NO-ARBITRAGE THEORY FOR DERIVATIVES PRICING

Nizar TOUZI, Peter TANKOV

Ecole Polytechnique Paris
Département de Mathématiques Appliquées

peter.tankov@polytechnique.edu

January 2012
# Contents

1 Introduction 4
   1.1 European and American options 5
   1.2 No dominance principle and first properties 7
   1.3 Put-Call Parity 9
   1.4 Bounds on call prices and early exercise of American calls 10
   1.5 Risk effect on options prices 11
   1.6 Some popular examples of contingent claims 13

2 The finite probability one-period model 16
   2.1 The model 16
   2.2 Arbitrage-free markets 18
   2.3 Complete markets and pricing of attainable claims 20
   2.4 Asset pricing in incomplete markets 22
   2.5 Portfolio optimization and equilibrium asset pricing 24
      2.5.1 The portfolio optimization problem 25
      2.5.2 Finding the optimal contingent claim 26
      2.5.3 Equilibrium asset pricing 29

3 The finite probability multiperiod model 31
   3.1 The model 31
   3.2 Arbitrage-free markets 33
   3.3 Evaluation of contingent claims 35
   3.4 Complete markets 36
6.2 Itô’s formula .......................................................... 93
   6.2.1 Extension to Itô processes ................................. 95
6.3 The Feynman-Kac representation formula .................. 98
6.4 The Cameron-Martin change of measure .................... 100
6.5 Complement: density of simple processes in $\mathbb{H}^2$ .... 102

7 Financial markets in continuous time .......................... 105
   7.1 Portfolio and wealth process ................................ 107
   7.2 Admissible portfolios and no-arbitrage ..................... 108
   7.3 The Merton consumption/investment problem .............. 112
      7.3.1 The Hamilton-Jacobi-Bellman equation ............... 113
      7.3.2 Pure investment problem with power utility ........... 114
      7.3.3 Investment-consumption problem in infinite horizon .. 116
      7.3.4 The martingale approach ............................... 118

8 The Black-Scholes valuation theory ............................. 120
   8.1 Pricing and hedging Vanilla options ....................... 120
   8.2 The PDE approach for the valuation problem .............. 122
   8.3 The Black and Scholes model for European call options ... 124
      8.3.1 The Black-Scholes formula ............................ 124
      8.3.2 The Black’s formula .................................. 127
      8.3.3 Option on a dividend paying stock .................... 128
      8.3.4 The practice of the Black-Scholes model .............. 130
Chapter 1

Introduction

Throughout these notes, we shall consider a financial market defined by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a set of primitive assets $(S^1, \ldots, S^d)$ which are available for trading. In addition, agents trade derivative contracts on these assets, which we call financial assets, and which are the main focus of this course. While the total offer of primitive assets is non-zero, the total offer in financial asset is zero: these are contracts between two parties.

In this introduction, we focus on some popular examples of derivative assets, and we provide some properties that their prices must satisfy independently of the distribution of the primitive assets prices. The only ingredient which will be used in order to derive these properties is the no dominance principle introduced in Section 1.2 below.

The first part of these notes considers finite discrete-time financial markets on a finite probability space in order to study some questions of theoretical nature. The second part concentrates on continuous-time financial markets. Although continuous-time models are more demanding from the technical viewpoint, they are widely used in the financial industry because of the simplicity of the implied formulae for pricing and hedging. This is related to the powerful tools of differential calculus which are available only in continuous-time. We shall first provide a self-contained introduction of the main concepts from stochastic analysis: Brownian motion, Itô’s formula,
Cameron-Martin Theorem, connection with the heat equation. We then consider the Black-Scholes continuous-time financial market and derive various versions of the famous Black-Scholes formula. A discussion of the practical use of these formulae is provided.

1.1 European and American options

The most popular examples of derivative securities are European and American call and put options. More examples of contingent claims are listed in Section 1.6 below.

A European call option on the asset $S^i$ is a contract where the seller promises to deliver the risky asset $S^i$ at the maturity $T$ for some given exercise price, or strike, $K > 0$. At time $T$, the buyer has the possibility (and not the obligation) to exercise the option, i.e. to buy the risky asset from the seller at strike $K$. Of course, the buyer would exercise the option only if the price which prevails at time $T$ is larger than $K$. Therefore, the gain of the buyer out of this contract is

$$B = (S_T^i - K)^+ = \max\{S_T^i - K, 0\},$$

i.e. if the time $T$ price of the asset $S^i$ is larger than the strike $K$, then the buyer receives the payoff $S_T^i - K$ which corresponds to the benefit from buying the asset from the seller of the contract rather than on the financial market. If the time $T$ price of the asset $S^i$ is smaller than the strike $K$, the contract is worthless for the buyer.

A European put option on the asset $S^i$ is a contract where the seller promises to purchase the risky asset $S^i$ at the maturity $T$ for some given exercise price, or strike, $K > 0$. At time $T$, the buyer has the possibility, and not the obligation, to exercise the option, i.e. to sell the risky asset to the seller at strike $K$. Of course, the buyer would exercise the option only if the price which prevails at time $T$ is smaller than $K$. Therefore, the gain of
the buyer out of this contract is

\[ B = (K - S_T^i)^+ = \max\{K - S_T^i, 0\}, \]

i.e. if the time \( T \) price of the asset \( S^i \) is smaller than the strike \( K \), then the buyer receives the payoff \( K - S_T^i \) which corresponds to the benefit from selling the asset to the seller of the contract rather than on the financial market. If the time \( T \) price of the asset \( S^i \) is larger than the strike \( K \), the contract is worthless for the buyer, as he can sell the risky asset for a larger price on the financial market.

An American call (resp. put) option with maturity \( T \) and strike \( K > 0 \) differs from the corresponding European contract in that it offers the possibility to be exercised at any time before maturity (and not only at the maturity).

The seller of a derivative security requires a compensation for the risk that he is bearing. In other words, the buyer must pay the price or the premium for the benefit of the contract. The main interest of this course is to determine this price. In the subsequent sections of this introduction, we introduce the no dominance principle which already allows to obtain some model-free properties of options which hold both in discrete and continuous-time models.

In the subsequent sections, we shall consider call and put options with exercise price (or strike) \( K \), maturity \( T \), and written on a single risky asset with price \( S \). At every time \( t \leq T \), the American and the European call option price are respectively denoted by

\[ C(t, S_t, T, K) \text{ and } c(t, S_t, T, K). \]

Similarly, the prices of the American and the European put options are respectively denoted by

\[ P(t, S_t, T, K) \text{ and } p(t, S_t, T, K). \]
The intrinsic value of the call and the put options are respectively:

\[
C(t, S_t, T, K) = c(t, S_t, T, K) = (S_t - K)^+ \quad \text{and} \quad P(t, S_t, T, K) = p(t, S_t, T, K) = (K - S_t)^+ ,
\]

i.e. the value received upon immediate exercising the option. An option is said to be in-the-money (resp. out-of-the-money) if its intrinsic value is positive. If \( K = S_t \), the option is said to be at-the-money. Thus a call option is in-the-money if \( S_t > K \), while a put option is in-the-money if \( S_t < K \).

Finally, a zero-coupon bond is the discount bond defined by the fixed income 1 at the maturity \( T \). We shall denote by \( B_t(T) \) its price at time \( t \). Given the prices of zero-coupon bonds with all maturity, the price at time \( t \) of any stream of deterministic payments \( F_1, \ldots, F_n \) at the maturities \( t < T_1 < \ldots < T_n \) is given by

\[
F_1 B_t(T_1) + \ldots + F_n B_t(T_n) .
\]

### 1.2 No dominance principle and first properties

We shall assume that there are no market imperfections as transaction costs, taxes, or portfolio constraints, and we will make use of the following concept.

**No Dominance principle**  Let \( X \) be the gain from a portfolio strategy with initial cost \( x \). If \( X \geq 0 \) in every state of the world, Then \( x \geq 0 \).

Notice that, choosing to exercise the American option at the maturity \( T \) provides the same payoff as the European counterpart. Then the portfolio consisting of a long position in the American option and a short position in the European counterpart has at least a zero payoff at the maturity \( T \). It then follows from the dominance principle that American calls and puts are at least as valuable as their European counterparts:

\[
C(t, S_t, T, K) \geq c(t, S_t, T, K) \quad \text{and} \quad P(t, S_t, T, K) \geq p(t, S_t, T, K)
\]
By a similar easy argument, we now show that American and European call (resp. put) options prices are decreasing (resp. increasing) in the exercise price, i.e. for $K_1 \geq K_2$:

$$C(t, S_t, T, K_1) \leq C(t, S_t, T, K_2) \quad \text{and} \quad c(t, S_t, T, K_1) \leq c(t, S_t, T, K_2)$$

$$P(t, S_t, T, K_1) \geq P(t, S_t, T, K_2) \quad \text{and} \quad p(t, S_t, T, K_1) \leq p(t, S_t, T, K_2)$$

Let us justify this for the case of American call options. If the holder of the low exercise price call adopts the optimal exercise strategy of the high exercise price call, the payoff of the low exercise price call will be higher in all states of the world. Hence, the value of the low exercise price call must be no less than the price of the high exercise price call.

4. American/European Call and put prices are convex in $K$. Let us justify this property for the case of American call options. For an arbitrary time instant $u \in [t, T]$ and $\lambda \in [0, 1]$, it follows from the convexity of the intrinsic value that

$$\lambda (S_u - K_1)^+ + (1 - \lambda) (S_u - K_2)^+ - (S_u - \lambda K_1 + (1 - \lambda) K_2)^+ \geq 0.$$  

We then consider a portfolio $X$ consisting of a long position of $\lambda$ calls with strike $K_1$, a long position of $(1 - \lambda)$ calls with strike $K_2$, and a short position of a call with strike $\lambda K_1 + (1 - \lambda) K_2$. If the two first options are exercised on the optimal exercise date of the third option, the resulting payoff is non-negative by the above convexity inequality. Hence, the value at time $t$ of the portfolio is non-negative.

4. We next show the following result for the sensitivity of European call options with respect to the exercise price:

$$-B_t(T) \leq \frac{c(t, S_t, T, K_2) - c(t, S_t, T, K_1)}{K_2 - K_1} \leq 0, \quad K_2 > K_1.$$  

The right hand-side inequality follows from the decreasing nature of the European call option $c$ in $K$. To see that the left hand-side inequality holds, consider the portfolio $X$ consisting of a short position of the European call
with exercise price $K_1$, a long position of the European call with exercise price $K_2$, and a long position of $K_2 - K_1$ zero-coupon bonds. The value of this portfolio at the maturity $T$ is

$$X_T = -(S_T - K_1)^+ + (S_T - K_2)^+ + (K_2 - K_1) \geq 0.$$  

By the dominance principle, this implies that $-c(S_t, \tau, K_1) + c(S_t, \tau, K_2) + B_t(\tau)(K_2 - K_1) \geq 0$, which is the required inequality.

5. **American call and put prices are increasing in maturity**, i.e. for $T_1 \geq T_2$:

$$C(t, S_t, T_1, K) \geq C(t, S_t, T_2, K) \text{ and } P(t, S_t, T_1, K_1) \geq P(t, S_t, T_2, K_2)$$

This is a direct consequence of the fact that all stopping strategies of the shorter maturity option are allowed for the longer maturity one. Notice that this argument is specific to the American case.

### 1.3 Put-Call Parity

When the underlying security pays no income before the maturity of the options, the prices of calls and puts are related by

$$p(t, S_t, T, K) = c(t, S_t, T, K) - S_t + KB_t(T).$$

Indeed, Let $X$ be the portfolio consisting of a long position of a European put option and one unit of the underlying security, and a short position of a European call option and $K$ zero-coupon bonds. The value of this portfolio at the maturity $T$ is

$$X_T = (K - S_T)^+ + S_T - (S_T - K)^+ - K = 0.$$  

By the dominance principle, the value of this portfolio at time $t$ is non-negative, which provides the required inequality.

Notice that this argument is specific to European options. We shall see in fact that the corresponding result does not hold for American options.
Finally, if the underlying asset pays out some dividends then, the above argument breaks down because one should account for the dividends received by holding the underlying asset \( S \). If we assume that the dividends are known in advance, i.e. non-random, then it is an easy exercise to adapt the put-call parity to this context. However, if the dividends are subject to uncertainty as in real life, there is no direct way to adapt the put-call parity.

### 1.4 Bounds on call prices and early exercise of American calls

1. From the monotonicity of American calls in terms of the exercise price, we see that

\[
c(S_t, \tau, K) \leq C(S_t, \tau, K) \leq S_t
\]

- When the underlying security pays no dividends before maturity, we have the following lower bound on call options prices:

\[
C(t, S_t, T, K) \geq c(t, S_t, T, K) \geq (S_t - KB_t(T))^+.
\]

Indeed, consider the portfolio \( X \) consisting of a long position of a European call, a long position of \( K \) maturity zero-coupon bonds, and a short position of one share of the underlying security. The required result follows from the observation that the final value at the maturity of the portfolio is non-negative, and the application of the dominance principle.

2. Assume that interest rates are positive. Then, an American call on a security that pays no dividend before the maturity of the call will never be exercised early.

   Indeed, let \( u \) be an arbitrary instant in \([t, T)\),

   - the American call pays \( S_u - K \) if exercised at time \( u \),
   - but \( S_u - K < S - KB_u(T) \) because interest rates are positive.
   - Since \( C(S_u, u, K) \geq S_u - KB_u(T) \), by the lower bound, the American
option is always worth more than its exercise value, so early exercise is never optimal.

3 Assume that the security price takes values as close as possible to zero. Then, early exercise of American put options may be optimal before maturity.

Suppose the security price at some time \( u \) falls so deeply that \( S_u < K - KB_u(T) \).

- Observe that the maximum value that the American put can deliver when exercised at maturity is \( K \).
- The immediate exercise value at time \( u \) is \( K - S_u > K - [K - KB_u(T)] = KB_u(T) \equiv \) the discounted value of the maximum amount that the put could pay if held to maturity,

Hence, in this case waiting until maturity to exercise is never optimal.

1.5 Risk effect on options prices

1 The value of a portfolio of European/american call/put options, with common strike and maturity, always exceeds the value of the corresponding basket option.

Indeed, let \( S^1, \ldots, S^n \) be the prices of \( n \) security, and consider the portfolio composition \( \lambda^1, \ldots, \lambda^n \geq 0 \). By sublinearity of the maximum,

\[
\sum_{i=1}^{n} \lambda^i \max \{ S^i_u - K, 0 \} \geq \max \left\{ \sum_{i=1}^{n} \lambda^i S^i_u - K, 0 \right\}
\]

i.e. if the portfolio of options is exercised on the optimal exercise date of the option on the portfolio, the payoff on the former is never less than that on the latter. By the dominance principle, this shows that the portfolio of options is more valuable than the corresponding basket option.

2 For a security with spot price \( S_t \) and price at maturity \( S_T \), we denote its return by

\[
R_t(T) := \frac{S_T}{S_t}.
\]
**Definition**  Let $R_i^2(T), i = 1, 2$ be the return of two securities. We say that security 2 is more risky than security 1 if

$$R_i^2(T) = R_i^1(T) + \varepsilon \quad \text{where} \quad \mathbb{E}[\varepsilon|R_i^1(T)] = 0.$$  

As a consequence, if security 2 is more risky than security 1, the above definition implies that

$$\text{Var}[R_i^2(T)] = \text{Var}[R_i^1(T)] + \text{Var}[\varepsilon] + 2\text{Cov}[R_i^1(T), \varepsilon] \geq \text{Var}[R_i^1(T)].$$

3  We now assume that the pricing functional is continuous in some sense to be precised below, and we show that the value of an European/American call/put is increasing in its riskiness.

To see this, let $R := R_i(T)$ be the return of the security, and consider the set of riskier securities with returns $R^i := R_i^i(T)$ defined by

$$R^i = R + \varepsilon_i \quad \text{where} \quad \varepsilon_i \text{ are iid and} \quad \mathbb{E}[\varepsilon_i|R] = 0.$$  

Let $C^i(t, S_t, T, K)$ be the price of the American call option with payoff $(S_tR^i - K)^+$, and $\overline{C}_n(t, S_t, T, K)$ be the price of the basket option defined by the payoff $(\frac{1}{n}\sum_{i=1}^n S_tR^i - K)^+ = (S_T + \frac{1}{n}\sum_{i=1}^n \varepsilon_i - K)^+.$

We have previously seen that the portfolio of options with common maturity and strike is worth more than the corresponding basket option:

$$C^i(t, S_t, T, K) = \frac{1}{n}\sum_{i=1}^n C^i(t, S_t, T, K) \geq \overline{C}_n(t, S_t, T, K).$$

Observe that the final payoff of the basket option $\overline{C}_n(T, S_T, T, K) \rightarrow (S_T - K)^+$ a.s. as $n \rightarrow \infty$ by the law of large numbers. Then assuming that the pricing functional is continuous, it follows that $\overline{C}_n(t, S_t, T, K) \rightarrow C(t, S_t, T, K)$, and therefore: that

$$C^i(t, S_t, T, K) \geq C(t, S_t, T, K).$$

Notice that the result holds due to the convexity of European/American call/put options payoffs.
1.6 Some popular examples of contingent claims

Example 1.1 (Basket call and put options) Given a subset $I$ of indices in $\{1, \ldots, n\}$ and a family of positive coefficients $(a_i)_{i \in I}$, the payoff of a Basket call (resp. put) option is defined by

$$B = \left( \sum_{i \in I} a_i S_T^i - K \right)^+ \quad \text{resp.} \quad \left( K - \sum_{i \in I} a_i S_T^i \right)^+.$$  

Example 1.2 (Option on a non-tradable underlying variable) Let $U_t(\omega)$ be the time $t$ realization of some observable state variable. Then the payoff of a call (resp. put) option on $U$ is defined by

$$B = (U_T - K)^+ \quad \text{resp.} \quad (K - U_T)^+.$$  

For instance, a Temperature call option corresponds to the case where $U_t$ is the temperature at time $t$ observed at some location (defined in the contract).

Example 1.3 (Asian option) An Asian call option on the asset $S^i$ with maturity $T > 0$ and strike $K > 0$ is defined by the payoff at maturity:

$$\left( \bar{S}_T^i - K \right)^+,$$

where $\bar{S}_T^i$ is the average price process on the time period $[0, T]$. With this definition, there is still choice for the type of Asian option in hand. One can define $\bar{S}_T^i$ to be the arithmetic mean over of given finite set of dates (outlined in the contract), or the continuous arithmetic mean...

Example 1.4 (Barrier call options) Let $B, K > 0$ be two given parameters, and $T > 0$ given maturity. There are four types of barrier call options on the asset $S^i$ with strike $K$, barrier $B$ and maturity $T$:

- When $B > S_0$: 

– an *Up and Out Call option* is defined by the payoff at the maturity $T$:

$$\text{UOC}_T = (S_T - K)^+1_{\{\max_{[0,T]} S_t \leq B}\}.$$  

The payoff is that of a European call option if the price process of the underlying asset never reaches the barrier $B$ before maturity. Otherwise it is zero (the contract knocks out).

– an *Up and In Call option* is defined by the payoff at the maturity $T$:

$$\text{UIC}_T = (S_T - K)^+1_{\{\max_{[0,T]} S_t > B\}}.$$  

The payoff is that of a European call option if the price process of the underlying asset crosses the barrier $B$ before maturity. Otherwise it is zero (the contract knocks out). Clearly,

$$\text{UOC}_T + \text{UIC}_T = C_T$$

is the payoff of the corresponding European call option.

– When $B < S_0$: 

  – an *Down and Out Call option* is defined by the payoff at the maturity $T$:

$$\text{DOC}_T = (S_T - K)^+1_{\{\min_{[0,T]} S_t > B\}}.$$  

The payoff is that of a European call option if the price process of the underlying asset never reaches the barrier $B$ before maturity. Otherwise it is zero (the contract knocks out).

