set up. We now compute the forward price in the case
that $X = X^1$.

$$f_t = \frac{E^Q\left[X^1_T\mid \mathcal{F}_t\right]}{E^Q\left[\frac{1}{X^T}\mid \mathcal{F}_t\right]} = \frac{X^1_t}{E^Q\left[\frac{1}{X^T}\mid \mathcal{F}_t\right]}.$$

If $X^0_T$ is deterministic, we obtain

$$f_t = \frac{X^1_t}{X^0_t} \cdot X^0_T = e^{\int_t^T r_s \, ds} X^1_t.$$

### 6.2 Put-Call Parity

This section is dedicated to the well-known put-call parity, that establishes the
relation between the price of a European put and a European call option in
an arbitrage-free market. Let $P = (K - X^1_T)^+$ be a put option on $X^1$ with
strike price $K$. Then we have

$$C - P = (X^1_T - K)^+ - (K - X^1_T)^+ = X^1_T - K,$$

or equivalently

$$P = C - (X^1_T - K), \quad (6.3)$$

which implies that having a long position in a put is equivalent to having 1) a
long position in a call and 2) a short position in a forward contract. Hence
the price and the hedging strategy for a put option can be derived by (6.3).
6.3 Futures Contracts

In this section we derive the arbitrage-free price of a futures contract. First we introduce the following definition following Björk [1].

**Definition 6.3.1.** Let \( X \) be a financial asset, e.g. a contingent claim or a commodity. A futures contract on \( X \) with delivery time \( T \) has the following characteristics.

i) For every \( t \in [0, T] \) there exists a quoted futures price or futures quotation \( F(t, T) = F(t, T, X) \) for \( X \) on the market for delivery at time \( T \).

ii) At the time of delivery the holder of the contract receives the claim \( X \) and in return pays \( F(T, T) \) to the seller of the contract.

iii) For any arbitrary time interval \((s, t]\), where \( 0 \leq s < t \leq T \), the holder of the contract at time \( t \) receives the amount \( F(t, T) - F(s, T) \).

iv) The price \( \pi_t \) of the futures contract at any time \( t < T \) prior to delivery is equal to zero.

The payment schedule described above is known as *marking to market* and is organized in such a way that the holder of a futures position (either short or long) must keep a certain amount of money with the broker as a safety margin against default.

In the following we will assume that the futures price \( F(t, T) \) follows an Itô process.

**Proposition 6.3.2.** Let \( X \) be an \( \mathcal{F}_t^B \)-measurable random variable such that \( X \geq 0 \) and \( X \in L^2(\mathbb{Q}) \). Then the futures price process is given by

\[
F(t, T) = E^Q \left[ X \mid \mathcal{F}_t^B \right].
\]  

(6.4)

*Proof.* The formula in (6.4) is often simply stated as a definition, however, following Filipović [6] we will provide the following heuristic justification.

By the above definition, the holder of a futures contract receives a continuum of cash-flows during the interval \((t, T] \). If \( t = t_0 < t_1 < \ldots < t_n = T \) is a partition of \((t, T]\), the present value of the cash-flows \( F(t_i, T) - F(t_{i-1}, T) \) at \( t_i, i = 1, \ldots, n \) is given by

\[
E^Q \left[ \sum_{i=1}^{n} \frac{1}{X_{t_i}^0} (F(t_i, T) - F(t_{i-1}, T)) \right],
\]

which by Definition 6.3.1 iv) is assumed to be equal to zero. We now write

\[
\sum_{i=1}^{n} \frac{1}{X_{t_i}^0} (F(t_i, T) - F(t_{i-1}, T)) = \sum_{i=1}^{n} \left( \frac{1}{X_{t_i}^0} - \frac{1}{X_{t_{i-1}}^0} + \frac{1}{X_{t_{i-1}}^0} \right) (F(t_i, T) - F(t_{i-1}, T)) =
\]
\[ \sum_{i=1}^{n} \frac{1}{X_{t_{i-1}}} (F(t_i, T) - F(t_{i-1}, T)) + \sum_{i=1}^{n} \left( \frac{1}{X_{t_i}} - \frac{1}{X_{t_{i-1}}} \right) (F(t_i, T) - F(t_{i-1}, T)), \]

and by Proposition 2.6.2 and Corollary 2.6.5, if the mesh size of the partition tends to zero, then this converges to

\[ \int_{t}^{T} \frac{1}{X_{s}} dF(s, T) + \int_{t}^{T} d\langle \frac{1}{X_{s}}, F \rangle_s = \int_{t}^{T} \frac{1}{X_{s}} dF(s, T), \]

since \( \frac{1}{X_{s}} \) is a continuous finite-variation process. If we suppose that appropriate integrability assumptions hold (e.g. uniform integrability), then we have

\[ E^Q \left[ \int_{t}^{T} \frac{1}{X_{s}} dF(s, T) \bigg| F_B \right] = 0, \]

hence

\[ M_t = \int_{0}^{t} \frac{1}{X_{s}} dF(s, T) = E^Q \left[ \int_{0}^{T} \frac{1}{X_{s}} dF(s, T) \bigg| F_B \right], \quad t \in [0, T] \]

is a \( Q \)-martingale. By Theorem 2.9.2 there exists \( \tilde{\Phi} \in \mathcal{A}_{\text{loc, d}}^2(T) \) with

\[ M_t = M_0 + \int_{0}^{t} \tilde{\Phi}_s \cdot d\tilde{B}_s \]

and \( dF(t, T) = X_{t}^0 dM_t = X_{t}^0 \tilde{\Phi}_t \cdot d\tilde{B}_t \). Since \( X \in L^2(Q) \) we have

\[ E[F(T, T)^2] = E[X^2] < \infty, \]

and because \( (F(t, T))_{t \in [0, T]} \) is a local martingale we obtain

\[ E[F(T, T)^2] = E[F(\cdot, T)_{\cdot \cdot}] = E \left[ \int_{0}^{T} (X_{t}^0 \| \tilde{\Phi}_t \|)^2 dt \right], \]

which implies that \( X_{t}^0 \| \tilde{\Phi}_t \| \in \mathcal{A}_{d}^2(T) \), and consequently

\[ F(t, T) = F(0, T) + \int_{0}^{t} X_{t}^0 dM_t, \quad t \in [0, T] \]

is a true \( Q \)-martingale. \( \square \)

We refer to Hull [7] for further references concerning futures contracts.
A Probability Essentials

The following theorem is very useful, e.g. when one is interested in a sequence of independent events, or in general when considering the probability of the limes superior of sets.