  – an *Down and In Call option* is defined by the payoff at the maturity $T$:

$$\text{DIC}_T = (S_T - K)^+1_{\{\min_{[0,T]} S_t \leq B\}}.$$
The payoff is that of a European call option if the price process of the underlying asset crosses the barrier $B$ before maturity. Otherwise it is zero. Clearly,

$$\text{DOC}_T + \text{DIC}_T = C_T$$

is the payoff of the corresponding European call option.

\[\Box\]

**Example 1.5 (Barrier put options)** Replace calls by puts in the previous example
Chapter 2

The finite probability one-period model

2.1 The model

In this chapter, we consider a simple one-period financial market model with \( d + 1 \) assets and two dates: \( t = 0 \) and \( t = 1 \). The randomness in the model is represented by a finite space of the states of nature \( \Omega = \{\omega_1, \ldots, \omega_K\} \) which correspond to scenarios of market evolution between \( t = 0 \) and \( t = 1 \), with \( \sigma \)-field \( \mathcal{F} \) containing all subsets of \( \Omega \) and associated probabilities \( \mathbb{P} = \{p_1, \ldots, p_K\} \) which are supposed to be all positive. The prices of assets at time \( t \) are denoted by \( (S^i_t)_{i=0}^d \). The asset \( S^0 \) represents the risk-free bank account with deterministic evolution: \( S^0_t = 1 \) and \( S^0_t = 1 + R \), where \( R \) is the one-period interest rate, known at time \( t = 0 \). The other assets are risky: their prices at time 1 are only revealed at \( t = 1 \).

A portfolio strategy \( \theta \in \mathbb{R}^{d+1} \), is a vector giving the number of units of each asset to be held in the portfolio between date 0 and date 1. The portfolio value \( X_t \) is given by

\[
X_t = \sum_{i=0}^d \theta^i S^i_t.
\]
Sometimes we will also use the notation $X_t^\theta$ to emphasize that the portfolio is constructed using the strategy $\theta$.

The gain of a portfolio strategy is defined by

$$G = X_1 - X_0 = \sum_{i=0}^{d} \theta^i \Delta S^i, \quad \text{with} \quad \Delta S^i = S^i_1 - S^i_0.$$  

It is convenient to express the prices of all assets in the units of the risk-free asset, or, in other words, choose the bank account as numéraire. We therefore introduce discounted prices

$$\tilde{S}^i_t := \frac{S^i_t}{S^0_t}$$

and the discounted portfolio value

$$\tilde{X}_t = \frac{X_t}{S^0_t} = \theta_0 + \sum_{i=1}^{d} \theta^i \tilde{S}^i_t$$

The discounted gain of the portfolio strategy does not depend on the number of units of the risk-free asset in the portfolio:

$$\tilde{G} = \tilde{X}_1 - \tilde{X}_0 = \sum_{i=1}^{d} \theta^i \Delta \tilde{S}^i.$$

**Example 2.6** Let $\Omega = (\omega_1, \omega_2, \omega_3, \omega_4)$ with associated probabilities $P = (0.3, 0.3, 0.2, 0.2)$ and $R = 0.25$. Assume that $S^1$ is a stock whose price equals 1 at time $t = 0$, and takes values $(0.5, 1, 1.5, 2)$ in the states $(\omega_1, \omega_2, \omega_3, \omega_4)$ respectively. Finally, let $S^2$ be a call option on $S^1$ with strike $K = 1$, which costs 0.3 at time $t = 0$ and takes values $(0, 0, 0.5, 1)$ in the respective states of nature. It is convenient to represent this market using a table (see table 2.1).

We consider a strategy with zero initial cost, which consists in short-selling one unit of stock, buying one call option on this stock (to limit the losses) and placing the remaining funds in a bank. We therefore have $\theta = (0.7, -1, 1)$ and

$$G = \frac{7}{8} - S^1 + S^2 = \begin{cases} 
3 \\ 3/8 \\ -1/8 
\end{cases} \text{ in } \omega_1, \omega_2, \omega_3.$$  

17
Table 2.1: Tabular representation of the market of example 2.6.

<table>
<thead>
<tr>
<th></th>
<th>$S_0^i$</th>
<th>$\omega_1$</th>
<th>$\omega_2$</th>
<th>$\omega_3$</th>
<th>$\omega_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i = 0$</td>
<td>1</td>
<td>1.25</td>
<td>1.25</td>
<td>1.25</td>
<td>1.25</td>
</tr>
<tr>
<td>$i = 1$</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1.5</td>
<td>2</td>
</tr>
<tr>
<td>$i = 2$</td>
<td>0.3</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
</tbody>
</table>

2.2 Arbitrage-free markets

**Definition 2.1** An arbitrage opportunity is a portfolio strategy $\theta$ such that $X^\theta_0 = 0$, $\mathbb{P}[X^\theta_1 \geq 0] = 1$ and $\mathbb{P}[X^\theta_1 > 0] > 0$.

In other words, an arbitrage opportunity is a zero-cost strategy which has nonnegative pay-off in all states of nature and strictly positive pay-off in at least one state. Equivalently, one can say that a strategy $\theta$ is an arbitrage opportunity if the discounted gain of this strategy satisfies $\mathbb{P}[\tilde{G}^\theta \geq 0] = 1$ and $\mathbb{P}[\tilde{G}^\theta > 0] > 0$.

**Exercise 2.1** Prove that the two definitions of arbitrage opportunity are equivalent.

The theory of arbitrage pricing, developed for the case of discrete-time financial markets by Harrison and Kreps [13] provides a simple characterization of arbitrage-free markets in terms of the existence of a so-called risk-neutral probability.

**Definition 2.2** A risk-neutral probability is a probability measure $\mathbb{Q}$ on $\Omega$ such that

- $\mathbb{Q} \sim \mathbb{P}$, that is, $\mathbb{Q}(\omega_k) > 0$ for all $k$.
- $E^\mathbb{Q}[\tilde{S}_1^i] = \tilde{S}_0^i$ for all $i$.

Under the risk-neutral probability, the price of an asset depends only on its expected return, but not on its risk.
Theorem 2.1 A financial market admits no arbitrage opportunity if and only if there exists at least one risk-neutral probability.

Proof. The if part. Let $Q$ be a risk-neutral probability and assume that $\theta$ is an arbitrage opportunity. Then, since $E^Q[\tilde{G}^{\theta}] = 0$, $\tilde{G}^{\theta} = 0$ in all states of nature and we have a contradiction with $P[\tilde{G}^{\theta} > 0] > 0$.

The only if part. Suppose that the market is arbitrage-free, and let $Q$ be the set of all probability measures on $\Omega$ which assign nonzero probability to every state:

$$Q = \{q \in \mathbb{R}^K : q_i > 0 \forall i, \sum_{i=1}^K q_i = 1\}.$$  

Further, let $C$ be the $d$-dimensional set of all expectations of $(\Delta \tilde{S})^d_{i=1}$ under probabilities in $Q$:

$$C = \{E^Q[\Delta \tilde{S}] | Q \in Q\}.$$  

$C$ is clearly a convex set, and to prove the existence of a risk-neutral probability it is sufficient to show that $0 \in C$. Suppose on the contrary that $0 \notin C$. Then by the separating hyperplane theorem [23] there exists $H \in \mathbb{R}^d$ such that $H^T x \geq 0$ for all $x \in C$ and $Hx_0 > 0$ for some $x_0 \in C$. Let $\theta = (H_0, H)$ with $H_0$ arbitrary. For this strategy therefore $E^Q[\tilde{G}^{\theta}] \geq 0$ for all $Q \in Q$ and $E^Q[\tilde{G}^{\theta}] > 0$ for some $Q \in Q$.

Suppose that $\tilde{G}^{\theta}(\omega_i) < 0$ for some $\omega_i \in \Omega$. Taking $Q_\varepsilon = \frac{\varepsilon}{K} + (1-\varepsilon)1_{\omega_i}$ we see that $E^{Q_\varepsilon}[\tilde{G}^{\theta}] < 0$ for $\varepsilon$ sufficiently small, which is a contradiction with the above. Therefore, $\tilde{G}^{\theta} \geq 0$ in all states of nature, and $\tilde{G}^{\theta} > 0$ in at least one state: $\theta$ is an arbitrage opportunity. Since we have supposed that the market is arbitrage free, this is a contradiction, and therefore we conclude that $0 \in C$ and there exists at least one risk-neutral probability.

Example 2.7 Let us look for a risk-neutral probability in the market of ex-
Example 2.6. We need to solve the system of equations

\[
\begin{align*}
-\frac{3}{5}q_1 - \frac{1}{5}q_2 + \frac{1}{5}q_3 + \frac{3}{5}q_4 &= 0 \\
-\frac{3}{10}q_1 - \frac{3}{10}q_2 + \frac{1}{10}q_3 + \frac{1}{2}q_4 &= 0 \\
q_1 + q_2 + q_3 + q_4 &= 1.
\end{align*}
\]

This system admits an infinity of positive solutions of the form

\[
(q_1, q_2, q_3, q_4) = \left(\frac{1}{4}, q, \frac{3}{4} - 2q, q\right), \quad q \in \left(0, \frac{3}{8}\right).
\]

The market of example 2.6 is therefore arbitrage-free.

Assume now that we add an additional asset to the market of example 2.6, which is a put option with strike \(\frac{3}{2}\), quoted at the price of 0.2 at \(t = 0\). We then need to add another equation to the above system, and the only remaining solution is \(\left(\frac{1}{4}, 0, \frac{3}{4}, 0\right)\), which is not a risk-neutral probability. This means that the enlarged market is not arbitrage-free. To find this opportunity we need to construct a portfolio which only pays off in states \(\omega_2\) and \(\omega_4\). For example, the strategy consisting in the purchase of a stock funded from a bank account and simultaneous sale of a put option is an arbitrage opportunity.

### 2.3 Complete markets and pricing of attainable claims

In this and the following section we assume that the market is arbitrage-free and hence there exists a risk-neutral probability \(\mathbb{Q}\).

**Definition 2.3** A contingent claim \(B\) is a random variable on \((\Omega, \mathcal{F}, \mathbb{P})\).

In our simple setting, a random variable is simply a vector of values taken in different states of nature, and the contingent claim is therefore identified with the vector of its pay-off values in different states.
A contingent claim $B$ is said to be attainable if there exists a strategy $\theta$ such that $B = X_1^\theta$. By the absence of arbitrage, the price of the contingent claim at time 0 must then be given by $p(B) = X_0^\theta$. Indeed, if the contingent claim is quoted at a price $p' < p(B)$, then a strategy which consists in (i) buying the contingent claim at $t = 0$; (ii) selling the portfolio $X^\theta$, has a zero terminal pay-off and generates an immediate profit of $p(B) - p'$. Similarly, if the quoted price is greater than $p(B)$, an arbitrage strategy can be constructed.

Since, under a risk-neutral probability $Q$, $X_0 = E^Q[\tilde{V}_1]$, we have the following result:

**Proposition 2.1** The fair price $p(B)$ of an attainable contingent claim satisfies

$$p(B) = E^Q[\tilde{B}] = E^Q \left[ \frac{B}{S_1^0} \right],$$

where $Q$ is any risk-neutral probability in the market.

**Example 2.8** In the market of example 2.6, a put option with strike $K = 1$ is attainable (using the put-call parity) and one can check that its expected discounted pay-off is equal to $\frac{1}{10}$ under all the risk-neutral probabilities in this market. The put option with strike $K = \frac{3}{2}$ is not attainable.

**Definition 2.4** A market is said to be complete if every contingent claim is attainable.

Let $A$ be the pay-off matrix: $A_{ik} = S_i^1(\omega_k)$. The contingent claim $B$ is attainable if the equation $\theta^T A = B$ has a solution $\theta \in \mathbb{R}^{d+1}$. Every contingent claim is attainable if $\theta^T A = B$ has a solution for every $B \in \mathbb{R}^K$, which is the case if $A$ has $K$ linearly independent lines. This means that for a market to be complete, the number of assets (including the risk-free asset) must be greater or equal to the number of states of nature.

**Example 2.9** If the market of example 2.6 is enlarged with a put option with strike $K = \frac{3}{2}$, it becomes complete (four linearly independent assets). If it is enlarged with a put option with strike $K = 1$, it remains incomplete (four assets but not linearly independent).
**Theorem 2.2** An arbitrage-free market is complete if and only if the risk-neutral probability is unique.

**Proof.** Suppose that the market is complete. Then the prices of the contingent claims of the form $1_{\omega_k}$ for $k = 1, \ldots, K$ are uniquely defined and for every risk-neutral probability $Q$ we have

$$Q(\omega_k) = S^0_t p(1_{\omega_i}).$$

Suppose that the risk-neutral probability $Q$ is unique. We will show that the pay-off matrix $A$ has $K$ linearly independent lines. If this is not the case, there exists $z \in \mathbb{R}^K$ with $z \neq 0$ and $Az = 0$. Then for $\varepsilon$ sufficiently small, $Q + \varepsilon z$ is a risk-neutral probability different from $Q$. \hfill \Box

### 2.4 Asset pricing in incomplete markets

In incomplete markets, the fair price of some contingent claims is not unique, but one can nevertheless compute bounds on their prices: the upper and lower superhedging price, defined by

$$p^+(B) = \inf \{ p(C) | C \text{ attainable, } C \geq B \},$$

$$p^-(B) = \sup \{ p(C) | C \text{ attainable, } C \leq B \}.$$

In our finite-space model, we actually have

$$p^+(B) = \min \{ p(C) | C \text{ attainable, } C \geq B \},$$

$$p^-(B) = \max \{ p(C) | C \text{ attainable, } C \leq B \}.$$

For example, in the case of $p^+$, we can always restrict the set of attainable claims over which we minimize to those satisfying $C(\omega_i)Q(\omega_i) \leq \max_i B(\omega_i)$ (where $Q$ is any risk-neutral measure). Since the latter set is compact, the inf is attained.

If the claim $B$ is attainable, $p^+(B) = p^-(B)$ but for genuinely inattainable claims $p^-(B) < p(B) < p^+(B)$ because equality would create arbitrage.
Theorem 2.3 In an arbitrage-free market, the upper and lower superhedging prices are given by

\[ p^-(B) = \inf_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}] \]
\[ p^+(B) = \sup_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}], \]

where \( \mathcal{Q}^{RN} \) is the set of all risk-neutral probabilities.

Proof. Assume first that \( p^+(B) < \sup_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}] \). Then there exists \( Q \in \mathcal{Q}^{RN} \) with \( p^+(B) < E^Q[\tilde{B}] \) and an attainable claim \( C \) with \( p(C) < E^Q[\tilde{B}] \) and \( C \geq B \). Then \( E^Q[\tilde{C}] < E^Q[\tilde{B}] \) which is in contradiction with \( C \geq B \). Therefore \( p^+(B) \geq \sup_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}] \). Similarly, we can show that \( p^-(B) \leq \inf_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}] \).

Assume now that \( p^-(B) < \inf_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}] \) and consider an extended market where \( B \) is quoted at price \( p(B) \) with \( p^-(B) < p(B) < \inf_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}] \). By theorem 2.1, in this market there exists an arbitrage opportunity, which will be denoted by \((\theta, \theta')\), where \( \theta' \neq 0 \) is the coefficient in front of \( B \). If \( \theta' > 0 \) then \( B \) dominates the attainable portfolio \(-\frac{1}{\theta'} \theta S\), whose price is equal to \( p(B) \) (by definition of the arbitrage opportunity) and hence greater than \( p^-(B) \), which gives a contradiction.

If \( \theta' < 0 \) then \( B \) is dominated by the attainable portfolio \(-\frac{1}{\theta'} \theta S\) whose price is equal to \( p(B) \). Therefore, \( p^+(B) < \inf_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}] \), which gives a contradiction with the first part of the proof.

Therefore, we conclude that \( p^-(B) \geq \inf_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}] \) and in the same way we can show that \( p^+(B) \leq \sup_{Q \in \mathcal{Q}^{RN}} E^Q[\tilde{B}] \). ♦

Example 2.10 Let us compute the arbitrage bounds for a put option with strike \( K = \frac{3}{2} \) in example 2.6. The set \( \{E^Q[\tilde{B}]|Q \in \mathcal{Q}^{RN}\} \) coincides with \( \left\{ \frac{1}{5} + \frac{2}{5} q|0 < q < \frac{3}{5}\right\} \), and the bounds are therefore \( \frac{1}{5} < p(B) < \frac{7}{20} \).
2.5 Portfolio optimization and equilibrium asset pricing

In this section, we consider the problem of an economic agent who disposes of \( w \) euros at date \( t = 0 \) and wants to invest them optimally using the available assets. Whereas the price of a contingent claim in a complete arbitrage-free market is uniquely determined, the optimal investment policy depends on the preferences of the agent: some agents will choose riskier portfolios than others.

To measure the preferences of an agent, we introduce the utility function: a function \( u : \mathbb{R} \rightarrow \mathbb{R} \) or \( u : [0, \infty) \rightarrow \mathbb{R} \), which is strictly increasing and strictly concave. The agent prefers holding portfolio \( X \) at date \( t = 1 \) to holding the portfolio \( Y \) if and only if \( E[u(X)] \geq E[u(Y)] \). The fact that \( u \) is increasing means that \( X \) is always preferred to \( Y \) if \( X \geq Y \), and the concavity implies (via Jensen’s inequality) that \( E[X] \) is always preferred to \( X \). Since increasing linear transformations of the utility function do not change the structure of preferences, the utility function is only defined up to such transformations.

For a random financial flow \( X \), the certainty equivalent of \( X \) is the constant \( c_X \in \mathbb{R} \) such that \( u(c_X) = E[u(X)] \). The risk premium is the difference between the expected value of \( X \) and its certainty equivalent: \( \rho(X) = E[X] - c_X \). Let \( m = E[X] \). Then \( u(c_X) = u(m - \rho) \approx u(m) - u'(m)\rho \). On the other hand, \( u(c_X) = E[u(X)] \approx u(m) + \frac{1}{2} u''(m) \text{Var}(X) \). Therefore,

\[
\rho \approx \alpha(m) \frac{\text{Var}(X)}{2},
\]

where \( \alpha(x) = \frac{u''(x)}{u'(x)} \) is called the absolute risk aversion coefficient at the level \( x \).

**Examples of utility functions** Setting \( \alpha(x) = \alpha \) for all \( x \), we obtain the constant absolute risk aversion (CARA) utility defined for \( x \in \mathbb{R} \):

\[
(- \log u')' = \alpha \quad \Rightarrow \quad u' = ce^{-\alpha x} \quad \Rightarrow \quad u(x) = 1 - e^{-\alpha x}.
\]
Setting $\alpha = \frac{1-\gamma}{x}$ for $\gamma \in [0,1)$ leads to hyperbolic absolute risk aversion (HARA) utility function, defined for $x \in [0,\infty)$, which models the situation where the agent’s risk aversion decays with increasing wealth. We get:

$$( - \log u' )' = \frac{1-\gamma}{x} \Rightarrow u' = Cx^{\gamma-1}.$$ 

If $\gamma = 0$, this leads to the logarithmic utility $u(x) = \log x$, and the case $\gamma > 0$ corresponds to power utility $u(x) = \frac{x^\gamma}{\gamma}$. One checks easily that

$$\lim_{\gamma \to 0} \frac{x^\gamma}{\gamma} = \log x.$$ 

2.5.1 The portfolio optimization problem

In this section, we are interested in the following problem:

$$\max_{\theta} E[u(X_1^\theta)] \quad (2.1)$$

subject to the constraint $X_0^\theta = w$, called the budget constraint. First, let us observe that if the market is not arbitrage-free, this problem does not have a (finite) solution. Indeed, let $\hat{\theta}$ be an optimal strategy with $E[u(X_1^{\hat{\theta}})] < \infty$ and $\theta_A$ be an arbitrage opportunity. Then the strategy $\hat{\theta} + \theta_A$ satisfies the budget constraint and

$$E[u(X_1^{\hat{\theta} + \theta_A})] > E[u(X_1^{\hat{\theta}})],$$

which means that $\hat{\theta}$ cannot be optimal. Therefore, from now on, we suppose that the market is arbitrage-free. This means that there exits a risk-neutral probability $Q$ such that

$$E^Q \left[ \frac{X_1^{\hat{\theta}}}{(1+R)} \right] = w \quad (2.2)$$

The budget constraint can therefore be replaced by the constraint (2.2), and the utility maximization problem reduces to two independent problems:
• Maximize the utility over the set of all hedgeable claims subject to the budget constraint:

\[
\max_{X \text{ hedgeable}} E[u(X)] \quad \text{subject to} \quad E^Q[X] = (1 + R)w.
\]

Let \( \bar{X} \) be the claim which realizes the maximum.

• Find a trading strategy \( \theta \) which replicates \( \bar{X} \). This is a problem that we already know how to solve from previous sections.

At this point, two remarks are in order. First, if the market is complete, all claims are hedgeable, and the first problem reduces to

\[
\max E[u(X)] \quad \text{subject to} \quad E^Q[X] = (1 + R)w. \quad (2.3)
\]

Secondly, exactly the same logic can be used to solve the portfolio optimization problem in a multiperiod discrete-time model.

In view of the above discussion, from now on we concentrate on the problem (2.3), but without necessarily supposing that the market is complete. If it is, this problem gives us an immediate solution to the portfolio optimization problem. In the incomplete market case, it will allow us to understand the optimal contingent claim profiles and explain the demand for derivative products in financial markets. The problem of finding an optimal portfolio in an incomplete market is more difficult and out of scope of this introductory course.

2.5.2 Finding the optimal contingent claim

In this section, in addition to what was assumed before, we suppose that the utility function \( u \) satisfies one of the two sets of conditions (called Inada conditions):

(A1) \( u : \mathbb{R} \to \mathbb{R} \), is continuously differentiable on \( \mathbb{R} \), and satisfies \( \lim_{x \to \infty} u'(x) = 0 \) and \( \lim_{x \to -\infty} u'(x) = +\infty \) (the typical example is exponential utility).
(A2) \( u : [0, \infty) \to \mathbb{R} \), is continuously differentiable on \((0, \infty)\), and satisfies 
\[
\lim_{x \to \infty} u'(x) = 0 \quad \text{and} \quad \lim_{x \to 0^+} u'(x) = +\infty \quad \text{(the typical example is power utility)}.
\]

We consider a possibly incomplete arbitrage-free financial market with \( K \) states of nature and associated probabilities \((p_i)_{i=1}^{K}\). Let \( Q = (q_i)_{i=1}^{K} \) be a risk-neutral probability. The problem (2.3) becomes
\[
\max_{x \in \mathbb{R}^K} \sum_{i=1}^{K} p_i u(x_i) \quad \text{subject to} \quad \sum_{i=1}^{K} q_i x_i = w(1 + R).
\]

To solve it, we construct the Lagrangian:
\[
L(x_1, \ldots, x_K, \lambda) = \sum_{i=1}^{K} p_i u(x_i) - \lambda \left\{ \sum_{i=1}^{K} q_i x_i - w(1 + R) \right\}
\]

The first order conditions are
\[
p_i u'(x_i) = \lambda q_i.
\]

This gives us a potential candidate for the optimal claim:
\[
x_i = (u')^{-1} \left( \lambda \frac{q_i}{p_i} \right) := I \left( \lambda \frac{q_i}{p_i} \right),
\]

which makes sense as long as \( \lambda > 0 \) due to the assumptions in the beginning of this section. The value of \( \lambda \) is found from the budget equation:
\[
\sum_{i=1}^{K} q_i I \left( \lambda \frac{q_i}{p_i} \right) = w(1 + R) \quad (2.4)
\]

The function
\[
f(\lambda) = \sum_{i=1}^{K} q_i I \left( \lambda \frac{q_i}{p_i} \right)
\]
is strictly decreasing and satisfies
\[
\begin{align*}
\lim_{\lambda \to 0} f(\lambda) &= +\infty \quad \text{and} \quad \lim_{\lambda \to \infty} f(\lambda) = -\infty \quad \text{under (A1)} \\
\lim_{\lambda \to 0^+} f(\lambda) &= +\infty \quad \text{and} \quad \lim_{\lambda \to \infty} f(\lambda) = 0 \quad \text{under (A2)}
\end{align*}
\]
Therefore, under (A1), for every initial wealth $w \in \mathbb{R}$ there exists a solution $\lambda^* > 0$ of equation (2.4), and under (A2), such a solution exists for every positive initial wealth $w > 0$. The optimal claim is then given by

$$x^*_i = I \left( \frac{\lambda^* q_i}{p_i} \right) \quad \text{or} \quad X^* = I(\lambda^* Z) \quad \text{with} \quad Z := \frac{dQ}{dP}.$$

To show that $X^*$ is indeed the unique optimizer of the constrained problem (2.3), assume that $X' = (x'_1, \ldots, x'_K)$ is another solution, different from $X^*$. We then have:

$$\sum_{i=1}^K p_i u(x'_i) = L(x'_1, \ldots, x'_K, \lambda^*) < L(x^*_1, \ldots, x^*_K, \lambda^*) = \sum_{i=1}^K p_i u(x^*_i),$$

where the first equality holds because $X'$ satisfies the budget equation, the inequality is due to the strict concavity of $L$ and the last equality follows from the budget equation for $X^*$.

**Example: exponential utility**  Let $u(x) = 1 - e^{-\alpha x}$. Then

$$I(z) = -\frac{1}{\alpha} \log \left( \frac{z}{\alpha} \right)$$

and the budget equation can be solved explicitly:

$$E_Q \left[ -\frac{1}{\alpha} \log \left( \frac{\lambda^* Z}{\alpha} \right) \right] = w(1 + R)$$

$$\Rightarrow \quad X^* = w(1 + R) + \frac{1}{\alpha} \left\{ E_Q[\log Z] - \log Z \right\}$$

Therefore, in the exponential utility framework, every investor will want to hold a share of the same risky claim with pay-off $M = \{ E_Q[\log Z] - \log Z \}$. The amount to buy depends on the risk aversion of a particular investor, but the claim itself is the same for all investors. Therefore, the aggregate portfolio of all investors will be proportional to $M$, and we recover the *mutual fund theorem*. Note also that if $Q \neq P$, by Jensen’s inequality,

$$E^P[M] = E^P[(Z - 1) \log Z] > 0$$

which means that the agents are rewarded for the risk they take.
2.5.3 Equilibrium asset pricing

In this section, our aim is to understand how asset prices can be determined endogeneously via a supply-demand equilibrium among market participants. We suppose that the market consists of a finite set $A$ of economic agents with utility functions $(u_a)_{a \in A}$ and initial endowments with pay-offs $(W_a)_{a \in A}$. The sum of all endowments $M := \sum_{a \in A} W_a$ is the market portfolio, that is, the aggregate portfolio of all assets available for trading. A feasible allocation $(X_a)_{a \in A}$ is a set of contingent claims satisfying the market clearing condition:

$$\sum_{a \in A} X_a = M.$$ 

The prices in the market will be determined via a pricing density $\phi$: a random variable $\phi > 0$ satisfying $E[\phi] = 1$. The price of a contingent claim $X$ is $p(X) = E[\phi X]$. 

A microeconomic or Arrow-Debreu equilibrium is a pricing density $\phi^*$ and a feasible allocation $(X^*_a)_{a \in A}$, such that for every $a \in A$, $X_a$ solves the utility maximization problem for agent $a$ subject to the budget constraint defined by the pricing rule $\phi^*$: $E[\phi^* X_a] = E[\phi^* W_a]$. 

In these introductory lecture notes we consider the simplest case where all agents have an exponential utility function: $u_a(x) = 1 - e^{-\alpha_a x}$. Let $\phi$ be a pricing density. Then the optimal allocation for agent $a$ has the form

$$X_a = E[\phi W_a] + \frac{1}{\alpha_a} \{E[\phi \log \phi] - \log \phi\},$$

and the market clearing condition becomes

$$M = E[\phi M] + \frac{1}{\alpha} \{E[\phi \log \phi] - \log \phi\}$$

where we set $\frac{1}{\alpha} := \sum_{a \in A} \frac{1}{\alpha_a}$. This implies that equilibrium pricing rule is necessarily of the form

$$\phi^* = \frac{e^{-\alpha M}}{E[e^{-\alpha M}]}.$$
with this pricing rule, the optimal allocation for each agent

$$X^*_a = E[\phi^*W_a] + \frac{\alpha}{\alpha_a}(M - E[\phi^*M])$$

evidently satisfies the market clearing condition. The couple \((\phi^*, X^*_a)\) is then the unique equilibrium in this market. In equilibrium, the agents hold linear shares of the market portfolio, which means that once again we recover the mutual fund theorem.
Chapter 3

The finite probability multiperiod model

3.1 The model

In this chapter we consider a discrete-time financial model with \( T \) periods. The randomness is once again represented by a finite probability space \((\Omega, \mathcal{F}, \mathbb{P})\), but now we need to specify how the information is revealed through time. This is described via the notion of filtration: an increasing sequence of \( \sigma \)-fields \( \mathcal{F} = (\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_T) \). The \( \sigma \)-field \( \mathcal{F}_t \) contains all events which are known at date \( t \). We assume that \( \mathcal{F}_0 = (\emptyset, \Omega) \) and \( \mathcal{T}_T = \mathcal{F} = 2^\Omega \): no information is available at date 0 and everything is known at date \( T \).

The risky asset prices are modelled by \( d \) discrete-time stochastic processes \((S_i^t)_{t=0,1,\ldots,T}^d\) which are \( \mathcal{F} \)-adapted, that is, \( S_i^t \) is \( \mathcal{F}_t \) measurable for every \( t \). This means that time-\( t \) prices of risky assets are revealed at date \( t \).

The risk-free bank account process \((S_0^0)_{t=0,1,\ldots,T}^d\) is defined by \( S_0^0 = 1 \) and \( S_i^0 = S_{i-1}^0(1 + R_t) \), where \( R_t \) is the risk-free interest rate for the period between \( t - 1 \) and \( t \). The process \((R_t)_{1\leq t\leq T}^d\) is assumed to be \( \mathcal{F} \)-predictable, meaning that \( R_t \) is \( \mathcal{F}_{t-1} \)-measurable for every \( t = 1, \ldots, T \): the risk-free rate for the period is fixed in the beginning of the period.
**Trading strategies**  The trading strategy is a $d + 1$-dimensional stochastic process $(\theta^i_t)_{t=1,\ldots,T}$, where $\theta^i_t$ represents the amount of $i$-th asset held in the portfolio between dates $t_{i-1}$ and $t_i$. The trading strategies are also assumed to be predictable: the portfolio composition for the period is known in the beginning of the period.

For example, a constant strategy $\theta^i_t = \theta^i_0$ corresponds to buying the assets at $t = 0$ and holding them until $t = T$. In the fixed-mix strategy, one fixes the proportions $\omega_0, \ldots, \omega_d$ of the total wealth to invest into different assets, and at each date the portfolio is readjusted to maintain the same proportions. Since the asset prices do not vary proportionally, the amounts of assets will be different at each date for this strategy.

The *gain process* of a strategy is the stochastic process describing the change in the portfolio value due to the variation of the asset prices:

$$G_0 = 0, \quad G_t = \sum_{s=1}^{t} \sum_{i=0}^{d} \theta^i_s \Delta S^i_s, \quad \Delta S^i_t = S^i_t - S^i_{t-1}.$$

**Self-financing strategies**  The value of a portfolio may change throughout time due to two factors:

- Variation of asset prices;
- Adding/withdrawal of funds during transactions.

A *self-financing* portfolio strategy is a strategy where no funds are added to/withdrawn from the portfolio after the initial date, and the change in portfolio value is therefore only due to variations of asset prices.

The portfolio value at date $t$ just before the transaction can be written as

$$\sum_{i=0}^{d} \theta^i_t S^i_t,$$

and that after the transaction is given by

$$\sum_{i=0}^{d} \theta^{i+1} S^i_t.$$
For a self-financing strategy these two values are equal and the value of portfolio at date \( t \) is defined uniquely by

\[
X_t = \sum_{i=0}^{d} \theta_i^t S_i^t = X_0 + G_t.
\]

**Discounting** Similarly to the one-period model, we define

\[
\tilde{S}_i^t = \frac{S_i^t}{S_0^t}, \quad \tilde{X}_i^t = \frac{X_i^t}{S_0^t} = \tilde{X}_0 + \tilde{G}_t,
\]

where

\[
\tilde{G}_t = \sum_{s=1}^{t} \sum_{i=1}^{d} \theta_s^i \Delta \tilde{S}_i^s.
\]

The discounting allows once again to eliminate all dependency of the gain process on the risky asset.

### 3.2 Arbitrage-free markets

**Definition 3.5** An arbitrage opportunity is a self-financing portfolio strategy \( \theta \) such that

\[
X_0^\theta = 0, \quad \mathbb{P}[X_T^\theta \geq 0] = 1 \quad \text{and} \quad \mathbb{P}[X_T^\theta > 0] > 0.
\]

Equivalently, an arbitrage opportunity is a strategy with \( \mathbb{P}[\tilde{G}_T^\theta \geq 0] = 1 \) and \( \mathbb{P}[\tilde{G}_T^\theta > 0] > 0 \). In multi-period models, in addition to static arbitrage opportunities, such as violation of the put-call parity, there may exist dynamic arbitrage opportunities, which cannot be realized by buy-and-hold strategies.

Arbitrage-free multiperiod markets can be characterized via the existence of a martingale probability, which is a generalization of the notion of risk-neutral probability from the one-period setting.

**Theorem 3.4** The \( T \)-period financial market is arbitrage-free if and only if there exists a probability \( \mathbb{Q} \) such that \( \mathbb{Q}(\omega) > 0 \) for all \( \omega \in \Omega \) and the discounted prices of all assets are \( \mathbb{Q} \)-martingales:

\[
E^Q[\tilde{S}_{t+1} | \mathcal{F}_t] = \tilde{S}_t, \quad i = 0, \ldots, d, \quad t = 0, \ldots, T - 1.
\]
Proof. Let $Q$ be a martingale probability. Then the discounted value $(\tilde{X}_t)$ of every self-financing portfolio is a $Q$-martingale, so that for every strategy $\theta$ with $X^0_0 = 0$ we have $E^Q[\tilde{X}_T^\theta] = 0$, which is in contradiction with $V^\theta_T \geq 0$ in all states and $V^\theta_T > 0$ in at least one state.

Conversely, assume that the market is arbitrage-free. We consider a one-period market with the same states of nature $\Omega$ as the original one, with one risk-free asset evolving with zero interest rate, and with the risky assets having pay-offs of the form

$$1_{A_t}(\tilde{S}^i_t - \tilde{S}^i_{t-1}) \quad \forall t = 1, \ldots, T, \quad \forall i = 1, \ldots, d, \quad \forall A_t \in \mathcal{F}_{t-1},$$

and quoted all at zero price at the initial date. Since $\Omega$ is finite, there is only a finite number of such risky assets. This number will be denoted by $D$. For $j = 1, \ldots, D$, we denote the pay-off of the $j$-th asset by $1_{A(j)}(\tilde{S}^j_{t(j)} - \tilde{S}^j_{t(j)-1})$

We claim that the new market is also arbitrage-free. Indeed, assume that $w \in \mathbb{R}^{D+1}$ is an arbitrage strategy in this market. The discounted gain $G$ associated to this strategy can be written as

$$G = \sum_{j=1}^D w_j 1_{A(j)}(\tilde{S}^j_{t(j)} - \tilde{S}^j_{t(j)-1}) = \sum_{s=1}^T \sum_{i=1}^d \sum_{j=1}^D w_j 1_{t(j)=s} 1_{A(j)}(\tilde{S}^i_s - \tilde{S}^i_{s-1})$$

$$= \sum_{s=1}^T \sum_{i=1}^d \theta^i_s (\tilde{S}^i_s - \tilde{S}^i_{s-1}),$$

where

$$\theta^i_s = \sum_{j=1}^D w_j 1_{t(j)=s} 1_{A(j)} \in \mathcal{F}_{s-1}.$$

Therefore, $G$ is the discounted gain of a self-financing strategy in the original market, and since the original market is arbitrage free, it is not possible that $P[G \geq 0] = 1$ and $P[G > 0] > 0$.

From Theorem 2.1 we then know that there exists a probability $Q$ on $\Omega$ which satisfies $Q(\omega) > 0$ for all $\omega \in \Omega$ and

$$E^Q \left[ 1_{A_t}(\tilde{S}^i_t - \tilde{S}^i_{t-1}) \right] = 0$$

34
for all $A_t \in \mathcal{F}_{t-1}$, which implies that

$$E^Q[\tilde{S}_t^{i} | \mathcal{F}_{t-1}] = \tilde{S}_t^{i-1},$$

and therefore $Q$ is a martingale probability for the original market.

\[\Diamond\]

### 3.3 Evaluation of contingent claims

A European contingent claim is a $\mathcal{F}_T$-measurable random variable, which represents the pay-off of an option at time $T$. For example, a European call option has pay-off $B = (S_T - K)^+$, where $S$ is the price process of the underlying asset; an Asian call option has pay-off $B = \left(\frac{1}{T} \sum_{t=0}^{T} S_t - K\right)^+$. A contingent claim $B$ is said to be attainable or hedgeable if there exists a self-financing strategy $\theta$ such that $X_0^\theta = B$. By arbitrage arguments, similar to the ones used in the previous chapter, the price of $B$ at any date $t$ is then given by $p_t(B) = X_t^\theta$. Since under any risk-neutral probability $Q$, $\tilde{X}_t^\theta$ is a martingale, we obtain the following result:

**Proposition 3.2** The fair price $p_t(B)$ of an attainable contingent claim at date $t$ satisfies

$$p_t(B) = S_0^t E^Q \left[ \frac{B}{S_T} \right]_{\mathcal{F}_t},$$

where $Q$ is any martingale probability.

The price of a non-attainable contingent claim is not uniquely defined, but one can, as in the one-period model, establish arbitrage bounds for this price in terms of the set of risk-neutral probabilities in the market:

$$\inf_{Q \in \mathcal{Q}^M} S_0^t E^Q \left[ \frac{B}{S_T} \right]_{\mathcal{F}_t} \leq p_t(B) \leq \sup_{Q \in \mathcal{Q}^M} S_0^t E^Q \left[ \frac{B}{S_T} \right]_{\mathcal{F}_t},$$

where $\mathcal{Q}^RN$ is the set of all martingale probabilities.
3.4 Complete markets

As in the one-period model, a market is said to be complete if every contingent claim is attainable. Moreover, the same characterization in terms of the uniqueness of the martingale probability holds.

**Theorem 3.5** An arbitrage-free market is complete if and only if the martingale probability is unique.

**Proof.** The only if part. Assume that the market is complete. This means that every contingent claim is attainable and has a unique price, in particular the claims $S^0_t \cdot 1_{\omega_i}$ for every $i$, which shows the uniqueness of the martingale probability.

The if part. By way of contradiction, assume that there exists a non-attainable contingent claim. We consider a one-period market with the same states of nature $\Omega$ as the original one, with the one risk-free asset evolving with zero interest rate, and with the risky assets having pay-offs of the form

$$1_{A_t}(\tilde{S}_t^i - \tilde{S}_{t-1}^i) \quad \forall t = 1, \ldots, T, \quad \forall i = 1, \ldots, d, \quad \forall A_t \in \mathcal{F}_{t-1},$$

and quoted all at zero price at the initial date. Since $\Omega$ is finite, there is only a finite number of such risky assets. In this one-period market there is also a non-attainable claim, and by theorem 2.2, there exists a probability $Q^*$ on $\Omega$ which is different from $Q$ and satisfies

$$E^{Q^*}[1_{A_t}(\tilde{S}_t^i - \tilde{S}_{t-1}^i)] = 0$$

for all $A_t \in \mathcal{F}_{t-1}$, which implies that

$$E^{Q^*}[\tilde{S}_t^i | \mathcal{F}_{t-1}] = \tilde{S}_{t-1}^i,$$

and therefore $Q^*$ is a martingale probability for the original market, different from $Q$. 

$\diamond$
3.5 On discrete-time martingales

In this section, we recall some results from the theory of discrete-time martingales.

**Definition 3.6** A $\mathcal{F}$-adapted process $M = \{M_t, t = 0, \ldots, T\}$ is an $(\mathcal{F}, \mathbb{P})$—supermartingale (resp. $(\mathcal{F}, \mathbb{P})$—submartingale) if $M_t$ is $\mathbb{P}$-integrable for every $t$ and

$$
\mathbb{E}[M_t | \mathcal{F}_{t-1}] \leq \text{(resp. } \geq \text{)} M_{t-1} \quad \text{for every } t = 1, \ldots, T.
$$

A process $M$ is an $(\mathcal{F}, \mathbb{P})$—martingale if it is both an $(\mathcal{F}, \mathbb{P})$—supermartingale and an $(\mathcal{F}, \mathbb{P})$—submartingale.

**Example 3.11** Let $M$ be an $(\mathcal{F}, \mathbb{P})$—martingale and $\phi$ an $\mathcal{F}$—adapted bounded process. Then, the process $Y$ defined by:

$$
Y_0 = 0 \quad \text{and} \quad Y_t = \sum_{k=1}^{t} \phi_{k-1}(M_k - M_{k-1}), \quad t = 1, \ldots, T
$$

is an $(\mathcal{F}, \mathbb{P})$—martingale.