**Theorem A.0.1 (Borel-Cantelli lemma).** Let \((A_n)_{n \in \mathbb{N}} \in \mathcal{F}\) be a sequence of events in \((\Omega, \mathcal{F}, P)\).

i) If \(\sum_{n=1}^{\infty} P(A_n) < \infty\), then \(P\left(\bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} A_k\right) = 0\).

ii) If \(\sum_{n=1}^{\infty} P(A_n) = \infty\) and \(A_n\) are mutually independent events, then \(P\left(\bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} A_k\right) = 1\).

**Proof.** For the proof we refer to Jacod and Protter [8], pp. 71–72.

**Theorem A.0.2 (Kolmogorov’s zero-one law).** Let \((X_n)_{n \in \mathbb{N}}\) be a sequence of independent random variables, all defined on \((\Omega, \mathcal{F}, P)\), and let \(\mathcal{C}_\infty\) be the corresponding tail \(\sigma\)-algebra, i.e. \(\mathcal{C}_\infty = \bigcap_{n=1}^{\infty} \sigma(\bigcup_{p \geq n} \sigma(X_p))\).

If \(C \in \mathcal{C}_\infty\), then \(P(C) = 0\) or \(P(C) = 1\).

**Proof.** For the proof we refer to Jacod and Protter [8], pp. 72.

We now review some fundamental concepts and results on convergence of sequences of random variables.

**Definition A.0.3 (convergence of random variables).** Let \((X_n)_{n \in \mathbb{N}}\) be a sequence of \(\mathbb{R}^n\)-valued random variables on a probability space \((\Omega, \mathcal{F}, P)\). We denote by \(\|x\|\) the Euclidean norm of \(x \in \mathbb{R}^n\).

i) We say that \((X_n)_{n \in \mathbb{N}}\) converges almost surely to the random variable \(X\) (or in short: \(X_n \xrightarrow{n \to \infty} X\) a.s., \(X_n \xrightarrow{a.s.} X\)), if

\[
P\left(\omega \in \Omega : \lim_{n \to \infty} X_n(\omega) = X(\omega)\right) = 1.
\]
ii) We say that \((X_n)_{n \in \mathbb{N}}\) converges in probability to \(X\) (or in short: \(X_n \xrightarrow{P} X\)), if for any \(\varepsilon > 0\) we have
\[
\lim_{n \to \infty} \mathbb{P} (\omega \in \Omega : \|X_n(\omega) - X(\omega)\| > \varepsilon) = 0,
\]

iii) For \(1 \leq p < \infty\) we say that \((X_n)_{n \in \mathbb{N}}\) converges to \(X\) in \(L^p\) (or in short: \(X_n \xrightarrow{L^p} X\)), if \(\|X_n\|, \|X\|\) are in \(L^p(\Omega, \mathcal{F}, \mathbb{P})\) and
\[
\lim_{n \to \infty} E[\|X_n - X\|^p] = 0.
\]

iv) We say that \((X_n)_{n \in \mathbb{N}}\) converges in distribution or weakly to \(X\) (or in short: \(X_n \xrightarrow{D} X\), \(X_n \xrightarrow{\mathcal{L}} X\)), if the distribution functions \(F_n\) of \(X_n\) converge to the distribution function \(F\) of \(X\) for all continuity points of \(F\):
\[
\lim_{n \to \infty} F_n(x) = F(x) \text{ for all continuity points } x.
\]

We state some relationships between the different types of convergence.

**Theorem A.0.4.** Let \((X_n)_{n \in \mathbb{N}}\) be a sequence of \(\mathbb{R}^n\)-valued random variables on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\).

i) \(X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X\),

ii) \(X_n \xrightarrow{L^p} X \Rightarrow X_n \xrightarrow{P} X\),

iii) \(X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{\mathcal{L}} X\),

iv) If \(X_n \xrightarrow{P} X\), then there exists a subsequence \((n_k)_{k \in \mathbb{N}}\) such that
\[
\lim_{k \to \infty} X_{n_k} = X \text{ a.s.}
\]
v) If \(X_n \xrightarrow{\mathcal{L}} X\), with \(\mathbb{P}(X = c) = 1\), \(c \in \mathbb{R}^n\), then \(X_n \xrightarrow{P} X\).

**Proof.** For the proof we refer to Jacod and Protter [8], pp. 141–163.

**Proposition A.0.5.** Given a sequence \((Y_n)_{n \in \mathbb{N}}\) of real-valued random variables on \((\Omega, \mathcal{F}, \mathbb{P})\), if \(Y_n \xrightarrow{n \to \infty} Y\) in probability and \(\sup_{n \in \mathbb{N}} E[|Y_n|^q] < +\infty\) for some \(1 \leq p < q\), then \(Y_n \xrightarrow{n \to \infty} Y\) also in \(L^p\).

**Proof.** By Fatou’s lemma
\[
E[|Y|^q] = E[\lim \inf_{n \to \infty} |Y_n|^q] \leq \lim \inf_{n \to \infty} E[|Y_n|^q] \leq \sup_{n \in \mathbb{N}} E[|Y_n|^q] < \infty.
\]

It then follows that \(\sup_{n \in \mathbb{N}} E[|Y_n - Y|^q] < \infty\) by the triangle inequality (Minkowski’s inequality). Furthermore by using Hölder’s inequality for \(\frac{q}{p}\) and \(\frac{q}{q-p}\), for every \(\varepsilon > 0\) we have
\[
E[|Y_n - Y|^p] \leq \varepsilon^p + E\{|Y_n - Y|^p 1_{\{|Y_n - Y| > \varepsilon\}}\}
\]
\[
\left(1 \leq p < q \Rightarrow \frac{q}{p} > 1 \quad \frac{1}{p} + \frac{1}{q} = 1\right)
\]
\[ \leq \varepsilon^p + \left( E(|Y_n - Y|^q) \right)^{\frac{1}{q}} \mathbb{P}(|Y_n - Y| > \varepsilon)^{\frac{1}{q}}. \]

We finally conclude that
\[ \lim_{n \to \infty} E[|Y_n - Y|^p] = 0. \]

Many problems in mathematics can be solved by transformations into other spaces. In stochastics, one of the most important transforms is the Fourier transform or characteristic function of a random variable.

**Definition A.0.6 (characteristic function).** The characteristic function \( \varphi_X : \mathbb{R}^n \to \mathbb{C} \) of an \( \mathbb{R}^n \)-valued random variable \( X \) on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) is defined as
\[
\varphi_X(t) = E\left[e^{i \langle t, X \rangle}\right] = \int_{\Omega} e^{i \langle t, X \rangle} d\mathbb{P} = \int_{\mathbb{R}^n} e^{i \langle t, x \rangle} d\mathbb{P}_X(dx),
\]
where \( \mathbb{P}_X \) is the distribution measure of \( X \) and \( \langle x, y \rangle \) denotes the scalar product of \( x, y \in \mathbb{R}^n \).