This example allows to build a large family of $(\mathcal{F}, \mathbb{P})$—martingales out of some given $(\mathcal{F}, \mathbb{P})$—martingale $M$. Notice however that the process $\phi$ needs to satisfy a convenient restriction, namely $\phi$ has to be essentially bounded. Without this restriction, we do not even have the guarantee that $Y$ is integrable, and we are led to the notion of local martingales; see Definition 3.8 below.

**Definition 3.7** An $\mathcal{F}$—stopping time $\nu$ is a random variable with values in $\{0, \ldots, T\}$ such that

$$
\{\nu = t\} \in \mathcal{F}_t \quad \text{for every } t = 0, \ldots, T.
$$

We shall denote by $\mathcal{T}$ the set of all $\mathcal{F}$—stopping times.
Remark 3.1 An equivalent characterization of an $F$–stopping time is that
\[ \{ \nu \leq t \} \in F_t \text{ for every } t = 0, \ldots, T, \]
(exercise !). This definition extends to the continuous-time setting.

Let $Y = \{ Y_t, t = 0, \ldots, T \}$ be an $F$-adapted process. The stopped process $Y^\nu$ at the $F$–stopping time $\nu$ is defined by :
\[ Y_t^\nu := Y_{t\wedge \nu} \text{ for } t = 0, \ldots, T. \]

Exercise 3.2 Let $Y$ be an $F$-adapted process and consider some $F$–stopping time $\nu \in \mathcal{T}$.
1. Prove that the stopped process $Y^\nu$ is $F$-adapted.
2. If $Y$ is an $(F, \mathbb{P})$–martingale (resp. $(F, \mathbb{P})$–supermartingale), prove that $Y^\nu$ an $(F, \mathbb{P})$–martingale (resp. $(F, \mathbb{P})$–supermartingale).

For every stopping time $\nu$, the available information at time $\nu$ is defined by :
\[ F_\nu := \{ A \in F : A \cap \{ \nu = t \} \in F_t \text{ for all } t = 0, \ldots, T \}. \]

We recall the following important result on stopped martingales.

Theorem 3.6 (Optional sampling, Doob). Let $M$ be a martingale (resp. supermartingale), and consider two stopping times $\nu, \bar{\nu} \in \mathcal{T}$ with $\nu \leq \bar{\nu}$ a.s. Then :
\[ \mathbb{E}[M_\nu | F_\nu] = M_\nu \text{ (resp. } \leq M_\nu \text{) } \mathbb{P} - \text{a.s.} \]

Proof. We only prove this result for martingales ; the case of supermartingales is treated similarly.
(i) We first show that $\mathbb{E}[M_T | F_\nu] = M_\nu$ for every $\nu \in \mathcal{T}$. Let $A$ be an arbitrary event set in $F_\nu$. Then, the event set $A \cap \{ \nu = t \} \in F_t$ and therefore :
\[ \mathbb{E} \left[ (M_T - M_\nu) 1_{A \cap \{ \nu = t \}} \right] = \mathbb{E} \left[ (M_T - M_t) 1_{A \cap \{ \nu = t \}} \right] = 0 \]
since $M$ is a martingale. Summing up these inequalities, it follows that:

$$0 = \sum_{t=0}^{T} \mathbb{E} \left[ (M_T - M_\nu)1_{\mathcal{A}\cap\{\nu=t\}} \right] = \mathbb{E} \left[ (M_T - M_\nu)1_{\mathcal{A}} \right].$$

Since $\mathcal{A}$ is arbitrary in $\mathcal{F}_\nu$, this proves that $\mathbb{E}[M_T - M_\nu|\mathcal{F}_\nu] = 0$.

(ii) By the second question of Exercise 3.2, the stopped process $M^\nu$ is a martingale. Then, by the above step (i), we see that:

$$\mathbb{E} [M_\nu | \mathcal{F}_\nu] = \mathbb{E} [M^\nu_T | \mathcal{F}_\nu] = M^\nu_\nu = M_\nu$$
since $\nu \leq \bar{\nu}$.

We next introduce the notion of local martingales.

**Definition 3.8** A process $M = \{M_t, t = 0, \ldots, T\}$ is said to be a local martingale if there exists a sequence of $(\mathcal{F}, \mathbb{P})$—stopping times $(\tau_n)_n$ such that $\tau_n \rightarrow \infty$ $P$-a.s. and the stopped process $M^{\tau_n}$ is an $(\mathcal{F}, \mathbb{P})$—martingale for every $n$.

**Example 3.12** Let $M$ be an $(\mathcal{F}, \mathbb{P})$—martingale and $\phi$ an arbitrary $\mathcal{F}$—adapted process. Then, the process $Y$ defined by:

$$Y_0 := 0 \quad \text{and} \quad Y_t := \sum_{k=1}^{t} \phi_{t-1}(M_k - M_{k-1}), \quad t = 1, \ldots, T$$

is a local martingale. To see this, consider the sequence of stopping times

$$\tau_n := \inf \{t : |\phi_t| \geq n\},$$

with the convention $\inf \emptyset = \infty$. Then, for every $n \geq 1$,

$$Y_t^{\tau_n} = \sum_{k=1}^{t} \tilde{\phi}_{t-1}(M_k - M_{k-1}) \quad \text{where} \quad \tilde{\phi}_t := \phi_t 1_{\{t < \tau_n\}}, \quad t = 1, \ldots, T.$$ 

Since $\tilde{\phi}$ is an $\mathcal{F}$—measurable bounded process, we are reduced to the context of Example 3.11, and the stopped process $Y^{\tau_n}$ is an $(\mathcal{F}, \mathbb{P})$—martingale.
The following result provides a sufficient condition in finite discrete time for a local martingale to be a martingale.

**Lemma 3.1** Let \( M = \{M_t, t = 0, \ldots, T\} \) be an \((\mathcal{F}, \mathbb{P})\)-local martingale with:
\[
\mathbb{E}[M_T^-] < \infty.
\]
Then, \( M \) is an \((\mathcal{F}, \mathbb{P})\)-martingale.

**Proof.** Let \((\tau_n)_n\) be a sequence of stopping times such that \( \tau_n \to \infty \) a.s. and the stopped process \( M^{\tau_n} \) is an \((\mathcal{F}, \mathbb{P})\)-martingale for every \( n \). We organize the proof into two steps.

1. We first prove that:
\[
\mathbb{E}[|M_t|^1_{\mathcal{F}_{t-1}}] < \infty \quad \text{and} \quad \mathbb{E}[M_t|\mathcal{F}_{t-1}] = M_{t-1} \quad \mathbb{P} \text{-a.s.} \quad 1 \leq t \leq T. \tag{3.1}
\]
Since \( \tau_n \) is an \((\mathcal{F}, \mathbb{P})\)-stopping time, we deduce that the event set \( \{\tau_n > t-1\} \) is \( \mathcal{F}_{t-1} \)-measurable. Then:
\[
\mathbb{E}[|M_t|_{\mathcal{F}_{t-1}}] = \mathbb{E}[|M^{\tau_n}_t|^1_{\mathcal{F}_{t-1}}] < \infty \quad \mathbb{P} \text{-a.s. on } \{\tau_n > t-1\},
\]
where the last inequality follows from the fact that \( M^{\tau_n} \) is \( \mathbb{P} \)-integrable. By sending \( n \) to infinity, we deduce that \( \mathbb{E}[|M_t|_{\mathcal{F}_{t-1}}] < \infty \) \( \mathbb{P} \)-a.s.

Similarly, using the martingale property of the stopped process \( M^{\tau_n} \), we see that:
\[
\mathbb{E}[M_t|\mathcal{F}_{t-1}] = \mathbb{E}[M^{\tau_n}_t|\mathcal{F}_{t-1}] = M^{\tau_n}_{t-1} = M_{t-1} \quad \text{on } \{\tau_n > t-1\},
\]
which implies that \( \mathbb{E}[M_t|\mathcal{F}_{t-1}] = M_{t-1} \) \( \mathbb{P} \text{-a.s.} \) by sending \( n \) to infinity (Notice that the condition \( \mathbb{E}[M_T^-] < \infty \) has not been used in this step).

2. It remains to prove that \( M_t \) is \( \mathbb{P} \)-integrable for all \( t = 0, \ldots, T \).
By the second property of (3.1), together with the convexity of the function \( x \mapsto x^- \) and the Jensen inequality, we see that:
\[
M_t^- \leq \mathbb{E}[M_{t+1}^-|\mathcal{F}_t] \quad \text{for every } t = 0, \ldots, T-1.
\]
Then, \( E[M_t^-] \leq E[M_{t+1}^-] \leq \ldots \leq E[M_T^-] < \infty \) by the condition of the lemma.

We next use Fatou’s lemma to obtain:

\[
E[M_t^+] = E \left( \liminf_{n \to \infty} M_{t \wedge \tau_n}^+ \right) \\
\leq \liminf_{n \to \infty} E \left[ M_{t \wedge \tau_n}^+ \right] = \liminf_{n \to \infty} E \left[ M_{t \wedge \tau_n}^+ + M_{t \wedge \tau_n}^- \right].
\]

Recall that the stopped process \( M_{\tau_n}^+ \) is an \((\mathcal{F}, \mathbb{P})\)-martingale. Then:

\[
E[M_t^+] \leq M_0 + \liminf_{n \to \infty} E \left[ M_{t \wedge \tau_n}^- \right] \leq M_0 + \sum_{t=0}^{T} E[M_t^-] < \infty.
\]

\[\Box\]

We next provide an example of a local martingale which fails to be a martingale.

**Example 3.13** Let \((A_n)_{n \geq 1}\) be a measurable partition of \(\Omega\) with \(\mathbb{P}[A_n] = 2^{-n}\). Let \((Z_n)_{n \geq 1}\) be a sequence of random variables independent of \(A_n\) with \(\mathbb{P}[Z_n = 1] = \mathbb{P}[Z_n = -1] = 1/2\). Set

\[
\mathcal{F}_0 := \sigma(A_n, n \geq 1) \quad \text{and} \quad \mathcal{F}_1 := \sigma(A_n, Z_n, n \geq 1),
\]

and consider the process

\[
X_0 := 0 \quad \text{and} \quad X_1 := Y_\infty = \lim_{n \to \infty} Y_n \quad \text{where} \quad Y_n := \sum_{1 \leq p \leq n} 2^p Z_p 1_{A_p}.
\]

Clearly, \(X = \{X_0, X_1\}\) is not an \((\mathcal{F}, \mathbb{P})\)-martingale as \(X_1 = Y_\infty\) is not integrable. We next show that \(X\) is an \((\mathcal{F}, \mathbb{P})\)-local martingale. Define the sequence of stopping times:

\[
\tau_n := +\infty \text{ on } \cup_{1 \leq p \leq n} A_p \quad \text{and} \quad \tau_n := 0 \text{ on } \Omega \setminus (\cup_{1 \leq p \leq n} A_p).
\]

For every \(n\), the stopped process \(X_\tau^n\) is given by:

\[
X_{0\tau_n}^n = 0 \quad \text{and} \quad X_{1\tau_n}^n = Y_n.
\]

Since the random variable \(Y_n\) is bounded and independent of \(\mathcal{F}_0\), we easily see that \(E[Y_n|\mathcal{F}_0] = 0\). Hence, \(X_\tau^n\) is an \((\mathcal{F}, \mathbb{P})\)-martingale. \[\Box\]
Remark 3.2  In finite discrete-time, Lemma 3.1 provides an easy necessary and sufficient condition for a local martingale to be a martingale (necessity is trivial). This condition does not extend to the continuous-time framework, as we will provide an example of a positive continuous-time (integrable) local martingale which fails to be a martingale.

3.6 American contingent claims in complete financial markets

3.6.1 Problem formulation

In the previous chapters, we have considered the so-called “European” contingent claims which are defined by some payoff at some given maturity $T$.

For instance, a European call option on the asset $S^1$ is defined by the payoff $(S^1_T - K)^+$ at time $T$. The buyer of such a contingent claim can exercise the option at the maturity date $T$ by buying the risky asset $S^1$ at the pre-fixed price $K$. Similarly, a European put option on the asset $S^1$ is defined by the payoff $(K - S^1_T)^+$ at time $T$, and the buyer of such a contingent claim can exercise the option at the maturity date $T$ by selling the risky asset $S^1$ at the pre-fixed price $K$.

Hence, the buyer of the European option decides whether to exercise or not the option at the final time $T$. The latter restriction is dropped in the case of American options, which allow the buyer to exercise the option at any time before the maturity $T$. Consequently, an American call option is defined by the process of payoffs $\{(S^1_t - K)^+, \; t = 1, \ldots, T\}$, i.e. the buyer of this contingent claim receives the payoff $(S_t - K)^+$ if he decides to exercise the option at time $t$.

Similarly, an American put option allows the buyer to sell the risky asset $S^1$, for the price defined by the strike $K$, at any time before the maturity $T$. Such an option is defined by the process of payoffs $\{(K - S_t)^+, \; t = 1, \ldots, T\}$, i.e. the buyer of an American put option receives the payoff $(K - S_t)^+$ if he
decides to exercise the option at time \( t \).

This leads to the following concept of American contingent claims.

**Definition 3.9** An American contingent claim is a process \( B = \{B_t, t = 1, \ldots, T\} \) adapted to the filtration \( \mathbb{F} \).

The seller of an American contingent claim has to face the promised payoff \( B_t \) at any time \( t \) before the maturity \( T \), as the buyer can decide to exercise his right. Then, the relevant super-hedging problem in this context is:

\[
V^a(B) := \inf \left\{ x \in \mathbb{R} : X^x,\theta_t \geq B_t, \quad 1 \leq t \leq T \text{ a.s. for a } \theta \in \mathcal{A} \right\}. \quad (3.2)
\]

### 3.6.2 Decomposition of supermartingales

In the previous chapters, the concept of martingales played an important role. In the setting of American contingent claim, we shall make use of the notion of supermartingales.

**Definition 3.10** Let \( \mathbb{Q} \) be some given probability measure, and \( Y \) an \( \mathbb{F} \)-adapted process. We say that \( Y \) is a \( \mathbb{Q} \)-supermartingale for the filtration \( \mathbb{F} \) if \( Y_t \) is \( \mathbb{Q} \)-integrable for all \( t \leq T \) and

\[
\mathbb{E}^{\mathbb{Q}}[Y_t|\mathcal{F}_k] \leq Y_k \quad \text{for } 0 \leq k < t \leq T.
\]

It is immediately checked that we may formulate an equivalent definition by taking \( k = t - 1 \). The analysis of supermartingales is usually reduced to that of martingales thanks the following decomposition result.

**Theorem 3.7** (Doob-Meyer). Let \( \mathbb{Q} \) be a probability measure and \( Y \) a \( \mathbb{Q} \)-supermartingale for \( \mathbb{F} \). Then \( Y \) has the following unique decomposition:

\[
Y_t = M_t - A_t, \quad t = 0, \ldots, T,
\]

where \( M \) is a \( \mathbb{Q} \)-martingale for \( \mathbb{F} \) and \( A \) is a predictable non-decreasing process with \( A_0 = 0 \).
Proof. For $t = 0$, the only possible choice is $M_0 = Y_0$ since $A_0 = 0$. Next, we must have:

$$Y_{t+1} - Y_t = M_{t+1} - M_t - (A_{t+1} - A_t).$$

Since $M$ has to be a martingale, this implies that:

$$-(A_{t+1} - A_t) = \mathbb{E}^Q [Y_{t+1} | \mathcal{F}_t] - Y_t, \quad t = 0, \ldots, T - 1.$$

Since $Y$ is a $\mathbb{Q}$-supermartingale and $A_0 = 0$, this defines a non-decreasing process $A$. Then, the unique choice for the process $M$ is:

$$M_0 = Y_0 \text{ and } M_{t+1} - M_t = Y_{t+1} - \mathbb{E}^Q [Y_{t+1} | \mathcal{F}_t], \quad t = 0, \ldots, T - 1.$$

Clearly, this process is a $\mathbb{Q}$-martingale. ♦

3.6.3 The Snell envelope

Let $B = \{B_t, t = 0 \ldots, T\}$ be an $\mathbb{F}$-adapted process (an American contingent claim !) satisfying:

$$\mathbb{E}^Q \left[ \sup_{t=0, \ldots, T} |B_t| \right] < \infty,$$

where $\mathbb{Q}$ is some given probability measure. In this paragraph, we study the process $Y$ defined by the backward induction

$$Y_T = B_T \text{ and } Y_t = \max \{B_t, \mathbb{E}^Q [Y_{t+1} | \mathcal{F}_t] \}, \quad t = 0, \ldots, T - 1.$$

Exercise 3.3 Provide a financial interpretation of the process $Y$.

Proposition 3.3 The process $Y$ is a $\mathbb{Q}$-supermartingale. It is the smallest $\mathbb{Q}$-supermartingale which is larger than the process $B$.

Proof. The supermartingale property of $Y$ together with the fact that $Y \geq B$ follows immediately from the definition of $Y$. Let $\bar{Y}$ be a $\mathbb{Q}$-supermartingale larger than $B$, and let us show by induction that $\bar{Y}_t \geq Y_t$ for all $t = 0, \ldots, T$ $\mathbb{Q}$-a.s.
By definition, \( \bar{Y}_T \geq B_T = Y_T \). Suppose that \( \bar{Y}_t \geq Y_t \). Since \( \bar{Y} \) is a \( \mathbb{Q} \)-supermartingale, we have:

\[
\bar{Y}_{t-1} \geq \mathbb{E}^\mathbb{Q}[\bar{Y}_t | \mathcal{F}_{t-1}] \geq \mathbb{E}^\mathbb{Q}[Y_t | \mathcal{F}_{t-1}].
\]

Since \( \bar{Y} \geq B \), this implies that

\[
\bar{Y}_{t-1} \geq \max \{ B_{t-1}, \mathbb{E}^\mathbb{Q}[Y_t | \mathcal{F}_{t-1}] \} = Y_{t-1}.
\]

\[\diamondsuit\]

**Definition 3.11** The process \( Y \) is called the Snell envelope of \( B \) for the probability measure \( \mathbb{Q} \) and the filtration \( \mathcal{F} \).

Let \( \mathcal{T} \) denote the collection of all stopping times with values in \( \{0, \ldots, T\} \). Our next result provides a representation of the Snell envelope in terms of the underlying process \( B \) and the stopping times of \( \mathcal{T} \).

**Proposition 3.4** The random variable

\[
\nu^* := \inf \{ t = 0, \ldots, T : Y_t = B_t \}
\]

is a stopping time such that the stopped process \( Y^{\nu^*} \) is a \( \mathbb{Q} \)-martingale and

\[
Y_0 = \sup_{\nu \in \mathcal{T}} \mathbb{E}^\mathbb{Q}[B_\nu] = \mathbb{E}^\mathbb{Q}[B_{\nu^*}].
\]

**Proof.** (i) We first show that \( \nu^* \) defines a stopping time. Since \( Y_T = B_T \), the random variable \( \nu^* \) takes values in \( \{0, \ldots, T\} \). Clearly \( \{ \nu^* = 0 \} = \{ Y_0 = B_0 \} \in \mathcal{F}_0 \) and for \( t \geq 1 \):

\[
\{ \nu^* = t \} = (\cap_{k=0}^{t-1} \{ Y_k > B_k \}) \cap \{ Y_t = B_t \} \in \mathcal{F}_t
\]

since \( Y \) and \( B \) are both \( \mathbb{F} \)-adapted.

(ii) We next prove that the stopped process \( Y^{\nu^*} \) is a \( \mathbb{Q} \)-martingale. By definition of \( \nu^* \), we have:

\[
Y^{\nu^*}_{t+1} - Y^{\nu^*}_t = (Y_{t+1} - Y_t) \mathbf{1}_{\{ \nu^* \geq t+1 \}}.
\]
By definition of the process $Y$, we have $Y_t > B_t$ and $Y_t = \mathbb{E}^Q[ Y_{t+1} | \mathcal{F}_t ]$ on the event set $\{ \nu^* \geq t + 1 \}$ ($\mathcal{F}_t$-measurable as the complement of $\{ \nu^* \leq t \}$).

Hence:

$$Y_{t+1}^{\nu^*} - Y_t^{\nu^*} = (Y_{t+1} - \mathbb{E}^Q[ Y_{t+1} | \mathcal{F}_t ]) 1_{\{ \nu^* \geq t + 1 \}} ,$$

and the required result is obtained by taking the expected values conditionally on $\mathcal{F}_t$ and using the fact that $\{ \nu^* \geq t + 1 \} \in \mathcal{F}_t$.

(iii) Since the stopped process $Y^{\nu^*}$ is a $Q$–martingale, we have:

$$Y_0 = \mathbb{E}^Q [ Y_T^{\nu^*} ] = \mathbb{E}^Q [ Y_{T \wedge \nu^*} ] = \mathbb{E}^Q [ Y_{\nu^*} ] = \mathbb{E}^Q [ B_{\nu^*} ] .$$

Moreover, for all $\nu \in \mathcal{T}$, the stopped process $Y^{\nu}$ is a $Q$–supermartingale. Therefore,

$$Y_0 \geq \mathbb{E}^Q [ Y_T^{\nu} ] = \mathbb{E}^Q [ Y_{\nu} ] \geq \mathbb{E}^Q [ B_{\nu} ]$$

by definition of $Y$. This completes the proof of the proposition. ♦

### 3.6.4 Valuation of American contingent claims in a complete market

In this section, we assume that the financial market satisfies the no-arbitrage condition and is complete. By Theorems 3.4 and 3.5, we know that the set of all equivalent martingale measures, or risk neutral probability measures, is reduced to a singleton:

$$\mathcal{M}(S) = \{ \mathbb{P}^0 \} .$$

We consider an American contingent claim satisfying the integrability condition

$$\mathbb{E}^{\mathbb{P}^0} \left[ \sup_{t=0,...,T} | \tilde{B}_t | \right] < \infty , \quad (3.3)$$

The next result expresses the super-replication problem for American contingent claim (3.2) as an optimal stopping problem.
Theorem 3.8 Let $B$ be an American contingent claim satisfying (3.3). Then,

\[ V^a(B) = \sup_{\nu \in \mathcal{T}} \mathbb{E}^{\mathbb{P}^0}[\tilde{B}_\nu] = \mathbb{E}^{\mathbb{P}^0}[\tilde{B}_{\nu^*}] , \]

where $\nu^* = \inf\{t = 0, \ldots, T : Y_t = B_t\}$ and $Y$ is the Snell envelope of the process $B$ for the probability measure $\mathbb{P}^0$ and the filtration $\mathbb{F}$.

Proof. (i) Let $\tilde{Y}$ be the Snell envelope of the process $\tilde{B}$. This defines the process $Y$ by the usual relation $Y := S^0\tilde{Y}$. As a $\mathbb{P}^0$-supermartingale, the Snell envelope $\tilde{Y}$ has the Doob-Meyer decomposition:

\[ \tilde{Y}_t = M_t - A_t , \quad t = 0, \ldots, T , \]

where $M$ is a $\mathbb{P}^0$-martingale and $A$ a non-decreasing process with $A_0 = 0$. Since the financial market is complete, the (European) contingent claim defined by the random variable $S^0_T M_T$ is attainable, i.e. there exists a self-financing portfolio strategy $\theta \in \mathcal{A}$ such that:

\[ M_T = \mathbb{E}^{\mathbb{P}^0}[M_T] + \sum_{t=0}^{T-1} \theta_t \cdot (\tilde{S}_{t+1} - \tilde{S}_t) \]

\[ = Y_0 + \sum_{t=0}^{T-1} \theta_t \cdot (\tilde{S}_{t+1} - \tilde{S}_t) \]

as $A_0 = 0$. Since the process $\tilde{S}$ is a $\mathbb{P}^0$-martingale, the process $\tilde{X}^{Y_0,\theta}$ is a $\mathbb{P}^0$-local martingale. Since $\tilde{X}^{Y_0,\theta}_T = M_T = \tilde{Y}_T + A_T = \tilde{B}_T + A_T \geq \tilde{B}_T$, we see that $(\tilde{X}^{Y_0,\theta}_T)^- \leq \tilde{B}_T$, so that $(\tilde{X}^{Y_0,\theta}_T)^-$ is $\mathbb{P}^0$-integrable by Condition (3.3). By Lemma 3.1, this proves that:

\[ M_t = \tilde{X}^{Y_0,\theta}_t = Y_0 + \sum_{k=0}^{t-1} \theta_k \cdot (\tilde{S}_{k+1} - \tilde{S}_k) , \quad t = 1, \ldots, T . \]

We have then found a self-financing portfolio strategy $\theta \in \mathcal{A}$ such that:

\[ \tilde{X}^{Y_0,\theta}_t = Y_0 + \sum_{k=0}^{t-1} \theta_k \cdot (\tilde{S}_{k+1} - \tilde{S}_k) \]

\[ = Y_t + A_t \]

\[ \geq Y_t \]

\[ \geq \tilde{B}_t \]
for \( t = 1, \ldots, T \). By definition of \( V^a \), this shows that \( Y_0 \geq V^a(B) \), and by Proposition 3.4, we obtain the first inequality
\[
V^a(B) \leq \sup_{\nu \in \mathcal{T}} E^\mathbb{P}_0 \left[ \tilde{B}_\nu \right] = E^\mathbb{P}_0[\tilde{B}_\nu].
\]
It remains to show that the converse inequality holds. We claim that
\[
V^a(B) \geq V(B_\nu) \quad \text{for all } \nu \in \mathcal{T}.
\] (3.4)
Then, since \( p(B_\nu) = E^\mathbb{P}_0[B_\nu] \) for all \( \nu \in \mathcal{T} \), this provides the required inequality. To prove (3.4), we observe that
\[
V^a(B) \geq \bar{V}(B_\nu, \nu) := \inf \{ x \in \mathbb{R} : \exists \theta \in \mathcal{A}, X^{x,\theta}_\nu \geq B_\nu \text{ a.s.} \}
\]
for every \( \nu \in \mathcal{T} \). We next check that \( V(B_\nu) = \bar{V}(B_\nu, \nu) \). For \( x > \bar{V}(B_\nu, \nu) \), there exists a self-financing portfolio strategy \( \theta \in \mathcal{A} \) such that \( X^{x,\theta}_\nu \geq B_\nu \) a.s. Consider the portfolio strategy
\[
\bar{\theta}_t(\omega) = \theta_t(\omega)1_{\{\nu(\omega) \leq t\}}.
\]
Then, \( X^{\bar{\theta},\bar{\theta}}_T = X^{x,\bar{\theta}}_T = X^{x,\theta}_T \geq B_\nu \), which shows that \( x \geq V(B_\nu) \). By the arbitrariness of \( x > \bar{V}(B_\nu, \nu) \), this proves that \( \bar{V}(B_\nu, \nu) \geq V(B_\nu) \).

Conversely, let \( x > V(B_\nu) \). Then, \( X^{x,\theta}_T \geq B_\nu \) a.s. for some self-financing portfolio strategy \( \theta \in \mathcal{A} \). Now observe that the process \( X^{x,\theta}_T \) is a \( \mathbb{P}^0 \)-martingale since \( \tilde{S} \) is a \( \mathbb{P}^0 \)-martingale and \( B^-_\nu \) is \( \mathbb{P}^0 \)-integrable. Consequently, by taking expected values under \( \mathbb{P}^0 \) conditionally to \( \mathcal{F}_\nu \), it follows from the optional sampling Theorem 3.6 that \( X^{x,\theta}_T \geq B_\nu \) a.s. Hence, \( x \geq \bar{V}(B_\nu, \nu) \) and therefore \( V(B_\nu) \geq \bar{V}(B_\nu, \nu) \).

The previous proof also provides as a by-product the optimal superhedging strategy, as the perfect hedging strategy of the European contingent claim \( B_\nu^* \).

We conclude this chapter by the following sufficient condition for an American contingent claim to coincide with the corresponding European contingent claim \( B_T \).

48
**Corollary 3.1** Suppose that the process \( \tilde{B} \) is a \( \mathbb{P}^0 \)-submartingale, i.e. \( \tilde{B}_k \leq \mathbb{E}^\mathbb{P}_0 [\tilde{B}_t|\mathcal{F}_k] \) for \( 0 \leq k \leq t \leq T \). Then:

\[
V^a(B) = V(B_T) = \mathbb{E}^\mathbb{P}_0 [\tilde{B}_T].
\]

**Proof.** By the optional sampling Theorem 3.6:

\[
V(B_T) = \mathbb{E}^\mathbb{P}_0 [\tilde{B}_T] \leq \sup_{\nu \in T} \mathbb{E}^\mathbb{P}_0 [\tilde{B}_{\nu}] = V^a(B_T) \leq \sup_{\nu \in T} \mathbb{E}^\mathbb{P}_0 [E^Q[\tilde{B}_T|\mathcal{F}_{\nu}]] = \mathbb{E}^\mathbb{P}_0 [\tilde{B}_T]
\]

by the tower property of conditional expectations. \( \diamond \)

**Remark 3.3** The last corollary applies to the case of American call options. Let \( r \) be a constant interest rate, to simplify, and let us check that the process \( \{e^{-rt}(S_t - K)^+, \ t = 0, \ldots, T\} \) is a \( \mathbb{P}^0 \)-submartingale. For \( 0 \leq k < t \leq T \), we have:

\[
\mathbb{E}^\mathbb{P}_0 [e^{-rt}(S_t - K)^+|\mathcal{F}_k] \geq \mathbb{E}^\mathbb{P}_0 [e^{-rt}(S_t - K)|\mathcal{F}_k] = e^{-rk}S_k - e^{-rt}K \geq e^{-r}(S_k - K)
\]

since \( \tilde{S} \) is a \( \mathbb{P}^0 \)-martingale and \( r \geq 0 \). Since \( \mathbb{E}^\mathbb{P}_0 [(S_t - K)^+|\mathcal{F}_k] \geq 0 \), this provides:

\[
\mathbb{E}^\mathbb{P}_0 [e^{-rt}(S_t - K)^+|\mathcal{F}_k] \geq e^{-rk}(S_k - K)^+.
\]

We may also obtain this inequality by the Jensen inequality.

From the previous remark, the no-arbitrage price of an American call option is equal to that of the corresponding European call option, and it is optimal to exercise the American option at the maturity \( T \). This property was obtained by a no-arbitrage argument without any model specification in section 1.4.

Observe however that the corollary does not apply to American put options whenever the interest rate parameter \( r \) is strictly positive.
Chapter 4

The Cox-Ross-Rubinstein model

In this chapter we study the best-known multiperiod discrete time model: the binomial tree, or Cox-Ross-Rubinstein (CRR) model. Despite its simplicity, this model is quite often used in practice, because it provides simple algorithms for pricing exotic derivatives and can be considered as an approximation to more complex continuous time models.

4.1 The one-period binomial model

We first study the simplest one-period financial market \( T = 1 \). Let \( \Omega = \{ \omega_u, \omega_d \} \), \( \mathcal{F} \) the \( \sigma \)-algebra consisting of all subsets of \( \Omega \), and \( \mathbb{P} \) a probability measure on \( (\Omega, \mathcal{F}) \) such that \( 0 < \mathbb{P}(\omega_u) < 1 \).

The financial market contains a non-risky asset with price process

\[
S_0^0 = 1, \quad S_1^0(\omega_u) = S_1^0(\omega_d) = e^r := 1 + R,
\]

and one risky asset \( (d = 1) \) with price process

\[
S_0 = s, \quad S_1(\omega_u) = su, \quad S_1(\omega_d) = sd,
\]

where \( s, r, u \) and \( d \) are given strictly positive parameters with \( u > d \). Such
a financial market can be represented by the binomial tree:

<table>
<thead>
<tr>
<th>Time 0</th>
<th>Time 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risky asset</td>
<td>$S_0 = s$</td>
</tr>
<tr>
<td>Non-risky asset</td>
<td>$S_0^0 = 1$</td>
</tr>
</tbody>
</table>

(i) The No-Arbitrage condition: By Theorem 2.1, the no-arbitrage condition is equivalent to the existence of a probability measure $Q$, defined by $q_u = Q(\omega_u)$ and $q_d = Q(\omega_d) = 1 - q_u$, equivalent to $P$, i.e. $0 < q_u < 1$, such that

$$q_u \frac{su}{1 + R} + (1 - q_u) \frac{sd}{1 + R} = S_0.$$

This linear equation admits a unique solution:

$$q_u = \frac{1 + R - d}{u - d},$$

and we directly check that $0 < q_u < 1$ if and only if:

$$u > 1 + R > d. \quad (4.1)$$

In addition, we observe that the set of risk-neutral probabilities is reduced to a singleton in the present setting and hence the market is complete.

(ii) Hedging contingent claims: A contingent claim is defined by its payoff $B_u := B(\omega_u)$ and $B_d := B(\omega_d)$ at time 1.

The discounted wealth at time 1 of a portfolio with initial value $x$ and the quantity of risky asset equal to $\theta$ is given by:

$$\tilde{X}^{x,\theta}_1 = x + \theta(\tilde{S}_1 - s).$$

In this context, it turns out that there exists a pair $(x^0, \theta^0)$ such that $X^{x^0,\theta^0}_1 = B$. Indeed, the equality $X^{x^0,\theta^0}_1 = B$ is a system of two (linear) equations with
two unknowns which can be solved straightforwardly:

\[ x_0 = q_u \frac{B_u}{1+R} + (1-q_u) \frac{B_d}{1+R} = \mathbb{E}^Q[B] \quad \text{and} \quad \theta_0 = \frac{B_u - B_d}{su - sd} \, . \]

We also observe in our simple context that such a pair \((x_0, \theta_0)\) is unique.

Then \(x_0 = \mathbb{E}^Q[Be^{-r}]\) and \((x_0, \theta_0)\) is a perfect replication strategy for \(B\), i.e. \(X_{x_0,\theta_0} = B\).

(iii) No arbitrage valuation: Let us denote by \(p(B)\) the market price of the contingent claim contract. Under the no-arbitrage condition we have that

\[ p(B) = \mathbb{E}^Q[Be^{-r}] \, . \quad (4.2) \]

Hence in the context of the simple binomial model, the no-arbitrage condition implies a unique valuation rule. This is the no-arbitrage price of the contingent claim \(B\).

In order to better understand this result, let us show directly this equality. Suppose that the contingent claim \(B\) is available for trading at time 0. We will now show that, under the no arbitrage condition, the price \(p(B)\) of the contingent claim contract is necessarily given by \(p(B) = x_0\).

(iii-a) Indeed, suppose that \(p(B) < x_0\), and consider the following portfolio strategy:

- at time 0, pay \(p(B)\) to buy the contingent claim, so as to receive the payoff \(B\) at time 1,
- perform the self-financing strategy \((-x_0, -\theta)\), this leads to paying \(-x_0\) at time 0, and receiving \(-B\) at time 1

The initial capital needed to perform this portfolio strategy is \(p(B) - x_0 < 0\).

At time 1, the terminal wealth induced by the self-financing strategy exactly compensates the payoff of the contingent claim. We have then built an arbitrage opportunity in the financial market augmented with the contingent claim contract, thus violating the no arbitrage condition on this enlarged financial market.

(iii-b) If \(p(B) > x_0\), we consider the following portfolio strategy:
- at time 0, receive \( p(B) \) by selling the contingent claim, so as to pay the payoff \( B \) at time 1,
- perform the self-financing strategy \((x_0, \theta)\), this leads to paying \( x_0 \) at time 0, and receiving \( B \) at time 1

The initial capital needed to perform this portfolio strategy is \(-p(B) + x_0 < 0\). At time 1, the terminal wealth induced by the self-financing strategy exactly compensates the payoff of the contingent claim. This again defines an arbitrage opportunity in the financial market augmented with the contingent claim contract, thus violating the no arbitrage condition on this enlarged financial market.

### 4.2 The dynamic CRR model

#### 4.2.1 Description of the model

In this chapter, we consider a dynamic version of the binomial model introduced in Section 4.1.

Let \( \Omega = \{-1, 1\}^N \), and let \( \mathcal{F} \) be the Borel \( \sigma \)-algebra on \( \Omega \). Let \((Z_k)_{k \geq 0}\) be a sequence of independent random variables with distribution \( \mathbb{P}[Z_k = 1] = \mathbb{P}[Z_k = -1] = 1/2 \). We shall see later that we may replace the value 1/2 by any parameter in \((0, 1)\), see Remark 4.6). We consider the standard filtration \( \mathcal{F}_0 = \{\emptyset, \mathcal{F}\} \), \( \mathcal{F}_k = \sigma(Z_0, \ldots, Z_k) \) and \( \mathbb{P}^n = \{\mathcal{F}_0, \ldots, \mathcal{F}_n\} \).

Let \( T > 0 \) be some fixed finite horizon, and \((b_n, \sigma_n)_{n \geq 1}\) the sequence defined by :

\[
b_n = b \frac{T}{n} \quad \text{and} \quad \sigma_n = \sigma \left( \frac{T}{n} \right)^{1/2},
\]

where \( b \) and \( \sigma \) are two given strictly positive parameters.

**Remark 4.4** All the results of this section hold true with a sequence \((b_n, \sigma_n)\) satisfying :

\[
nb_n \to bT \quad \text{and} \quad \sqrt{n}\sigma_n \to \sigma\sqrt{T} \quad \text{whenever} \ n \to \infty.
\]
For \( n \geq 1 \), we consider the price process of a single risky asset \( S^n = \{S^n_k, k = 0, \ldots, n\} \) defined by:

\[
S^n_0 = s \quad \text{and} \quad S^n_k = s \exp \left( k b_n + \sigma_n \sum_{i=1}^{k} Z_i \right), \quad k = 1, \ldots, n.
\]

The non-risky asset is defined by a constant interest rate parameter \( r \), so that the return from a unit investment in the bank during a period of length \( T/n \) is

\[
R_n := e^{r(T/n)} - 1.
\]

For each \( n \geq 1 \), we have then defined a financial market with time step \( T/n \).

In order to ensure that these financial markets satisfy the no-arbitrage condition, we assume that:

\[
d_n < R_n - 1 < u_n \quad \text{where} \quad u_n = e^{b_n + \sigma_n}, \quad d_n = e^{b_n - \sigma_n}, \quad n \geq 1 \quad (4.3)
\]

**Exercise 4.4** Show that (4.3) is a necessary and sufficient condition for the financial markets of this section to satisfy the no-arbitrage condition.

Under this condition, there exists a unique risk neutral measure \( \mathbb{P}^0_n \) defined by:

\[
\mathbb{P}_n[Z_i = 1] = q_n := \frac{1 + R_n - d_n}{u_n - d_n}.
\]

By Theorem 2.2, the above financial markets are complete.

### 4.2.2 Valuation and hedging of European options

In the rest of this chapter, we consider the contingent claims

\[
B^n := g(S^n) \quad \text{where} \quad g(s) = (s - K)^+ \quad \text{and} \quad K > 0.
\]

From the results of the previous chapters, the no-arbitrage price of this European call option is:

\[
p^n(B^n) = e^{-rT} \mathbb{E}^{\mathbb{P}^0_n} [(S^n - K)^+].
\]
Under the probability measure $\mathbb{P}_n^0$, the random variables $(1 + Z_i)/2$ are independent and identically distributed as a Bernoulli with parameter $q_n$. Then:

$$\mathbb{P}_n^0 \left[ \sum_{i=1}^{n} \frac{1 + Z_i}{2} = j \right] = C_n^j q_n^j (1 - q_n)^{n-j} \quad \text{for} \quad j = 0, \ldots, n .$$

This provides

$$p_n^0(B^n) = e^{-rT} \sum_{j=0}^{n} g \left( su_n^i d_n^{n-j} \right) C_n^j q_n^j (1 - q_n)^{n-j} .$$

By moving the time origin to $kT/n$, we obtain a trivial extension of the problem with no-arbitrage price at time $kT/n$ given by:

$$p_k^n(B^n) = e^{-r(1-(k/n))T} E_n^0 \left[ p_{k+1}^n(B^n) \mid \mathcal{F}_k \right]$$

$$= e^{-r(1-(k/n))T} \sum_{j=0}^{n-k} g \left( S_n^k u_n^j d_n^{n-k-j} \right) C_{n-k}^j q_n^j (1 - q_n)^{n-k-j} .$$

By the Tower property of conditional expectations, we observe that this provides:

$$p_k^n(B^n) = \frac{1}{1 + R_n} E_n^0 \left[ p_{k+1}^n(B^n) \mid \mathcal{F}_k \right] \quad \text{for} \quad k = 0, \ldots, n-1 . \quad (4.4)$$

This formula shows that the no-arbitrage price of the European call option at time 0 can be computed by a backward induction, avoiding the computation of the combinatorial terms $C_n^p$. Indeed $p_n^0(B^n) = B_n$ is given and (4.4) reduces to:

$$p_k^n(B^n) = \frac{1}{1 + R_n} \left[ q_n p_{k+1}^n(B^n)_{u_n} + (1 - q_n) p_{k+1}^n(B^n)_{d_n} \right] . \quad (4.5)$$

Expression (4.4) reduces the multi-period valuation and hedging problem into a sequence “local” one-period valuation and hedging problems. Indeed,
\( p^n_k(B^n) \) is the no-arbitrage price of the contingent claim \( p^n_{k+1}(B^n) \):

\[
\begin{array}{ll}
\text{Risky asset} & S^n_k \\
\text{Contingent claim} & p^n_k(B^n) \\
\end{array}
\]

\[
\begin{array}{ll}
\text{date } kT/n & \text{date } (k+1)T/n \\
& u_n S^n_k & d_n S^n_k \\
& p^n_{k+1}(B^n)_{u_n} & p^n_{k+1}(B^n)_{d_n} \\
\end{array}
\]

Therefore, we can describe the (perfect) hedging strategy of the contingent claim \( B^n \) by applying the one-period results of Section 4.1:

\[
\theta^n_{k+1} = \frac{p^n_{k+1}(B^n)_{u_n} - p^n_{k+1}(B^n)_{d_n}}{u_n S^n_k - d_n S^n_k}, \quad k = 0, \ldots, n - 1.
\]

**Remark 4.5** The hedging strategy is the finite differences approximation (on the binomial tree) of the partial derivative of the price of the contingent claim with respect to the spot price of the underlying asset. This observation will be confirmed in the continuous-time model section.