Note that characteristic functions always exist, because \( |e^{i \langle t, X \rangle}| = 1 \). We state some useful properties of the characteristic function.

**Theorem A.0.7.**

i) The characteristic function \( \varphi_X \) of an \( \mathbb{R}^n \)-valued random variable \( X \) is a continuous function with \( \varphi_X(0) = 1 \) and \( |\varphi_X(t)| \leq \varphi_X(0) = 1 \) for all \( t \in \mathbb{R}^n \).

ii) For \( A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m \): \( \varphi_{AX+b}(t) = e^{i \langle t, b \rangle} \varphi_X(A^T t) \).

iii) Uniqueness: If two random variables have the same characteristic functions, then they have the same distribution function.

iv) Let \( X \) be an \( \mathbb{R}^n \)-valued random variable and suppose \( E[|X|^m] < \infty \) for some integer \( m \in \mathbb{N} \). Then the characteristic function \( \varphi_X \) of \( X \) has continuous partial derivatives up to order \( m \), and
\[
\frac{\partial^m}{\partial x_{j_1} \cdots \partial x_{j_m}} \varphi_X(t) = i^m E[X_{j_1} \cdots X_{j_m} e^{i \langle t, X \rangle}].
\]

This property is often used to calculate the moments of a random variable.

v) Characteristic functions and independence: Let \( X = (X_1, \ldots, X_n) \) be an \( \mathbb{R}^n \)-valued random variable. Then the real-valued random variables \( X_j, j = 1, \ldots, n \), are independent if and only if
\[
\varphi_X(t_1, \ldots, t_n) = \prod_{j=1}^n \varphi_{X_j}(t_j)
\]
for all \( t = (t_1, \ldots, t_n) \in \mathbb{R}^n \).
vi) The Lévy-Cramér continuity theorem: Let \((X_n)_{n \in \mathbb{N}}\) be a sequence of \(\mathbb{R}^n\)-valued random variables on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\). Then

\[ X_n \xrightarrow{L} X \iff \varphi_{X_n}(t) \xrightarrow{n \to \infty} \varphi_X(t) \text{ for all } t \in \mathbb{R}^n. \]

**Proof.** For the proof we refer to Jacod and Protter [8], pp. 103–113 and 167–170.
Conditional Expectation

Consider a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and a sub-\(\sigma\)-algebra \(\mathcal{B} \subset \mathcal{F}\). Given \(X \in L^1(\Omega, \mathcal{F}, \mathbb{P})\), we wish to evaluate \(X\) knowing only the “smaller” information \(\mathcal{B}\). We will need the following theorem.

**Theorem B.0.1 (Radon-Nikodym).** Consider the measurable space \((\Omega, \mathcal{F}, \nu)\), where \(\nu\) is \(\sigma\)-finite. Let \(\mu\) be a signed measure on \((\Omega, \mathcal{F})\) such that \(\mu \ll \nu\). Then there exists a unique \(\Phi \in L^1(\Omega, \mathcal{F}, \nu)\) such that for all \(A \in \mathcal{F}\)

\[
\mu(A) = \int_A \Phi \, d\nu.
\]

If \(\mu\) is also \(\sigma\)-finite, then \(\Phi \geq 0\).

*Proof.* For the proof we refer to Jacod and Protter \[8\].

The next theorem shows the existence of the conditional expectation (assuming \(X \in L^1(\Omega, \mathcal{F}, \mathbb{P})\)).

**Theorem B.0.2.** Let \(X \in L^1(\Omega, \mathcal{F}, \mathbb{P})\) and \(\mathcal{B} \subset \mathcal{F}\) be a sub-\(\sigma\)-algebra. Then there exists a unique \(Y \in L^1(\Omega, \mathcal{B}, \mathbb{P})\) such that

\[
\int_B X \, d\mathbb{P} = \int_B Y \, d\mathbb{P} \text{ for all } B \in \mathcal{B}.
\]

We call \(Y\) the conditional expectation of \(X\) and denote it by \(E[X|\mathcal{B}]\).

*Proof.* Consider the signed measure on \((\Omega, \mathcal{B}, \mathbb{P})\) defined as

\[
\mu_X(B) := \int_B X \, d\mathbb{P}_B
\]

* A measure \(\nu\) on \((\Omega, \mathcal{F})\) is \(\sigma\)-finite if \(\exists (A_n)_{n \in \mathbb{N}} \subseteq \mathcal{F}\) such that \(\nu(A_n) < \infty\) for all \(n \in \mathbb{N}\) and \(\Omega = \bigcup_{n \in \mathbb{N}} A_n\).
for all \( B \in \mathcal{B} \), where \( \mathbb{P}_B := \mathbb{P}|_B \). Then \( \mu_X \ll \mathbb{P}_B \) and by Theorem B.0.1 there exists \( Y = \frac{d\mu_X}{d\mathbb{P}_B} \in L^1(\Omega, \mathcal{B}, \mathbb{P}) \) such that

\[
\mu_X(B) = \int_B Y \, d\mathbb{P}_B.
\]

Hence

\[
\mu_X(B) = \int_B X \, d\mathbb{P} = \int_B Y \, d\mathbb{P}_B = \int_B Y \, d\mathbb{P},
\]

and \( E[X|\mathcal{B}] = \frac{d\mu_X}{d\mathbb{P}_B} \).

If we assume that \( X \in L^2(\Omega, \mathcal{F}, \mathbb{P}) \) then we can derive the existence of the conditional expectation \( Y = E[X|\mathcal{B}] \) by means of the theory of Hilbert spaces. Namely, since \( L^2(\Omega, \mathcal{F}, \mathbb{P}) \) is a Hilbert space and \( L^2(\Omega, \mathcal{B}, \mathbb{P}) \) is a (closed, linear) subspace of \( L^2(\Omega, \mathcal{F}, \mathbb{P}) \), there exists a unique element \( \tilde{Y} \in L^2(\Omega, \mathcal{B}, \mathbb{P}) \) such that

\[
E[XZ] = E[\tilde{Y}Z] \quad \text{for all} \quad Z \in L^2(\Omega, \mathcal{B}, \mathbb{P}).
\]

By the defining property of the conditional expectation we already know that

\[
E[X1_B] = E[Y1_B] \quad \text{for all} \quad B \in \mathcal{B}.
\] (B.1)

Then (B.1) holds for every simple function, and by the monotone class and Beppo Levi theorems also for every non-negative \( Z \in L^2(\Omega, \mathcal{B}, \mathbb{P}) \). Since every \( Z \in L^2(\Omega, \mathcal{B}, \mathbb{P}) \) can be written as \( Z = Z^+ - Z^- \), we obtain that

\[
\tilde{Y} = Y.
\]

Therefore we conclude that if \( X \in L^2(\Omega, \mathcal{F}, \mathbb{P}) \), then \( Y = E[X|\mathcal{B}] \) is the unique element in \( L^2(\Omega, \mathcal{B}, \mathbb{P}) \) such that for every \( Z \in L^2(\Omega, \mathcal{B}, \mathbb{P}) \),

\[
E[XZ] = E[YZ].
\]

This means that the conditional expectation \( Y = E[X|\mathcal{B}] \) is simply the Hilbert space projection of \( X \) on the closed, linear subspace \( L^2(\Omega, \mathcal{B}, \mathbb{P}) \) of \( L^2(\Omega, \mathcal{F}, \mathbb{P}) \). In a similar way we can prove the following.

**Proposition B.0.3.**

i) Let \( X \geq 0 \). Then for every \( \mathcal{B} \)-measurable \( Z \geq 0 \), \( Y = E[X|\mathcal{B}] \) verifies

\[
E[XZ] = E[YZ].
\]

ii) Let \( X \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \). Then for every bounded \( \mathcal{B} \)-measurable \( Z \), \( Y = E[X|\mathcal{B}] \) verifies

\[
E[XZ] = E[YZ].
\]
We finally recall some of the more important properties of conditional expectations.

**Proposition B.0.4.** We assume all the following conditional expectations defined without ambiguity.

1. If $\mathcal{B} = \{\emptyset, \Omega\}$, $E[ X | \mathcal{B} ] = E[ X ]$.
2. $E[ aX + bY | \mathcal{B} ] = aE[ X | \mathcal{B} ] + bE[ Y | \mathcal{B} ]$.
3. If $X \leq Y$ a.s. then $E[ X | \mathcal{B} ] \leq E[ Y | \mathcal{B} ]$ a.s.
5. If $X$ is $\mathcal{B}$-measurable, $E[ XY | \mathcal{B} ] = XE[ Y | \mathcal{B} ]$ a.s.
6. If $X$ is independent of $\mathcal{B}$, then $E[ X | \mathcal{B} ] = E[ X ]$.
7. If $\mathcal{C} \subset \mathcal{B}$ is a sub-$\sigma$-algebra of $\mathcal{B}$, then $E[ E[ X | \mathcal{B} ] | \mathcal{C} ] = E[ X | \mathcal{C} ]$.
8. Schwartz’s inequality: If $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, then $E[ XY | \mathcal{B} ] \leq E[ X^2 | \mathcal{B} ]^{1/2} E[ Y^2 | \mathcal{B} ]^{1/2}$.
9. Jensen’s inequality: If $f : \mathbb{R} \to \mathbb{R}$ is a convex function, then $f( E[ X | \mathcal{B} ] ) \leq E[ f(X) | \mathcal{B} ]$.
10. Beppo Levi’s theorem (monotone convergence theorem): If $(X_n)_{n \in \mathbb{N}} \geq 0$ and $X_n \uparrow X$ a.s., then $E[ X_n | \mathcal{B} ] \uparrow E[X | \mathcal{B}]$ a.s.
11. Fatou’s lemma: If $(X_n)_{n \in \mathbb{N}} \geq 0$, then $E\left[ \liminf_{n \to \infty} X_n | \mathcal{B} \right] \leq \liminf_{n \to \infty} E[ X_n | \mathcal{B} ]$ a.s.
12. Lebesgue’s theorem (dominated convergence theorem): If for all $n \in \mathbb{N}$, $|X_n| \leq V$ for $V \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $X_n \to X$ a.s., then $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $E[ X_n | \mathcal{B} ] \to E[X | \mathcal{B}]$ a.s.
13. If $X : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ is independent of $\mathcal{B}$ and $Y : (\Omega, \mathcal{B}) \to (H, \mathcal{H})$ is $\mathcal{B}$-measurable, then for every positive or bounded $\Phi$ on $(E \times H, \mathcal{E} \otimes \mathcal{H})$, we have $E[ \Phi(X, Y) | \mathcal{B} ] = E[ \Phi(X, y) ] |_{y=Y}$. 

B  Conditional Expectation  131
Let $F : \mathbb{R} \to \mathbb{R}$ be a deterministic function that is right-continuous and increasing. The function $F$ induces a measure on $\mathbb{R}$ in the following way. Consider the algebra

$$A = \left\{ A = \bigcup_{k=1}^{n} (a_k, b_k], \ -\infty \leq a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n \leq \infty \right\}.$$

Define on $A$ the measure

$$\mu(A) = \sum_{k=1}^{n} F(b_k) - F(a_k),$$

where

$$F(+\infty) = \lim_{x \to +\infty} F(x)$$
$$F(-\infty) = \lim_{x \to -\infty} F(x).$$

It is clear that $\mu$ is additive on $A$, one can also show that $\mu$ is $\sigma$-finite on $A$ (see e.g. Jacod and Protter [8], Theorem 7.2). In addition

$$\mu((a, b]) = F(b) - F(a).$$

We use the following theorem to uniquely extend $\mu$ from $A$ to $\sigma(A) = \mathcal{B}(\mathbb{R})$.

**Theorem C.0.1 (Carathéodory).** Let $\mu$ be a measure on an algebra $A$, then $\mu$ can be extended to $\sigma(A)$. If in addition $\mu$ is $\sigma$-finite, this extension is unique.

We can summarize what we have proved.

**Theorem C.0.2.** Let $F : \mathbb{R} \to \mathbb{R}$ be a right-continuous increasing function. Then there exists a unique measure $\mu$ on $\mathcal{B}(\mathbb{R})$ such that for every $a < b$,

$$\mu((a, b]) = F(b) - F(a).$$
We conclude this appendix with the following definition.

**Definition C.0.3.** Let $f : \mathbb{R}^+ \to \mathbb{R}$ with $f(0) = 0$. Consider

$$S^f(t) = \sup_{\sigma} \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|,$$

where $\sigma : 0 = t_0 < t_1 < \cdots < t_n = t$ is a partition of $[0, t]$. If for every $t \in \mathbb{R}^+$, $S^f(t)$ is finite, we say that $f$ has finite variation. The function $\ t \to S^f(t)$

is called the total variation of $f$. 
Lévy Processes

The main reference for this appendix is Cont and Tankov [3]. Throughout the whole chapter, assume given a probability space \((\Omega, \mathcal{F}, \mathbb{P})\). We start with the definition of Lévy Processes.