**Remark 4.6** The reference measure \( \mathbb{P} \) is not involved neither in the valuation formula, nor in the hedging formula. This is due to the fact that the no-arbitrage price in our complete market framework coincides with the super-hedging cost, which in turn depends on the reference measure only through the corresponding zero-measure sets.

**4.2.3 American contingent claims in the CRR model**

In the CRR model, American options can be priced by the same backward induction algorithm with only a minor modification of Equation 4.5. For simplicity we consider the case of an American put option with strike \( K \).
At the terminal date, the price $P^n_n(B^n)$ of the American put option coincides with the price $p^n_n(B^n)$ of the European put option with the same strike, since both coincide with the pay-off. At a date $k < n$, the buyer has the choice between exercising the option and receiving $(K - S^n_k)^+$ or not exercising. In the latter case, the option becomes “locally European” and its price is $\frac{1}{1 + R^n_n} E^Q[P^n_{k+1}(B^n)|\mathcal{F}_k]$ (called continuation value). The optimal exercise policy is to choose the largest of these two values:

$$P^n_k = \max\{(K - S^n_k)^+, \frac{1}{1 + R^n_n} [q^n_n P^n_{k+1}(B^n)_{u^n} + (1 - q^n_n) P^n_{k+1}(B^n)_{d^n}]\}.$$ 

Therefore, the rational buyer will exercise her option as soon as the $P^n_k$ becomes equal to $(K - S^n_k)$. For the seller of the option, the same hedging strategy

$$\theta^n_{k+1} = \frac{P^n_{k+1}(B^n)_{u^n} - P^n_{k+1}(B^n)_{d^n}}{u^n S^n_k - d^n S^n_k}$$

allows to hedge the put option until the exercise date. If the buyer does not exercise optimally, after the optimal exercise date, the hedging strategy is no longer self-financing and generates profit for the seller: the seller can withdraw some money from the account and still be able to face her obligations.

### 4.2.4 Continuous-time limit: A first approach to the Black-Scholes model

In this paragraph, we examine the asymptotic behavior of the Cox-Ross-Rubinstein model when the time step $T/n$ tends to zero, i.e. when $n \to \infty$. Our final goal is to show that the limit of the discrete-time valuation formulae coincides with the Black-Scholes formula which was originally derived in [1] in the continuous-time setting. We will also obtain this formula in the context of a continuous-time formulation in Chapter 8.

Although the following computations are performed in the case of European call options, the convergence argument holds for a large class of contingent claims.
Introduce the sequence:
\[
\eta_n := \inf\{ j = 0, \ldots, n : su_n^j d_n^{n-j} \geq K \},
\]
and let
\[
B(n, p, \eta) := \text{Prob} [\text{Bin}(n, p) \geq \eta],
\]
where Bin(n, p) is a Binomial random variable with parameters (n, p).

The following Lemma provides an interesting reduction of our problem.

**Lemma 4.2** For \( n \geq 1 \), we have:
\[
p^n(B^n) = sB \left( n, \frac{q_n u_n}{1 + R_n}, \eta_n \right) - K e^{-rT} B(n, q_n, \eta_n).
\]

**Proof.** Using the expression of \( p^n(B^n) \) obtained in the previous paragraph, we see that
\[
p^n(B^n) = (1 + R_n)^{-n} \sum_{j=\eta_n}^{n} (su_n^j d_n^{n-j} - K) C_n^j q_n^j (1 - q_n)^{n-j}
\]
\[
= s \sum_{j=\eta_n}^{n} C_n^j \left( \frac{q_n u_n}{1 + R_n} \right)^j \left( \frac{(1 - q_n)d_n}{1 + R_n} \right)^{n-j} - \frac{K}{(1 + R_n)^n} \sum_{j=\eta_n}^{n} C_n^j q_n^j (1 - q_n)^{n-j}.
\]
The required result follows by noting that \( q_n u_n + (1 - q_n)d_n = 1 + R_n \). \( \diamond \)

Hence, in order to derive the limit of \( p^n(B^n) \) when \( n \rightarrow \infty \), we have to determine the limit of the terms \( B(n, q_n u_n/(1 + R_n), \eta_n) \) et \( B(n, q_n, \eta_n) \).

We only provide a detailed exposition for the second term; the first one is treated similarly.

The main technical tool in order to obtain these limits is the following.

**Lemma 4.3** Let \((X_{k,n})_{1 \leq k \leq n}\) be a triangular sequence of iid Bernoulli random variables with parameter \( \pi_n \):
\[
\mathbb{P}[X_{k,n} = 1] = 1 - \mathbb{P}[X_{k,n} = 0] = \pi_n.
\]
Then:

\[
\sum_{k=1}^{n} X_{k,n} - n\pi_n \quad \Rightarrow \quad \mathcal{N}(0, 1) \quad \text{in distribution}.
\]

The proof of this lemma is reported at the end of this paragraph.

**Exercise 4.5** Use Lemma 4.3 to show that

\[
\ln \left( \frac{S^n}{s} \right) \quad \Rightarrow \quad \mathcal{N}(bT, \sigma^2T) \quad \text{in distribution under } \mathbb{P}.
\]

This shows that the Cox-Ross-Rubinstein model is a discrete-time approximation of a continuous-time model where the risky asset price has a log-normal distribution.

**Theorem 4.9** In the context of the Cox-Ross-Rubinstein model, the no-arbitrage price \( p^n(B^n) \) of a European call option converges, as \( n \to \infty \), to the Black-Scholes price:

\[
p(B) = s \mathcal{N} \left( d_+(s, \tilde{K}, \sigma^2T) \right) - \tilde{K} \mathcal{N} \left( d_-(s, \tilde{K}, \sigma^2T) \right)
\]

where

\[
\tilde{K} := Ke^{-rT}, \quad d_\pm(s, k, v) := \frac{\ln(s/k)}{\sqrt{v}} \pm \frac{\sqrt{v}}{2},
\]

and \( \mathcal{N}(x) = \int_{-\infty}^{x} e^{-v^2/2} dv/\sqrt{2\pi} \) is the cumulative distribution function of standard \( \mathcal{N}(0, 1) \) distribution.

**Proof.** Under the probability measure \( Q_n \), notice that \( B_i := (Z_i + 1)/2, i \geq 0 \), defines a sequence of iid Bernoulli random variables with parameter \( q_n \). Then

\[
B(n, q_n, \eta_n) = Q_n \left[ \sum_{j=1}^{n} B_i \geq \eta_n \right].
\]

We shall only develop the calculations for this term.
1. By definition of \( \eta_n \), we have

\[
 su_n^{\eta_n-1} d_n^{\eta_n+1} \leq K \leq su_n^{\eta_n} d_n^{\eta_n}.
\]

Then,

\[
 2 \eta_n \sigma \sqrt{T/n} + n \left( b \frac{T}{n} - \sigma \sqrt{T/n} \right) = \ln \left( \frac{K}{s} \right) + O \left( n^{-1/2} \right),
\]

which provides, by direct calculation that

\[
 \eta_n = \frac{n}{2} + \sqrt{n} \frac{\ln(K/s) - bT}{2\sigma \sqrt{T}} + o(\sqrt{n}) \quad (4.6)
\]

We also compute that

\[
 nq_n = \frac{1}{2} + \frac{\left( r - b - \frac{\sigma^2}{2} \right) T}{2\sigma \sqrt{T}} \sqrt{n} + o(\sqrt{n}) \quad (4.7)
\]

By (4.6) and (4.7), it follows that

\[
 \lim_{n \to \infty} \frac{\eta_n - nq_n}{\sqrt{nq_n(1-q_n)}} = -d_-(s, \tilde{K}, \sigma^2 T).
\]

2. Applying Lemma 4.3 to the sequence \((Z_1, \ldots, Z_n)\), we see that:

\[
 \mathcal{L}^{Q_n} \left( \frac{\sum_{k=1}^{n} Z_j - nq_n}{\sqrt{nq_n(1-q_n)}} \right) \to \mathcal{N}(0,1),
\]

where \( \mathcal{L}^{Q_n}(Z) \) denotes the distribution under \( Q_n \) of the random variable \( Z \). Then:

\[
 \lim_{n \to \infty} B(n, q_n, \eta_n) = \lim_{n \to \infty} Q_n \left[ \frac{\sum_{k=1}^{n} Z_j - nq_n}{\sqrt{nq_n(1-q_n)}} \geq \frac{\eta_n - nq_n}{\sqrt{nq_n(1-q_n)}} \right]
\]

\[
 = 1 - N \left( -d_-(s, \tilde{K}, \sigma^2 T) \right) = N \left( d_-(s, \tilde{K}, \sigma^2 T) \right).
\]

\( \diamond \)
Proof of Lemma 4.3. (i) We start by recalling a well-known result on characteristic functions. Let $X$ be a random variable with $\mathbb{E}[X^n] < \infty$. Then:

$$
\phi_X(t) := \mathbb{E}[e^{itX}] = \sum_{k=0}^{n} \frac{(it)^n}{n!} \mathbb{E}[X^k] + o(t^n).
$$

To prove this result, we denote $F(t, x) := e^{itx}$ and $f(t) = \mathbb{E}[F(t, X)]$. The function $t \mapsto F(t, x)$ is differentiable with respect to the $t$ variable. Since $|F_t(t, X)| = |iXF(t, X)| \leq |X| \in \mathbb{L}^1$, it follows from the dominated convergence theorem that the function $f$ is differentiable with $f'(t) = E[iXe^{itX}]$. In particular, $f'(0) = iE[X]$. Iterating this argument, we see that the function $f$ is $n$ times differentiable with $n-$th order derivative at zero given by:

$$
f^{(n)}(0) = i^n E[X^n].
$$

The expansion (4.8) is an immediate consequence of the Taylor-Young formula.

(ii) We now proceed to the proof of Lemma 4.3. Let

$$
Y_j := \frac{X_{j,n} - \pi_n}{\sqrt{n\pi_n(1-\pi_n)}} \quad \text{and} \quad \Sigma Y_n := \sum_{k=1}^{n} Y_j.
$$

Since the random variables $Y_j$ are independent and identically distributed, we have:

$$
\phi_{\Sigma Y_n}(t) = (\phi_{Y_1}(t))^n.
$$

Moreover, we compute directly that $E[Y_j] = 0$ and $E[Y_j^2] = 1/n$. Then, it follows from (4.8) that:

$$
\phi_{Y_1}(t) = 1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right).
$$

Sending $n$ to $\infty$, this provides:

$$
\lim_{n \to \infty} \phi_{\Sigma Y_n}(t) = e^{-t^2/2} = \phi_{N(0,1)}(t).
$$

This shows the convergence in distribution of $\Sigma Y_n$ towards the standard normal distribution. $\diamond$
Chapter 5

The Brownian motion

The purpose of this and the following chapter is to introduce the concepts from stochastic calculus which will be needed for the future development of the course, namely the Black and Scholes model for pricing and hedging European Vanilla options. We provide a treatment of all of these applications without appealing to the general theory of stochastic integration.

5.1 Filtration and stopping times

Throughout this chapter, \((\Omega, \mathcal{F}, \mathbb{P})\) is a given probability space.

A stochastic process with values in a set \(E\) is a map

\[
V : \mathbb{R}_+ \times \Omega \rightarrow E \\
(t, \omega) \mapsto V_t(\omega)
\]

The index \(t\) is conveniently interpreted as the time variable. In the context of these lectures, the state space \(E\) will be a subset of a finite dimensional space, and we shall denote by \(\mathcal{B}(E)\) the corresponding Borel \(\sigma\)-field. The process \(V\) is said to be measurable if the mapping

\[
V : (\mathbb{R}_+ \times \Omega, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}) \rightarrow (E, \mathcal{B}(E)) \\
(t, \omega) \mapsto V_t(\omega)
\]
is measurable. For a fixed $\omega \in \Omega$, the function $t \in \mathbb{R}_+ \mapsto V_t(\omega)$ is the sample path (or trajectory) of $V$ corresponding to $\omega$.

A filtration $\mathcal{F} = \{\mathcal{F}_t, \ t \geq 0\}$ is an increasing family of sub-$\sigma$–algebras of $\mathcal{F}$. As in the discrete-time context, $\mathcal{F}_t$ is intuitively understood as the information available up to time $t$. The increasing feature of the filtration, $\mathcal{F}_s \subset \mathcal{F}_t$ for $0 \leq s \leq t$, means that information can only increase as time goes on.

**Definition 5.12** A stochastic process $V$ is said to be

(i) adapted to the filtration $\mathcal{F}$ if the random variable $V_t$ is $\mathcal{F}_t$–measurable for every $t \in \mathbb{R}_+$,

(ii) progressively measurable with respect to the filtration $\mathcal{F}$ if the mapping

$$V : ([0, t] \times \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t) \rightarrow (E, \mathcal{B}(E))$$

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

is measurable for every $t \in \mathbb{R}_+$.

Given a stochastic process $V$, we define its canonical filtration by $\mathcal{F}^V_t := \sigma(V_s, s \leq t), \ t \in \mathbb{R}_+$. This is the smallest filtration to which the process $V$ is adapted.

Obviously, any progressively measurable stochastic process is measurable and adapted. The following result states that these two notions are equivalent for processes which are either right-continuous or left-continuous.

**Proposition 5.5** Let $V$ be a stochastic process with right-continuous sample paths or else left-continuous sample paths. Then, if $V$ is adapted to a filtration $\mathcal{F}$, it is also progressively measurable with respect to $\mathcal{F}$.

**Proof.** Assume that every sample path of $V$ is right-continuous (the case of left-continuous sample paths is treated similarly), and fix an arbitrary $t \geq 0$. Observe that $V_s(\omega) = \lim_{n \to \infty} V^n_s(\omega)$ for every $s \in [0, t]$, where $V^n$ is defined by

$$V^n_s(\omega) = V_{kt/n}(\omega) \text{ for } (k - 1)t < sn \leq kt \text{ and } k = 1, \ldots, n.$$
Since the map $V^n$ is obviously $\mathcal{B}([0, t]) \otimes \mathcal{F}$–measurable, we deduce that the measurability of the limit map $V$ defined on $[0, t] \times \Omega$.

A random time is a random variable $\tau$ with values in $[0, \infty]$. It is called
- a stopping time if the event set $\{ \tau \leq t \}$ is in $\mathcal{F}_t$ for every $t \in \mathbb{R}_+$,
- an optional time if the event set $\{ \tau < t \}$ is in $\mathcal{F}_t$ for every $t \in \mathbb{R}_+$.

Obviously, any stopping time is an optional time. It is an easy exercise to show that these two notions are in fact identical whenever the filtration $\mathcal{F}$ is right-continuous, i.e.

$$\mathcal{F}_{t+} := \cup_{s>t} \mathcal{F}_s = \mathcal{F}_t \quad \text{for every } t \geq 0.$$ 

This will be the case in all of the financial applications of this course. An important example of a stopping time is:

**Exercise 5.6 (first exit time)** Let $X$ be a stochastic process with continuous paths adapted to $\mathbb{F}$, and consider a closed subset $\Gamma \in \mathcal{B}(E)$. Show that the random time

$$T_\Gamma := \inf \{ t \geq 0 : X_t \not\in \Gamma \}$$

(with the convention $\inf \emptyset = \infty$) is a stopping time.

**Exercise 5.7** (i) If $\tau_1$ and $\tau_2$ are stopping times, then so are $\tau_1 + \tau_2$, $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$.

(ii) If $(\tau_n)_{n \geq 1}$ is a sequence of stopping times, then so is $\sup_{n \geq 1} \tau_n$.

(iii) If $(\tau_n)_{n \geq 1}$ is a sequence of optional times, then so are $\sup_{n \geq 1} \tau_n$, $\inf_{n \geq 1} \tau_n$, $\limsup_{n \to \infty} \tau_n$, $\liminf_{n \to \infty} \tau_n$.

As in the discrete-time framework, we provide a precise definition of the information available up to some stopping time $\tau$ of a filtration $\mathbb{F}$:

$$\mathcal{F}_\tau := \{ A \in \mathcal{F} : A \cap \{ \tau \leq t \} \in \mathcal{F}_t \text{ for every } t \in \mathbb{R}_+ \}$$

Finally, given a stopping time $\tau$, we shall frequently introduce the sequence

$$\tau_n := \frac{\lfloor n\tau \rfloor + 1}{n}, \quad n \geq 1,$$
which defines a decreasing sequence of stopping times converging a.s. to $\tau$. Here \( \lfloor t \rfloor \) denotes the largest integer less than or equal to $t$. Notice that the random time $\frac{\lfloor n \tau \rfloor}{n}$ is not a stopping time in general.

5.2 Martingales and optional sampling

In this section, we shall consider real-valued adapted stochastic processes $V$ on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The notion of martingales is defined similarly as in the discrete-time case.

Definition 5.13 Let $V$ be an $\mathbb{F}$-adapted stochastic process with $E|V_t| < \infty$ for every $t \in \mathbb{R}_+$.

(i) $V$ is a submartingale if $E[V_t|\mathcal{F}_s] \geq V_s$ for $0 \leq s \leq t$,

(ii) $V$ is a supermartingale if $E[V_t|\mathcal{F}_s] \leq V_s$ for $0 \leq s \leq t$,

(iii) $V$ is a martingale if $E[V_t|\mathcal{F}_s] = V_s$ for $0 \leq s \leq t$.

The following Doob’s optional sampling theorem states that submartingales and supermartingales satisfy the same inequalities when sampled along random times, under convenient conditions.

Theorem 5.10 (Optional sampling) Let $V$ be a right-continuous submartingale such that $V_\infty := \lim_{t \to \infty} V_t$ exists for almost every $\omega \in \Omega$. If $\tau_1 \leq \tau_2$ are two stopping times, then

$$E[V_{\tau_2}|\mathcal{F}_{\tau_1}] \geq V_{\tau_1} \quad \mathbb{P} \text{ – a.s.}$$

Proof. For stopping times $\tau_1$ and $\tau_2$ taking values in a finite set, the proof is identical to that of Theorem 3.6. In order to extend the result to general stopping times, we approximate the stopping times $\tau_i$ by the decreasing sequence $\tau_i^n := \frac{\lfloor n \tau_i \rfloor}{n} + 1$, and we use a limiting argument. \( \diamond \)

The optional sampling theorem requires the existence of a last element $V_\infty$ as defined in the statement of the theorem. For completeness, we observe that this condition is verified for right-continuous submartingales $V$. 

65
with $\sup_{t \geq 0} \mathbb{E}[V_t^+] < \infty$. This is the so-called submartingale convergence theorem, see e.g. Karatzas and Shreve [16] Theorem 1.3.15. In the context of these lectures, we shall simply apply the following consequence of the optional sampling theorem:

**Exercise 5.8** For a right-continuous submartingale $V$ and two stopping times $\tau_1 \leq \tau_2$, the optional sampling theorem holds under either of the following conditions:

(i) $\tau_2 \leq a$ for some constant $a > 0$,

(ii) there exists an integrable r.v. $Y$ such that $V_t \leq \mathbb{E}[Y|\mathcal{F}_t] \mathbb{P}-a.s.$ for every $t \geq 0$.

### 5.3 The Brownian motion

The Brownian motion was introduced by the scottish botanist Robert Brown in 1828 to describe the movement of pollen suspended in water. Since then it has been widely used to model various irregular movements in physics, economics, finance and biology. In 1905, Albert Einstein built a model for the trajectory of atoms subject to shocks, and obtained a Gaussian density. Louis Bachelier (1870-1946) was the very first to use the Brownian motion as a model for stock prices in his thesis in 1900, but his work was not recognized until the recent history. It is only sixty years later that Samuelson (Nobel Prize in economics 1970) suggested the Brownian motion as a model for stock prices. The real success of Brownian motion in the financial application was however realized by Fisher Black, Myron Scholes, et Robert Merton (Nobel Prize in economics 1997) who founded between 1969 and 1973 the modern theory financial mathematics by introducing the portfolio theory and the no-arbitrage pricing argument.

The first rigorous construction of the Brownian motion was achieved by Norbert Wiener in 1923, who provided many applications in signal theory and telecommunications. Paul Lévy, (X1904, Professor at Polytechnique) contributed to the mathematical study of the Brownian motion and proved
many surprising properties. Kyōshi Itō (1915-) developed the stochastic
differential calculus. The theory benefitted from the considerable activity on
martingales theory, in particular in France around P.A. Meyer.

**Definition 5.14** Let \( W = \{ W_t, \ t \in \mathbb{R}_+ \} \) be a stochastic process on the
probability space \((\Omega, \mathcal{F}, \mathbb{P})\), and \( \mathbb{F} \) a filtration. \( W \) is an \( \mathbb{F} \)-standard Brownian
motion if

(i) \( W \) is \( \mathbb{F} \)-adapted,

(ii) \( W_0 = 0 \) and the sample paths \( W(\omega) \) are continuous for a.e. \( \omega \in \Omega \),

(iii) For all \( 0 \leq s < t \), \( W_t - W_s \) is independent of \( \mathcal{F}_s \),

(iv) For all \( 0 \leq s < t \), the distribution of \( W_t - W_s \) is \( \mathcal{N}(0, t - s) \).

Let us observe that, for any given filtration \( \mathbb{F} \), an \( \mathbb{F} \)-standard Brownian
motion is also a \( \mathbb{F}^W \)-standard Brownian motion where \( \mathbb{F}^W \) is the canonical
filtration of \( W \). This justifies the consistency of the above definition with the
following one which does not refer to any filtration:

**Definition 5.15** Let \( W = \{ W_t, \ t \in \mathbb{R}_+ \} \) be a stochastic process on the
probability space \((\Omega, \mathcal{F}, \mathbb{P})\). \( W \) is a standard Brownian motion if

(i) \( W_0 = 0 \) and the sample paths \( W(\omega) \) are continuous for a.e. \( \omega \in \Omega \),

(ii) \( W \) has independent increments: for all \( 0 \leq t_1 < t_2 \leq t_3 < t_4 \), \( W_{t_4} - W_{t_3} \)
and \( W_{t_2} - W_{t_1} \) are independent.

(iii) For all \( 0 \leq s < t \), the distribution of \( W_t - W_s \) is \( \mathcal{N}(0, t - s) \).

Before discussing the properties of the Brownian motion, let us comment
on its existence as a continuous-time limit of a random walk. Given a family
\( \{ Y_i, \ i = 1, \ldots, n \} \) of \( n \) independent random variables defined by the distribution

\[
\mathbb{P}[Y_i = 1] = 1 - \mathbb{P}[Y_i = -1] = \frac{1}{2}, \quad (5.1)
\]
we define the symmetric random walk

\[ M_0 = 0 \text{ and } M_k = \sum_{j=1}^{k} Y_j \text{ for } k = 0, \ldots, n. \]

A continuous-time process can be obtained from the sequence \( \{M_k, k = 0, \ldots, n\} \) by linear interpolation:

\[ M_t := M_{\lfloor t \rfloor} + (t - \lfloor t \rfloor) Y_{\lfloor t \rfloor + 1} \text{ for } t \geq 0, \]

where \( \lfloor t \rfloor \) denotes the largest integer less than or equal to \( t \). The following figure shows a typical sample path of the process \( M \).

![Sample path of a random walk](image)

Figure 5.1: Sample path of a random walk

We next define a stochastic process \( W^n \) from the previous process by speeding up time and conveniently scaling:

\[ W^n_t := \frac{1}{\sqrt{n}} M_{nt}, \quad t \geq 0. \]
In the above definition, the normalization by $\sqrt{n}$ is suggested by the Central Limit Theorem. We next set

$$t_k := \frac{k}{n} \quad \text{for} \quad k \in \mathbb{N}$$

and we list some obvious properties of the process $W^n$:

- for $0 \leq i \leq j \leq k \leq \ell \leq n$, the increments $W^n_{t_k} - W^n_{t_i}$ and $W^n_{t_j} - W^n_{t_i}$ are independent,
- for $0 \leq i \leq k$, the two first moments of the increment $M_{t_k} - M_{t_i}$ are given by

$$\mathbb{E}[W^n_{t_k} - W^n_{t_i}] = 0 \quad \text{and} \quad \mathbb{V}[ar[W^n_{t_k} - W^n_{t_i}] = t_k - t_i,$$

which shows in particular that the normalization by $n^{-1/2}$ in the definition of $W^n$ prevents the variance of the increments from blowing up,

- with $F^n_t := \sigma(Y_j, j \leq \lfloor nt \rfloor)$, $t \geq 0$, the sequence $\{W^n_{t_k}, k \in \mathbb{N}\}$ is a discrete $\{F^n_{t_k}, k \in \mathbb{N}\}$-martingale:

$$\mathbb{E}[W^n_{t_k} | F^n_{t_i}] = W^n_{t_i} \quad \text{for} \quad 0 \leq i \leq k.$$

Hence, except the Gaussian feature of the increments, the discrete-time process $\{W^n_{t_k}, k \in \mathbb{N}\}$ is approximately a Brownian motion. One could even obtain Gaussian increments with the required mean and variance by replacing the distribution (5.1) by a convenient normal distribution. However, since our objective is to imitate the Brownian motion in the asymptotics $n \to \infty$, the Gaussian distribution of the increments is expected to hold in the limit by a central limit type of argument.

The following figures represent a typical sample path of the process $W^n$.

Another interesting property of the rescaled random walk, which will be inherited by the Brownian motion, is the following quadratic variation result:

$$[W^n, W^n]_{t_k} = \sum_{j=1}^{k} \left( W^n_{t_j} - W^n_{t_{j-1}} \right)^2 = t_k \quad \text{for} \quad k \in \mathbb{N}.$$
A possible proof of the existence of the Brownian motion consists in proving the convergence in distribution of the sequence $W^n_t$ toward a Brownian motion, i.e. a process with the properties listed in Definition 5.15. This is the so-called Donsker’s invariance principle. The interested reader may consult a rigorous treatment of this limiting argument in Karatzas and Shreve [16] Theorem 2.4.20.

We conclude this section by extending the definition of the Brownian motion to the vector case.

**Definition 5.16** Let $W = \{W_t, \ t \in \mathbb{R}_+\}$ be an $\mathbb{R}^n$—valued stochastic process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathbb{F}$ a filtration. $W$ is an $\mathbb{F}$—standard
Brownian motion if the components $W^i$, $i = 1, \ldots, n$, are independent $\mathbb{F}$–standard Brownian motions, i.e. (i)-(ii)-(iii) of Definition 5.14 hold, and (iv) the distribution of $W_t - W_s$ is $\mathcal{N}(0, (t-s)I_n)$ for all $t > s \geq 0$, where $I_n$ is the identity matrix of $\mathbb{R}^n$.

## 5.4 Distribution of the Brownian motion

Let $W$ be a standard real Brownian motion. In this section, we list some properties of $W$ which are directly implied by its distribution.

- The Brownian motion is a martingale:

$$\mathbb{E}[W_t|\mathcal{F}_s] = W_s \quad \text{for} \quad 0 \leq s \leq t,$$

where $\mathcal{F}$ is any filtration containing the canonical filtration $\mathcal{F}^W$ of the Brownian motion. From the Jensen inequality, it follows that the squared Brownian
Figure 5.4: A sample path of the two-dimensional Brownian motion

motion $W^2$ is a submartingale:

$$\mathbb{E} [W^2_t|\mathcal{F}_s] \geq W^2_s, \quad \text{for } 0 \leq s < t.$$  

The precise departure from a martingale can be explicitly calculated

$$\mathbb{E} [W^2_t|\mathcal{F}_s] = W^2_s + (t - s), \quad \text{for } 0 \leq s < t,$$

which means that the process \{\(W^2_t - t, \ t \geq 0\)\} is a martingale.

- The Brownian motion is a Markov process, i.e.

$$\mathbb{E} [\phi (W_s, s \geq t) |\mathcal{F}_t] = \mathbb{E} [\phi (W_s, s \geq t) |W_t]$$

for every $t \geq 0$ and every bounded continuous function $\phi : C^0(\mathbb{R}_+) \rightarrow \mathbb{R}$, where $C^0(\mathbb{R}_+)$ is the set of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}$. This follows immediately from the fact that $W_s - W_t$ is independent of $\mathcal{F}_t$ for every $s \geq t$, a
consequence of the independence of the increments of the Brownian motion. We shall see later in Corollary 5.2 that the Markov property holds in a stronger sense by replacing the deterministic time \( t \) by an arbitrary stopping time \( \tau \).

- The Brownian motion is a centered Gaussian process as it follows from its definition that the vector random variable \((W_{t_1}, \ldots, W_{t_n})\) is Gaussian for every \( 0 \leq t_1 < \cdots < t_n \). Centered Gaussian processes can be characterized in terms of their covariance function. A direct calculation provides the covariance function of the Brownian motion

\[
\text{Cov}(W_t, W_s) = E[W_t W_s] = t \wedge s = \min\{t, s\}
\]

The Kolmogorov theorem provides an alternative construction of the Brownian motion as a centered Gaussian process with the above covariance function, we will not elaborate more on this and we send the interested reader to Karatzas and Shreve [16], Section 2.2.

We conclude this section by the following property which is very useful for the purpose of simulating the Brownian motion.

**Exercise 5.9** For \( 0 \leq t_1 < i < t_2 \), show that the conditional distribution of \( W_i \) given \((W_{t_1}, W_{t_2}) = (x_1, x_2)\) is Gaussian, and provided its mean and variance in closed form.

- By definition of the Brownian motion, for \( 0 \leq t < T \), the conditional distribution of the random variable \( W_T \) given \( W_t = x \) is a \( N(x, t - s) \):

\[
p(t, x, T, y)dy := \mathbb{P}[W_T \in [y, y + dy] | W_t = x] = \frac{1}{\sqrt{2\pi(T - t)}} e^{-\frac{(y-x)^2}{2(T-t)}}, dy
\]

An important observation is that this density function satisfies the heat equation for every fixed \((t, x)\):

\[
\frac{\partial p}{\partial T} = \frac{1}{2} \frac{\partial^2 p}{\partial y^2},
\]
as it can be checked by direct calculation. One can also fix $(T, y)$ and express the heat equation in terms of the variables $(t, x)$:

$$\frac{\partial p}{\partial t} + \frac{1}{2} \frac{\partial^2 p}{\partial x^2} = 0.$$  \hfill (5.2)

We next consider a function $g$ with polynomial growth, say, and we define the conditional expectation:

$$V(t, x) = \mathbb{E}[g(W_T) | W_t = x] = \int g(y)p(t, x, T, y)\,dy.$$  \hfill (5.3)

Since $p$ is $C^\infty$, it follows from the dominated convergence theorem that $V$ is also $C^\infty$. By direct differentiation inside the integral sign, it follows that the function $V$ is a solution of

$$\frac{\partial V}{\partial t} + \frac{1}{2} \frac{\partial^2 V}{\partial x^2} = 0 \quad \text{and} \quad V(T, \cdot) = g.$$  \hfill (5.4)

We shall see later in Section 6.3 that the function $V$ defined in (5.3) is the unique solution of the above linear partial differential equation in the class of polynomially growing functions.

### 5.5 Scaling, symmetry, time reversal and translation

The following easy properties follow from the properties of the centered Gaussian distribution.

**Proposition 5.6** Let $W$ be a standard Brownian motion, $t_0 > 0$, and $c > 0$. Then, so are the processes

- $\{-W_t, \ t \geq 0\}$ (symmetry),
- $\{\sqrt{c}W_t, \ t \geq 0\}$ (scaling),
- $\{W_{t_0+t} - W_{t_0}, \ t \geq 0\}$ (time translation),
- $\{W_{T-t} - W_T, \ 0 \leq t \leq T\}$ (time reversal).
Proof. Properties (i), (ii) and (iii) of Definition 5.15 are immediately checked.

Remark 5.7 For a Brownian motion $W$ in $\mathbb{R}^n$, the symmetry property of the Brownian motion extends as follows: for any $(n \times n)$ matrix $A$, with $AA^T = I_n$, the process $\{AW_t, \ t \geq 0\}$ is a Brownian motion.

Another invariance property for the Brownian motion will be obtained by time inversion in subsection 5.7 below. Indeed, the process $B$ defined by $B_0 := 0$ and $B_t := tW_{1/t}$, $t > 0$, obviously satisfies properties (ii) and (iii); property (i) will be obtained as a consequence of the law of large numbers.

We next investigate whether the translation property of the Brownian motion can be extended to the case where the deterministic time $t_0$ is replaced by some random time. The following result states that this is indeed the case when the random time is a stopping time.

Proposition 5.7 Let $W$ be a Brownian motion, and consider some stopping time $\tau$. Then, the process $B$ defined by

$$B_t := W_{t+\tau} - W_{\tau}, \ t \geq 0,$$

is a Brownian motion independent of $\mathcal{F}_\tau$.

Proof. Clearly $B_0 = 0$ and $B$ has a.s. continuous sample paths. In the rest of this proof, we show that, for $0 \leq t_1 < t_2 < t_3 < t_4$, $s > 0$, and bounded continuous functions $\phi$, $\psi$ and $f$:

$$\mathbb{E} \left[ \phi (B_{t_4} - B_{t_3}) \psi (B_{t_2} - B_{t_1}) f (W_s) 1_{s \leq \tau} \right] = \mathbb{E} \left[ \phi (W_{t_4} - W_{t_3}) \right] \mathbb{E} \left[ \psi (W_{t_2} - W_{t_1}) \right] \mathbb{E} [f (W_s) 1_{s \leq \tau}] \quad (5.5)$$

This would imply that $B$ has independent increments with the required Gaussian distribution.

Observe that we may restrict our attention to the case where $\tau$ has a finite support $\{s_1, \ldots, s_n\}$. Indeed, given that (5.5) holds for such stopping times,
one may approximate a general stopping time \( \tau \) by the decreasing sequence of stopping times \( \tau^n := ([n\tau] + 1)/n \), apply (5.5) for each \( n \geq 1 \), and pass to the limit by the monotone convergence theorem thus proving that (5.5) holds for \( \tau \).

For a stopping time \( \tau \) with finite support \( \{s_1, \ldots, s_n\} \), we have:

\[
\mathbb{E} \left[ \phi (B_{t_4} - B_{t_3}) \psi (B_{t_2} - B_{t_1}) f (W_s) 1_{\{s \leq \tau\}} \right] \\
= \sum_{i=1}^{n} \mathbb{E} \left[ \phi (B_{t_4} - B_{t_3}) \psi (B_{t_2} - B_{t_1}) f (W_s) 1_{\{s \leq \tau\}} 1_{\{\tau = s_i\}} \right] \\
= \sum_{i=1}^{n} \mathbb{E} \left[ \phi (W_{t_4} - W_{t_3}) \psi (W_{t_2} - W_{t_1}) f (W_s) 1_{\{\tau = s_i \geq s\}} \right]
\]

where we denoted \( t'_k := s_i + t_k \) for \( i = 1, \ldots, n \) and \( k = 1, \ldots, 4 \). We next condition upon \( F_{s_i} \) for each term inside the sum, and recall that \( 1_{\{\tau = s_i\}} \) is \( F_{s_i} \)-measurable as \( \tau \) is a stopping time. This provides

\[
\mathbb{E} \left[ \phi (B_{t_4} - B_{t_3}) \psi (B_{t_2} - B_{t_1}) f (W_s) 1_{\{s \leq \tau\}} \right] \\
= \sum_{i=1}^{n} \mathbb{E} \left\{ \mathbb{E} \left[ \phi (W_{t_4} - W_{t_3}) \psi (W_{t_2} - W_{t_1}) \mid F_{s_i} \right] f (W_s) 1_{\{\tau = s_i \geq s\}} \right\} \\
= \sum_{i=1}^{n} \mathbb{E} \left\{ \mathbb{E} \left[ \phi (W_{t_4} - W_{t_3}) \right] \mathbb{E} \left[ \psi (W_{t_2} - W_{t_1}) \right] f (W_s) 1_{\{\tau = s_i \geq s\}} \right\}
\]

where the last equality follows from the independence of the increments of the Brownian motion and the symmetry if the Gaussian distribution. Hence

\[
\mathbb{E} \left[ \phi (B_{t_4} - B_{t_3}) \psi (B_{t_2} - B_{t_1}) f (W_s) 1_{\{s \leq \tau\}} \right] \\
= \mathbb{E} \left[ \phi (W_{t_4} - W_{t_3}) \right] \mathbb{E} \left[ \psi (W_{t_2} - W_{t_1}) \right] \sum_{i=1}^{n} \mathbb{E} \left[ f (W_s) 1_{\{\tau = s_i \geq s\}} \right]
\]

which is exactly (5.5).

An immediate consequence of Proposition 5.7 is the strong Markov property of the Brownian motion.
Corollary 5.2 The Brownian motion satisfies the strong Markov property:

\[ \mathbb{E} [\phi (W_{s+\tau}, s \geq 0) | \mathcal{F}_\tau] = \mathbb{E} [\phi (W_{s+\tau}, s \geq 0) | W_\tau] \]

for every stopping time \( \tau \), and every bounded function \( \phi : C^0(\mathbb{R}_+) \rightarrow \mathbb{R} \).

**Proof.** Since \( B_s := W_{s+\tau} - W_\tau \) is independent of \( \mathbb{F}_\tau \) for every \( s \geq 0 \), we have

\[ \mathbb{E} [\phi (W_{s+\tau}, s \geq 0) | \mathcal{F}_\tau] = \mathbb{E} [\phi (B_s + W_\tau, s \geq 0) | \mathcal{F}_\tau] = \mathbb{E} [\phi (B_s + W_\tau, s \geq 0) | W_\tau]. \]

\( \diamond \)

We next use the symmetry property of Proposition 5.7 in order to provide explicitly the joint distribution of the Brownian motion \( W \) and the corresponding running maximum process:

\[ W^*_t := \sup_{0 \leq s \leq t} W_s, \quad t \geq 0. \]

The key-idea for this result is to make use of the Brownian motion started at the first hitting time of some level \( y \):

\[ T_y := \inf \{ t > 0 : W_t > y \}. \]

Observe that

\[ \{W^*_t \geq y\} = \{T_y \leq t\}, \]

which implies in particular a connection between the distributions of the running maximum \( W^*_t \) and the first hitting time \( T_y \).

**Proposition 5.8** Let \( W \) be a Brownian motion and \( W^* \) the corresponding running maximum process. Then, for \( t > 0 \), the random variables \( W^*_t \) and \( |W_t| \) have the same distribution, i.e.

\[ \mathbb{P} [W^*_t \geq y] = \mathbb{P} [|W_t| \geq y]. \]
Furthermore, the joint distribution of the Brownian motion and the corresponding running maximum is characterized by

\[ P[W_t \leq x, W^*_t \geq y] = P[W_t \geq 2y - x] \text{ for } y > 0 \text{ and } x \leq y. \]

**Proof.** From Exercise 5.6 and Proposition 5.7, the first hitting time \( T_y \) of the level \( y \) is a stopping time, and the process

\[ B_t := (W_{t+T_y} - W_{T_y}), \quad t \geq 0, \]

is a Brownian motion independent of \( \mathcal{F}_{T_y} \). Since \( B_t \) and \( -B_t \) have the same distribution and \( W_{T_y} = y \), we compute that

\[ P[W_t \leq x, W^*_t \geq y] = P[y + B_{t-T_y} \leq x, T_y \leq t] \]
\[ = P[y - B_{t-T_y} \leq x, T_y \leq t] \]
\[ = P[W_t \geq 2y - x, W^*_t \geq y] = P[W_t \geq 2y - x], \]

where the last equality follows from the fact that \( \{W^*_t \geq y\} \subset \{W_t \geq 2y - x\} \) as \( x \geq y \). As for the marginal distribution of the running maximum, we decompose:

\[ P[W^*_t \geq y] = P[W_t < y, W^*_t \geq y] + P[W_t \geq y, W^*_t \geq y] \]
\[ = P[W_t < y, W^*_t \geq y] + P[W_t \geq y] \]
\[ = 2P[W_t \geq y] = P[|W_t| \geq y] \]

where the two last equalities follow from the first part of this proof together with the symmetry of the Gaussian distribution. \( \Box \)

The following property is useful for the approximation of Lookback options prices by Monte Carlo simulations.