**Definition D.0.1.** A càdlàg stochastic process \((X_t)_{t \in \mathbb{R}^+}\) on \((\Omega, \mathcal{F}, \mathbb{P})\) with values in \(\mathbb{R}^d\) and such that \(X_0 = 0\) is called a Lévy process if it has the following properties:

i) independent increments

ii) stationary increments

iii) stochastic continuity:

\[
\forall \varepsilon > 0 \quad \lim_{h \to 0} \mathbb{P}(\|X_{t+h} - X_t\| \geq \varepsilon) = 0.
\]

We remark that this third property does not imply that the sample paths are continuous. It means that for a given time \(t > 0\), the probability of seeing a jump at a deterministic time \(t\) is zero, i.e. jumps occur at random times. By Definition D.0.1 it follows that a Brownian motion is a Lévy process. In the sequel we now characterize the structure of a Lévy process in full generality. Given a Lévy process \(X\), we can associate to it a random walk as follows. We sample \(X\) at regular intervals \(0, \Delta, 2\Delta, \ldots, n\Delta\) and define the random walk

\[
S_n(\Delta) := X_n\Delta = \sum_{k=0}^{n-1} Y_k
\]

with \(Y_k = X_{(k+1)\Delta} - X_k\Delta\). If we choose \(n\Delta = t\), we see that for any \(t > 0\) and \(n \geq 1\), \(X_t = S_n(\Delta)\) can be represented as the sum of \(n\) random variables whose distribution is the same as the one of \(X_{t/n}\). A distribution with this property is called infinitely divisible.

**Definition D.0.2.** A probability distribution \(F\) on \(\mathbb{R}^d\) is said to be infinitely divisible if for any integer \(n \geq 2\), there exist \(n\) i.i.d. random variables \(Y_1, \ldots, Y_n\) such that \(Y_1 + \ldots + Y_n\) has distribution \(F\).
We obtain that the distribution of a Lévy process must be infinitely divisible. In particular also the converse is true, as shown in the following result.

**Proposition D.0.3.** Let \((X_t)_{t \in \mathbb{R}^+}\) be a Lévy process. Then for every \(t\), \(X_t\) has an infinitely divisible distribution. Conversely, if \(F\) is an infinitely divisible distribution, then there exists a Lévy process \((X_t)_{t \in \mathbb{R}^+}\), such that the distribution of \(X_1\) is given by \(F\).

Examples of infinitely divisible laws are the Gaussian, gamma, \(\alpha\)-stable, Poisson, log-normal, Pareto and Student’s \(t\)-distribution. A probability distribution that is not infinitely divisible is the uniform law on an interval.

We now consider the characteristic function of a Lévy process:

\[
\Phi_t(z) = \Phi_{X_1}(z) = \mathbb{E}[e^{iz \cdot X_t}], \quad z \in \mathbb{R}^d.
\]

By using the fact that the increments of a Lévy process are independent and stationary, we obtain

\[
\Phi_{t+s}(z) = \mathbb{E}[e^{iz \cdot ((X_{t+s} - X_s) + X_s)}]
= \mathbb{E}[e^{iz \cdot (X_{t+s} - X_s)}]\Phi_s(z)
= \mathbb{E}[e^{iz \cdot X_s}]\Phi_s(z)
= \Phi_t(z)\Phi_s(z).
\]  

(D.1)

Equation (D.1) implies that \(t \mapsto \Phi_t(z)\) is an exponential function. In particular we have the following proposition.

**Proposition D.0.4.** Let \((X_t)_{t \in \mathbb{R}^+}\) be a Lévy process with values in \(\mathbb{R}^d\). There exists a continuous function \(\Psi : \mathbb{R}^d \to \mathbb{R}\) called the characteristic exponent of \(X\), such that

\[
\mathbb{E}[e^{iz \cdot X_t}] = e^{t \cdot \Psi(z)}, \quad z \in \mathbb{R}^d.
\]

The law of \(X_t\) is then determined by the knowledge of the probability distribution of \(X_1\).

We now characterize Lévy processes with constant sample paths.

**Proposition D.0.5.** \((X_t)_{t \in \mathbb{R}^+}\) is a compound Poisson process if and only if it is a Lévy process with piecewise constant sample paths.

**Proof.** We only show that if \((X_t)_{t \in \mathbb{R}^+}\) is a compound Poisson process, then it is a Lévy process. First we recall the definition of a compound process: a compound Poisson process with intensity \(\lambda\) and jump size distribution \(F\) is a stochastic process \((X_t)_{t \in \mathbb{R}^+}\) defined as

\[
X_t = \sum_{i=1}^{N_t} Y_i,
\]
where \((Y_i)_{i \in \mathbb{N}}\) are i.i.d. random variables with distribution \(F\) and \((N_t)_{t \in \mathbb{R}_+}\) is a Poisson process with intensity \(\lambda\), independent of \((Y_i)_{i \in \mathbb{N}}\).

Now let \((X_t)_{t \in \mathbb{R}_+} = \left(\sum_{i=1}^{N_t} Y_i\right)_{t \in \mathbb{R}_+}\) be a compound Poisson process. We first prove the independence of increments. Let \(s < t\) and \(f, g\) be bounded Borel functions with values in \(\mathbb{R}^d\). We prove only that \(X_s\) and \(X_t - X_s\) are independent, but the same arguments also hold for any finite number of increments.

We have

\[
E[f(X_s)g(X_t - X_s)] =
E \left[ f \left( \sum_{i=1}^{N_s} Y_i \right) g \left( \sum_{i=N_s+1}^{N_t} Y_i \right) \right]
= E \left[ E \left[ f \left( \sum_{i=1}^{N_s} Y_i \right) \left| N_u, u \leq s \right. \right] g \left( \sum_{i=N_s+1}^{N_t} Y_i \right) \left| N_u, u \leq s \right. \right]
= E \left[ f \left( \sum_{i=1}^{N_s} Y_i \right) \cdot E \left[ g \left( \sum_{i=x+1}^{y} Y_i \right) \right] \right]_{x=N_s, y=N_t}
= E[f(N_s)\varphi(N_t - N_s)]
= E[f(N_s)]E[\varphi(N_t - N_s)]
= E \left[ f \left( \sum_{i=1}^{N_s} Y_i \right) \right] E \left[ g \left( \sum_{i=N_s+1}^{N_t} Y_i \right) \left| N_u, u \leq t \right. \right]
= E[f(X_s)]E[g(X_t - X_s)].
\]

Here we have used the fact that \(N, Y_i, i \in \mathbb{N}\) are independent. Implicitly we have already proved that the increments of a compound Poisson process are stationary. In fact, we have that for a given bounded Borel function \(f\),

\[
E[f(X_t - X_s)] = E \left[ f \left( \sum_{i=N_s+1}^{N_t} Y_i \right) \left| N_u, u \leq s \right. \right]
= E \left[ f \left( \sum_{i=1}^{N_t - N_s} Y_i \right) \left| N_u, u \leq s \right. \right]
= E \left[ f \left( \sum_{i=1}^{N_t - N_s} Y_i \right) \left| N_u, u \leq s \right. \right]
\]
To prove the stochastic continuity, we note that $X_t$ jumps if $N_t$ does. We have that

$$N_s \overset{s \to t}{\longrightarrow} N_t \quad \text{a.s.,} \quad t > 0,$$

hence

$$X_s \overset{s \to t}{\longrightarrow} X_t \quad \text{a.s.,} \quad t > 0,$$

i.e. $X_s \overset{s \to t}{\longrightarrow} X_t$ also in probability. \hfill \Box

We now compute the characteristic function of a compound Poisson process.

**Proposition D.0.6.** Let $(X_t)_{t \in \mathbb{R}^+}$ be a compound Poisson process with values in $\mathbb{R}^d$, jump intensity $\lambda$ and jump size distribution $f$. Its characteristic function is given by

$$E[e^{iu \cdot X_t}] = \exp \left[ t \lambda \int_{\mathbb{R}^d} (e^{iu \cdot x} - 1)f(dx) \right], \quad u \in \mathbb{R}^d.$$ 

**Proof.** If we put $\hat{f}(u) = E[e^{iu \cdot Y_1}]$, we have

$$E[e^{iu \cdot X_t}] = E[E[e^{iu \cdot X_t} \mid N_t]] = E[E[e^{iu \cdot Y_1}]^{N_t}] = E[\hat{f}(u)^{N_t}] = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \hat{F}(u)^n e^{-\lambda t} = \exp[\lambda t(\hat{F}(u) - 1)] = \exp \left[ t \lambda \int_{\mathbb{R}^d} (e^{iu \cdot x} - 1)F(dx) \right].$$ \hfill (D.2)

If we introduce a new measure

$$\nu(A) = \lambda F(A), \quad A \in \mathcal{B}(\mathbb{R}^d),$$

then

$$E[e^{iu \cdot X_t}] = \exp \left[ t \int_{\mathbb{R}^d} (e^{iu \cdot x} - 1)\nu(dx) \right].$$

We have that $\nu(A) > 0$, $A \in \mathcal{B}(\mathbb{R}^d)$, but $\nu$ is not a probability measure, since
Formula (D.2) is a particular case of the Lévy-Khinchin representation that we show in the sequel. Every càdlàg process \((X_t)_{t \in \mathbb{R}_+}\) on \(\mathbb{R}^d\) can be associated with a random measure on \(\mathbb{R}_+ \times \mathbb{R}^d\) describing the jumps of \(X\), given by

\[
\mathcal{I}_X(B) = \#\{(t, X_t - X_{t^-}) \in B\}
\]

for every \(B \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^d)\). Given \(A \in \mathcal{B}(\mathbb{R}^d)\), \(\mathcal{I}_X([t_1, t_2] \times A)\) counts the jumps between \(t_1\) and \(t_2\) whose jump sizes are in \(A\). In particular we have that

**Proposition D.0.7.** Let \((X_t)_{t \in \mathbb{R}_+}\) be a compound Poisson process with intensity \(\lambda\) and jump size distribution \(F\). Its jump measure \(\mathcal{I}_X\) is a Poisson random measure on \(\mathbb{R}_+ \times \mathbb{R}^d\) with intensity measure

\[
\mu(dt \times dx) = \nu(dx)dt = \lambda F(dx)dt.
\]

**Definition D.0.8 (Poisson random measure).** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space, \(E \subset \mathbb{R}^d\), \(\mathcal{E}\) a σ-algebra on \(E\) and \(\mu\) a given positive Radon measure* on \((E, \mathcal{E})\). A Poisson random measure on \(E\) with intensity measure \(\mu\) is an integer valued random measure

\[
M : \Omega \times E \to \mathbb{N}
\]

\[
(\omega, A) \mapsto M(\omega, A)
\]

such that

i) for almost every \(\omega \in \Omega\), \(M(\omega, \cdot)\) is an integer-valued Radon measure on \(E\),

ii) for each \(A \in \mathcal{E}\), \(M(\cdot, A)\) is a Poisson random variable with parameter \(\mu(A)\):

\[
\forall k \in \mathbb{N} : \mathbb{P}(M(\omega, A) = k) = e^{-\mu(A)} \frac{\mu(A)^k}{k!},
\]

iii) for disjoint sets \(A_1, \ldots, A_n \in \mathcal{E}\), the random variables \(M(\cdot, A_1), \ldots, M(\cdot, A_n)\) are independent.

**Definition D.0.9.** Let \((X_t)_{t \in \mathbb{R}_+}\) be a Lévy process on \(\mathbb{R}^d\). The measure \(\mu\) on \(\mathbb{R}^d\) defined by

\[
\nu(A) = E[\#\{t \in [0, 1] : \triangle X_t \neq 0, \triangle X_t \in A\}], \quad A \in \mathcal{B}(\mathbb{R}^d)
\]

is called Lévy measure of \(X\).

* A Radon measure on \((E, \mathcal{E})\) is a measure \(\mu\) such that \(\mu(B) < \infty\) for every compact measurable set \(B \in \mathcal{E}\).
Here $\nu(A)$ is the expected number of jumps of size $A$ per unit of time. As a consequence of Proposition D.0.7 and Definition D.0.8, we obtain that every compound Poisson process can be written as

$$X_t = \sum_{s \in [0,t]} \triangle X_s = \int_{[0,t] \times \mathbb{R}^d} x I_X(ds \times dx),$$

where $I_X$ is a Poisson random measure with intensity $\nu(dx)dt$. In particular the Lévy measure of a compound Poisson process is equal to $\nu(dx) = \lambda F(dx)$. We now consider a compound Poisson process $(X^0_t)_{t \in \mathbb{R}_+}$ and a Brownian motion with drift $(B_t + \gamma t)_{t \in \mathbb{R}_+}, \gamma \in \mathbb{R}_+$, independent of $X^0$. Then the sum

$$X_t = X^0_t + \gamma t + B_t$$

defines another Lévy process, that can be decomposed as

$$X_t = \gamma t + B_t + \sum_{s \in [0,t]} \triangle X_s$$

$$= \gamma t + B_t + \sum_{s \in [0,t]} \triangle X^0_s$$

$$= \gamma t + B_t + \int_{[0,t] \times \mathbb{R}^d} x I_X(ds \times dx), \quad (D.3)$$

where $I_X$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ with intensity $\nu(dx)dt$. A natural question is then if any Lévy process has a decomposition as in (D.3). The answer is given by the following theorem.

**Theorem D.0.10 (Lévy-Itô decomposition).** Let $(X_t)_{t \in \mathbb{R}_+}$ be a Lévy process with values in $\mathbb{R}^d$ and $\nu$ its Lévy measure. Then

i) $\nu$ is a Radon measure on $\mathbb{R}^d \setminus \{0\}$ with

$$\int_{\|x\| \leq 1} \|x\|^2 \nu(dx) < \infty, \quad \int_{\|x\| \geq 1} \nu(dx) < \infty.$$

ii) The jump measure of $X$, denoted by $I_X$, is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ with intensity measure $\nu(dx)dt$.

iii) There exists a vector $\gamma \in \mathbb{R}^d$ and a $d$-dimensional Brownian motion $(B_t)_{t \in \mathbb{R}_+}$ with covariance matrix $A \in \mathbb{R}^{d \times d}$ such that

$$X_t = \gamma t + B_t + X^0_t + \lim_{\varepsilon \to 0^+} \tilde{X}^\varepsilon_t, \quad (D.4)$$

where

$$X^0_t = \int_{s \in [0,t], \|x\| \leq 1} x I_X(ds \times dx)$$

and

$$\tilde{X}^\varepsilon_t = \int_{s \in [0,t], \varepsilon \leq \|x\| < 1} x (I_X(ds \times dx) - \nu(dx)ds).$$
Proof. For the proof we refer to Cont and Tankov [3].

The terms appearing in decomposition (D.4) are independent and the convergence in the last term is almost sure and uniform in \( t \). By the Lévy-Itô decomposition it follows that for every Lévy process there exist a vector \( \gamma \in \mathbb{R}^d \), a positive definite matrix \( A \in \mathbb{R}^{d \times d} \) and a positive measure \( \nu \) that uniquely characterize its distribution. The triplet \((A, \nu, \gamma)\) is thus called characteristic triplet or Lévy triplet of the process \( X \). By (D.4) we obtain that \( \gamma t + AB_t \) is a continuous Gaussian Lévy process and that every Gaussian Lévy process is continuous and can be written as \( \gamma t + AB_t \). The other two terms generated by the jumps of \( X \) are discontinuous and described by the Lévy measure \( \nu \).

The condition

\[
\int_{\|y\| \geq 1} \nu(dy) < \infty
\]

means that \( X \) has a finite number of jumps with jump size norm larger than one. Hence the sum

\[
X^I_t = \sum_{s \in [0,t] \backslash \{ \alpha \in \mathbb{R} \backslash \mathbb{Z} \}} \triangle X_s
\]

contains almost surely a finite number of terms and \((X^I_t)_{t \in \mathbb{R}_+}\) is a compound Poisson process. There is nothing special about the threshold \( \triangle X = 1 \). For example the sum of jumps with amplitude between \( \varepsilon \) and 1

\[
X^I_t = \sum_{s \in [0,t] \backslash \{ \alpha \in \mathbb{R} \backslash \mathbb{Z} \}} \triangle X_s = \int_{s \in [0,t] \backslash \{ \alpha \in \mathbb{R} \backslash \mathbb{Z} \}} x^{JX}(ds \times dx)
\]

is again a well-defined compound Poisson process. However it may happen that \( \nu \) has a singularity at 0, due to the presence of infinitely many small jumps, whose sum does not necessarily converge. In order to obtain the convergence of the last term, one has to replace the jump integral by its compensated version

\[
\tilde{X}^\varepsilon_t := \int_{s \in [0,t] \backslash \{ \alpha \in \mathbb{R} \backslash \mathbb{Z} \}} x(JX(ds \times dx) - \nu(dx)ds),
\]

that is a martingale. The convergence of \( \tilde{X}^\varepsilon_t \) for \( \varepsilon \to 0 \) follows by a central-limit type argument. The Lévy-Itô decomposition of a Lévy process implies that every Lévy process is a combination of a Brownian motion with drift and a possibly infinite sum of independent compound Poisson processes. Decomposition (D.4) was originally provided by Paul Lévy and then completed by Kiyoshi Itô. A consequence of the Lévy-Itô decomposition is the following

**Theorem D.0.11 (Lévy-Khinchin representation).** Let \((X_t)_{t \in \mathbb{R}_+}\) be a Lévy process on \( \mathbb{R}^d \) with characteristic triplet \((A, \nu, \gamma)\). Then the characteristic function of \( X_t \) is given by
\[ E[e^{iz \cdot X_t}] = e^{i \psi(z)}, \quad z \in \mathbb{R}^d \]

with

\[ \psi(z) = -\frac{1}{2} z \cdot Az + i \gamma \cdot z + \int_{\mathbb{R}^d} (e^{iz \cdot x} - 1 - iz \cdot x\mathbb{1}_{\|x\| \leq 1}) \nu(dx). \]

In particular, for \( d = 1 \) we have

\[ \psi(z) = -\frac{1}{2} Az^2 + i \gamma z + \int_{\mathbb{R}} (e^{iz \cdot x} - 1 - iz \cdot x\mathbb{1}_{|x| \leq 1}) \nu(dx). \]

**Proof.** By Theorem D.0.10 we have that for every \( t \geq 0 \), the random variable

\[ X_t^\varepsilon = \gamma t + AB_t + X^\varepsilon_t + \tilde{X}^\varepsilon_t \]

converges almost surely to \( X_t \) for \( \varepsilon \to 0 \), hence it converges to \( X_t \) also in distribution, i.e. the characteristic function of \( X_t^\varepsilon \) converges to the one of \( X_t \) for \( \varepsilon \to 0 \). Since \( \gamma t + AB_t, X^\varepsilon_t \) and \( \tilde{X}^\varepsilon_t \) are independent, we have

\[
E[\exp\{iz \cdot (\gamma t + AB_t + X^\varepsilon_t + \tilde{X}^\varepsilon_t)\}] = \\
= E[\exp\{iz \cdot (\gamma t + AB_t)\}] E[\exp\{iz \cdot X^\varepsilon_t\}] E[\exp\{iz \cdot \tilde{X}^\varepsilon_t\}] \\
= \exp(t i \gamma \cdot z) \exp\left(-\frac{1}{2} z \cdot Az\right) \exp\left(t \int_{\|x\| \geq 1} (e^{iz \cdot x} - 1) \nu(dx)\right) \\
\cdot \exp\left(t \int_{\varepsilon \leq \|x\| < 1} (e^{iz \cdot x} - 1 - iz \cdot x) \nu(dx)\right).
\]

The last integral converges to

\[
\int_{\|x\| \leq 1} (e^{iz \cdot x} - 1 - iz \cdot x\mathbb{1}_{\|x\| \leq 1}) \nu(dx)
\]

for \( \varepsilon \to 0 \). Note that here we have used the fact that

\[
E \left[ \exp \left\{ iz \cdot \left( - \int_{\varepsilon \leq \|x\| < 1} x \nu(dx) dx \right) \right\} \right] = \exp \left( -t \int_{\varepsilon \leq \|x\| < 1} iz \cdot x \nu(dx) \right),
\]

since \( \nu(dx) \) is a measure on \( \mathbb{R}^d \). \( \square \)

By using the characteristic triplet of a Lévy process, we can deduce some properties of its sample paths.

**Proposition D.0.12.** A Lévy process has piecewise constant trajectories if and only if its characteristic triplet is given by

\[
A = 0, \quad \int_{\mathbb{R}^d} \nu(dx) < \infty \quad \text{and} \quad \gamma = \int_{\|x\| \leq 1} x \nu(dx),
\]

142  D Lévy Processes
or equivalently, its characteristic exponent is

$$\psi(z) = \int_{\mathbb{R}^d} (e^{iz \cdot x} - 1)\nu(dx),$$

with $$\nu(\mathbb{R}^d) < \infty$$.

Proof. This is a consequence of Theorem D.0.10 and the Lévy-Khinchin representation Theorem D.0.11.

Other properties of Lévy processes are the following.

**Proposition D.0.13.** A Lévy process is of finite variation if and only if its characteristic triplet $$(A, \nu, \gamma)$$ satisfies

$$A = 0 \quad \text{and} \quad \int_{\|x\| \leq 1} \|x\|\nu(dx) < \infty.$$ 

**Corollary D.0.14.** Let $$(X_t)_{t \in \mathbb{R}_+}$$ be a Lévy process of finite variation with Lévy triplet given by $$(0, \nu, \gamma)$$. Then $$X$$ can be expressed as a sum of its jumps between 0 and $$t$$ and a linear drift term:

$$X_t = bt + \int_{[0,t] \times \mathbb{R}^d} x \mathbb{1}_X(ds \times dx) = bt + \sum_{s \in [0,t] \atop \Delta X_s \neq 0} \Delta X_s,$$

with characteristic function given by

$$E[e^{iz \cdot X_t}] = \exp \left \{ t \left [ ib \cdot z + \int_{\mathbb{R}^d} (e^{iz \cdot x} - 1)\nu(dx) \right ] \right \},$$

with $$b = \gamma - \int_{\|x\| \leq 1} x\nu(dx)$$. 
References


σ-algebra of progressive sets, 14
adapted process, 13
admissible strategy, 66
arbitrage (strategy), 67
arbitrage-free, 67
attainable (contingent claim), 73
backward partial differential equation, 91
barrier option, 100
Bayes rule, 75
Black-Scholes model, 62
generalized, 72
pricing formula, 85
Brownian motion, 5
construction, 6
definition, 5, 14
geometric, 56
properties, 11
with drift, 96
characteristic exponent, 120
characteristic function, 111
characteristic triplet of a Lévy process, 125
closed submartingale, 16
complete market, 74
conditional expectation, 113
existence, 113
properties, 113
contingent claim, 73
European, 73
continuous a.s. process, 3
continuous process, 3
convergence (of random variables), 109
almost sure, 109
in $L^p$, 109
in distribution/weak, 109
in probability, 109
delta- and gamma-neutral (portfolio), 94
derivative, 73
drift, 62
equivalent (local) martingale measure, 67
equivalent processes, 4
European call (option), 82
European contingent claim, 73
European put (option), 82
exotic derivative, 96
expiration date, 73
Feynman-Kac formula, 91
filtered probability space, 13
filtration, 13
complete, 13
natural, 14
right-continuous, 13
standard, 13
usual conditions, 13
finite variation, 118
finite variation (FV) process, 28
forward (contract), 82
forward price, 86
forward value, 86
Futures (contract), 82
futures price, 87
futures quotation, 87
Gaussian process, 5
geometric Brownian motion, 56
hedge, 74
hedgeable (contingent claim), 73
implied volatility, 95
independent increments process, 4
indistinguishable processes, 4
infinitely divisible probability distribution, 119
integration by parts, 46
Itô process, 44
   vector, 54
Itô’s formula, 44
Lévy measure, 123
Lévy process, 119
Lévy triplet, 125
Lévy-Itô decomposition, 124
Lévy-Khinchin representation, 125
left-continuous a.s. process, 3
left-continuous process, 3
Lipschitz, 54
local martingale, 32
locally bounded, 54
long position, 63
market price of risk, 73
marking to market, 88
martingale, 15
closed submartingale, 16
generalized positive, 15
square integrable continuous, 22
submartingale, 15
supermartingale, 15
maturity, 73
mean-reverting Ornstein-Uhlenbeck process, 57
measurable process, 4
modification of processes, 4
money market account, 62
Ornstein-Uhlenbeck process, 56
   mean-reverting, 57
plain vanilla option, 96
Poisson random measure, 123
price of a claim
   at time $t = 0$, 78
   at time $t > 0$, 81
primary assets, 73
(stochastic) process, 3
progressive process, 14
progressively measurable process, 14
put-call parity, 86
quadratic covariation, 33
quadratic variation, 33
Radon measure, 123
realization, 3
reflection equality, 97
replicating strategy, 74
Riemann-Stieltjes integral (path-wise), 29
right-continuous a.s. process, 3
right-continuous process, 3
sample path, 3
self-financing strategy, 63
semimartingale, 44
sharp bracket, 33
short position, 63
simple process, 35
stationary increments process, 5
Stieltjes integral, 27
stochastic differential, 44
stochastic differential equation, 53
   solution, 54
stochastic process, 3
stochastic volatility models, 96
stopped process, 16
stopping time, 16
stopping time $\sigma$-algebra, 16
sub-linear growth, 54
tame strategy, 66
total variation, 118
total variation process, 28
trajectory, 3
underlying, 73
vector Itô process, 54
volatility, 62
volatility smile, 96
Wiener process, 5