**Exercise 5.10** For a Brownian motion \( W \) and \( t > 0 \), show that

\[ P[W^*_t \geq y | W_t = x] = \exp \left( -\frac{2}{t} y(y - x) \right) \text{ for } y \geq x^+. \]
5.6 On the filtration of the Brownian motion

Because the Brownian motion has a.s. continuous sample paths, the corresponding canonical filtration is left-continuous, i.e. $\cup_{s<t} F^W_s = F^W_t$. However, $F^W$ is not right-continuous. To see this, observe that the event set \{W has a local maximum at $t$\} is in $F^W_{t+} := \cap_{s>t} F^W_s$, but is not in $F^W_t$.

This difficulty can be overcome by slightly enlarging the canonical filtration by the collection of zero-measure sets. We define $N := \{A \in \Omega : P(A) = 0\}$ and the filtration by $\tilde{F}^W_t := \sigma(F^W_t \cup N)$.

The resulting filtration $\tilde{F}^W$ is called the completed canonical filtration, and $W$ remains a Brownian motion with respect to $\tilde{F}^W$, because the four properties of the Definition 5.14 do not depend on the null sets of $\mathbb{P}$. Importantly, $\tilde{F}^W$ can be shown to be continuous. We send the interested reader to Karatzas and Shreve [16] for a rigorous treatment of this question. In the rest of these notes, we always work with the completed filtration, and since there is no ambiguity, we shall from now on omit the bar in $\tilde{F}^W$, that is, $F^W$ will denote the completed filtration.

The following result appears simple, but has far-reaching consequences for the study of local behavior of Brownian paths.

**Proposition 5.9 (Blumenthal Zero-One law)** Let $W$ be a Brownian motion and $F^W$ its completed canonical filtration. If $A \in F^B_t$ then either $P(A) = 0$ or $P(A) = 1$.

5.7 Small/large time behavior of the Brownian sample paths

The discrete-time approximation of the Brownian motion suggests that $\frac{W_t}{t}$ tends to zero at least along natural numbers, by the law of large numbers. With a little effort, we obtain the following strong law of large numbers for the Brownian motion.
Theorem 5.11 For a Brownian motion \( W \), we have
\[
\frac{W_t}{t} \rightarrow 0 \quad \mathbb{P} \text{-a.s. as } t \to \infty.
\]

Proof. We first decompose
\[
\frac{W_t}{t} = \frac{W_t - W_{\lfloor t \rfloor}}{t} + \frac{\lfloor t \rfloor}{t} \frac{W_{\lfloor t \rfloor}}{\lfloor t \rfloor}
\]
By the law of large numbers, we have
\[
\frac{W_{\lfloor t \rfloor}}{\lfloor t \rfloor} = 1 - \frac{1}{\lfloor t \rfloor} \sum_{i=1}^{\lfloor t \rfloor} (W_i - W_{i-1}) \longrightarrow 0 \quad \mathbb{P} \text{-a.s.}
\]
We next estimate that
\[
\frac{|W_t - W_{\lfloor t \rfloor}|}{t} \leq \frac{\lfloor t \rfloor}{t} \frac{|\Delta_{\lfloor t \rfloor}|}{\lfloor t \rfloor}, \quad \text{where } \Delta_n := \sup_{n-1<t\leq n} |W_t - W_{n-1}|, \quad n \geq 1.
\]
Clearly, \( \{\Delta_n, n \geq 1\} \) is a sequence of independent identically distributed random variables. The distribution of \( \Delta_n \) is explicitly given by Proposition 5.8. In particular, by a direct application of the Chebychev inequality, it is easily seen that \( \sum_{n \geq 1} \mathbb{P}[\Delta_n \geq n\varepsilon] = \sum_{n \geq 1} \mathbb{P}[\Delta_1 \geq n\varepsilon] < \infty \). By the Borel Cantelli Theorem, this implies that \( \Delta_n/n \longrightarrow 0 \quad \mathbb{P}-\text{a.s.} \)
\[
\frac{W_t - W_{\lfloor t \rfloor}}{\lfloor t \rfloor} \longrightarrow 0 \quad \mathbb{P} \text{-a.s. as } t \to \infty,
\]
and the required result follows from the fact that \( t/\lfloor t \rfloor \longrightarrow 1 \) as \( t \to \infty \). \( \diamond \)

As an immediate consequence of the law of large numbers for the Brownian motion, we obtain the invariance property of the Brownian motion by time inversion:

Proposition 5.10 Let \( W \) be a standard Brownian motion. Then the process
\[
B_0 = 0 \quad \text{and} \quad B_t := tW_{\frac{1}{t}} \text{ for } t > 0
\]
is a Brownian motion.
Proof. All of the properties (i)-(ii)-(iii) of Definition 5.15 are obvious except the sample path continuity at zero. But this is equivalent to the law of large numbers stated in Proposition 5.11.

The following two results show the complexity of the sample paths of the Brownian motion.

**Proposition 5.11** For any $t_0 \geq 0$, we have

$$\liminf_{t \to t_0} \frac{W_t - W_{t_0}}{t - t_0} = -\infty \quad \text{and} \quad \limsup_{t \to t_0} \frac{W_t - W_{t_0}}{t - t_0} = \infty.$$ almost surely.

Proof. From the invariance of the Brownian motion by time translation, it is sufficient to consider $t_0 = 0$. Observe that the random variable $\limsup_{t \to 0} \frac{W_t}{t}$ is in $\mathcal{F}^B_0$, and therefore is deterministic. Assume that $\limsup_{t \to 0} \frac{W_t}{t} = C < \infty$. By symmetry of the Brownian motion, $\liminf_{t \to 0} \frac{W_t}{t} = -C$, and therefore $\lim \frac{W_t}{\sqrt{t}} \to 0$ p.s. However, by the scaling property $\frac{W_t}{\sqrt{t}} \sim N(0, 1)$, which provides a contradiction.

**Proposition 5.12** Let $W$ be a Brownian motion in $\mathbb{R}$. Then, $\mathbb{P}$-a.s. $W$ changes sign infinitely many times in any time interval $[0, t]$, $t > 0$.

Observe that the event “$W$ changes sign infinitely many times in any time interval $[0, t]$” is in $\mathcal{F}^B_0$, and therefore its probability is either zero or one. Assume that it is zero. Then, with probability one, either $\liminf_{t \to 0} \frac{W_t}{t} = 0$ or $\limsup_{t \to 0} \frac{W_t}{t} = 0$, which is a contradiction with Proposition 5.11.

The following result shows that the sample paths of the Brownian motion are unbounded a.s., and the Brownian motion is recurrent, meaning that this process returns to the point 0 an infinite number of times.
Proposition 5.13 For a standard Brownian motion $W$, we have

$$\limsup_{t \to \infty} W_t = \infty \quad \text{and} \quad \liminf_{t \to \infty} W_t = -\infty$$

almost surely.

Proof.

From Proposition 5.10, $B_t := tW_1/t$ defines a Brownian motion. Since $W_1/t = B_{1/t}$, it follows that the behavior of $W_1/t$ for $t \searrow 0$ corresponds to the behavior of $B_u$ for $u \nearrow \infty$. The required limit result is then a restatement of Proposition 5.11 for $t_0 = 0$.

We conclude this section by stating, without proof, the law of the iterated logarithm for the Brownian motion. The interested reader may consult [16] for the proof.

Theorem 5.12 For a Brownian motion $W$, we have

$$\limsup_{t \to 0} \frac{W_t}{\sqrt{2t \ln(\ln(1/t))}} = 1 \quad \text{and} \quad \liminf_{t \to 0} \frac{W_t}{\sqrt{2t \ln(\ln(1/t))}} = -1$$

5.8 Quadratic variation

The purpose of this section is to study the variation of the Brownian motion along dyadic sequences:

$$t^n_i := i2^{-n}, \quad \text{for integers} \quad i \geq 0 \quad \text{and} \quad n \geq 1.$$ 

As we shall see shortly, the Brownian motion has infinite total variation. In particular, this implies that classical integration theories are not suitable for the case of the Brownian motion. The key-idea in order to define an integration theory with respect to the Brownian motion is that the quadratic variation is finite. In the context of these notes, we shall not present the
complete stochastic integration theory. We will restrict our attention to the quadratic variation along dyadic sequences defined by

\[ V^n_t := \sum_{\ell_n^i \leq t} |W_{t_{i+1}^n} - W_{t_i^n}|^2 \quad \text{for} \quad t > 0. \]

More generally, given a continuous adapted stochastic process \( \varphi = \{ \varphi_t, \ t \geq 0 \} \), we define:

\[ V^n_t(\varphi) := \sum_{\ell_n^i \leq t} \varphi_{t_i^n} |W_{t_{i+1}^n} - W_{t_i^n}|^2 \quad \text{for} \quad t > 0. \]

Before stating the main result of this section, we observe that the quadratic variation (along any subdivision) of a continuously differentiable function \( f \) converges to zero. Indeed,

\[ \sum_{t_i \leq t} |f(t_{i+1}) - f(t_i)|^2 \leq \| f' \|_{L^\infty([0,t])} \sum_{t_i \leq t} |t_{i+1} - t_i|^2 \to 0. \]

Because of the non-differentiability property stated in Proposition 5.11, this result does not hold for the Brownian motion:

**Proposition 5.14** Let \( W \) be a standard Brownian motion in \( \mathbb{R} \), and \( \varphi \) a continuous adapted stochastic process with \( \int_{\mathbb{R}^+} |\varphi_t|^2 dt < \infty \) a.s. Then:

\[ \mathbb{P} \left[ \lim_{n \to \infty} V^n_t(\varphi) = \int_0^t \varphi_s ds, \ for \ every \ t \geq 0 \right] = 1. \]

**Proof.** (i) We first assume that \( \mathbb{E} \left[ \int_{\mathbb{R}^+} |\varphi_t|^2 dt \right] < \infty \), a condition which will be relaxed in Step (iii) below. We first fix \( t > 0 \) and show that \( V^n_t(\varphi) \to \int_0^t \varphi_s ds \ \mathbb{P}-\text{a.s. as} \ n \to \infty \), or equivalently:

\[ \sum_{\ell_n^i \leq t} Z_i \xrightarrow{n \to \infty} 0 \ \mathbb{P} - \text{a.s. where} \ Z_i := \varphi_{t_i^n} \left( (W_{t_{i+1}^n} - W_{t_i^n})^2 - 2^{-n} \right). \]

Observe that \( \mathbb{E}[Z_i Z_j] = 0 \) for \( i \neq j \), and \( \mathbb{E}[Z_i^2] = C 2^{-2n} \mathbb{E} [ |\varphi_{t_i^n}|^2 ] \) for some constant \( C > 0 \). Then

\[ \sum_{n \leq N} \mathbb{E} \left[ \left( \sum_{\ell_n^i \leq t} Z_i \right)^2 \right] = \sum_{n \leq N} \sum_{\ell_n^i \leq t} \mathbb{E} [Z_i^2] = C \sum_{n \leq N} 2^{-n} \sum_{\ell_n^i \leq t} \mathbb{E} [ |\varphi_{t_i^n}|^2 ] 2^{-n}. \]
Since $\sum_{t_n^i \leq t} \mathbb{E} \left[ |\varphi_{t_n^i}|^2 \right] 2^{-n} \rightarrow \int_0^t \mathbb{E} \left[ |\varphi_t|^2 \right] dt$, we conclude that

$$\liminf_{N \to \infty} \mathbb{E} \left[ \sum_{n \leq N} \left( \sum_{t_n^i \leq t} Z_i \right)^2 \right] < \infty.$$ 

By Fatou’s lemma, this implies that

$$\mathbb{E} \left[ \sum_{n \geq 1} \left( \sum_{t_n^i \leq t} Z_i \right)^2 \right] \leq \liminf_{N \to \infty} \sum_{n \leq N} \mathbb{E} \left[ \left( \sum_{t_n^i \leq t} Z_i \right)^2 \right] < \infty.$$ 

In particular, this shows that the series $\sum_{n \geq 1} \left( \sum_{t_n^i \leq t} Z_i \right)^2$ is a.s. finite, and therefore $\sum_{t_n^i \leq t} Z_i \rightarrow 0 \text{ a.s. as } n \to \infty$.

(ii) From the first step of this proof, we can find a zero measure set $N_s$ for each rational number $s \in \mathbb{Q}$. For an arbitrary $t \geq 0$, let $(s_p)$ and $(s'_p)$ be two monotonic sequences of rational numbers with $s_p \nearrow t$ and $s'_p \searrow t$. Then, except on the zero-measure set $N := \bigcup_{s \in \mathbb{Q}} N_s$, it follows from the monotonicity of the quadratic variation that

$$\int_0^{s_p} \varphi_u du = \lim_{n \to \infty} V^n_{s_p}(\varphi) \leq \liminf_{n \to \infty} V^n_t(\varphi) \leq \limsup_{n \to \infty} V^n_t(\varphi) \leq \lim_{n \to \infty} V^n_{s'_p}(\varphi) = \int_0^{s'_p} \varphi_u du.$$ 

Sending $p \to \infty$ shows that $V^n_t(\varphi) \rightarrow \int_0^t \varphi_u du$ as $n \to \infty$ for every $\omega$ outside the zero-measure set $N$.

(iii) We now consider the general case where $\varphi$ is only known to satisfy $\int_{\mathbb{R}^+} |\varphi|^2 dt < \infty$ a.s. We introduce the sequence of stopping times

$$T_k := \inf \left\{ t > 0 : \int_0^t |\varphi_u|^2 du \geq k \right\}, \quad k \geq 1,$$

and we observe that $T_k \rightarrow \infty \text{ a.s. as } k \to \infty$. From the previous steps of this proof, we have $\mathbb{P}$-a.s.

$$\sum_{t_n^i \leq t \wedge T_k} \varphi_{t_n^i} \left| W_{t_{n+1}^i} - W_{t_n^i} \right|^2 = V^n_t (\varphi 1_{[0,T_k]}) \rightarrow \int_0^t \varphi_u 1_{[0,T_k]}(u) du = \int_0^{t \wedge T_k} \varphi_u du.$$
The required limit result follows immediately by sending $k \to \infty$. 

Remark 5.8 Proposition 5.14 has a natural direct extension to the multi-dimensional setting. Let $W$ be a standard Brownian motion in $\mathbb{R}^d$, and let $\varphi$ be an adapted process with $(d \times d)-$matrix values. Then, $\mathbb{P}$–a.s.

$$
\sum_{t_i^n \leq t} \text{Tr} \left[ \varphi^n_t \left( W^n_{t_{i+1}} - W^n_{t_i} \right) \left( W^n_{t_{i+1}} - W^n_{t_i} \right)^T \right] \longrightarrow \int_0^t \text{Tr} [\varphi_t] \, dt, \quad t \geq 0.
$$

We leave the verification of this result as an exercise.

Remark 5.9 Inspecting the proof of Proposition 5.14, we see that the quadratic variation along any subdivision $0 = s^n_0 < \ldots < s^n_n = t$ satisfies:

$$
\sum_{i=1}^n \left| W^n_{s^n_{i+1}} - W^n_{s^n_i} \right|^2 \longrightarrow t \quad \mathbb{P}–\text{a.s. whenever } \sum_{n \geq 1} \sup_{1 \leq i \leq n} |s^n_{i+1} - s^n_i| \longrightarrow 0.
$$

We also observe that the above convergence result holds in probability for an arbitrary partition $(s^n_i, 1 \leq i \leq n)_n$ of $[0, t]$ with mesh $\sup_{1 \leq i \leq n} |s^n_{i+1} - s^n_i|$ shrinking to zero. See e.g. Karatzas and Shreve [16], Theorem 1.5.8 for the case $\varphi \equiv 1$.

Proposition 5.14 implies, in particular, that the total variation of the Brownian motion is not finite. This follows from the inequality

$$
V^n_t \leq \max_{t_i^n \leq t} \left| W^n_{t_{i+1}} - W^n_{t_i} \right| \sum_{i \in \mathbb{N}} \left| W^n_{t_{i+1}} - W^n_{t_i} \right| 1_{\{t_i^n \leq t\}}, \quad (5.6)
$$

together with the continuity of the Brownian motion which implies that $\max_{t_i^n \leq t} \left| W^n_{t_{i+1}} - W^n_{t_i} \right| \longrightarrow 0$. For this reason, the Stieltjes theory of integration does not apply to the Brownian motion.
Chapter 6

Integration to stochastic calculus

6.1 Stochastic integration with respect to the Brownian motion

Recall from (5.6) that the total variation of the Brownian motion is infinite:

$$\lim_{n \to \infty} \sum_{t^n_i \leq t} \left| W_{t^n_{i+1}} - W_{t^n_i} \right| = \infty \quad \mathbb{P} \text{-a.s.}$$

Because of this property of the Brownian, one cannot hope to define the stochastic integral with respect to the Brownian motion pathwise. To understand this, let us forget for a moment about stochastic processes. Let \( \varphi, f : [0, 1] \to \mathbb{R} \) be continuous functions, and consider the Riemann sum:

$$S_n := \sum_{t^n_i \leq 1} \varphi (t^n_{i-1}) \left[ f (t^n_i) - f (t^n_{i-1}) \right] .$$

The, if the total variation of \( f \) is infinite, one cannot guarantee that the above sum converges for every continuous function \( \varphi \).

In order to circumvent this limitation, we shall make use of the finiteness of the quadratic variation of the Brownian motion, which allows to obtain an \( L^2 \)-definition of stochastic integration.
6.1.1 Stochastic integrals of simple processes

Throughout this section, we fix a final time $T > 0$. A process $\phi$ is called simple if there exists a strictly increasing sequence $(t_n)_{n \geq 0}$ in $\mathbb{R}$ and a sequence of random variables $(\varphi_n)_{n \geq 0}$ such that

$$\phi_t = \varphi_0 1_{\{0\}}(t) + \sum_{n=0}^{\infty} \varphi_n 1_{(t_n, t_{n+1}]}(t), \quad t \geq 0,$$

and

$$\varphi_n \text{ is } \mathcal{F}_{t_n} - \text{measurable for every } n \geq 0 \text{ and } \sup_{n \geq 0} \|\varphi_n\|_\infty < \infty.$$  

We shall denote by $\mathcal{S}$ the collection of all simple processes. For $\phi \in \mathcal{S}$, we define its stochastic integral with respect to the Brownian motion by:

$$I^0_t(\phi) := \sum_{n \geq 0} \varphi_n (W_{t \wedge t_{n+1}} - W_{t \wedge t_n}), \quad 0 \leq t \leq T. \quad (6.1)$$

By this definition, we immediately see that:

$$\mathbb{E}[I^0_t(\phi) | \mathcal{F}_s] = I^0_s(\phi) \text{ for } 0 \leq s \leq t, \quad (6.2)$$

i.e. $\{I^0_t(X), t \geq 0\}$ is a martingale. We also calculate that

$$\mathbb{E}[I^0_t(\phi)^2] = \mathbb{E}\left[\int_0^t |\phi_s|^2 ds\right] \text{ for } t \geq 0. \quad (6.3)$$

Exercise 6.11 Prove properties (6.2) and (6.3).

Our objective is to extend $I^0$ to a stochastic integral operator $I$ acting on the larger set

$$\mathbb{H}^2 := \left\{ \phi : \text{measurable, } \mathbb{F} - \text{adapted processes with } \mathbb{E}\left[\int_0^T |\phi_t|^2 dt\right] < \infty \right\},$$

which is a Hilbert space when equipped with the norm

$$\|\phi\|_{\mathbb{H}^2} := \left(\mathbb{E}\left[\int_0^T |\phi_t|^2 dt\right]\right)^{1/2}.$$

This is crucially based on the following density result.

87
Proposition 6.15 The set of simple processes \( S \) is dense in \( \mathbb{H}^2 \), i.e. for every \( \phi \in \mathbb{H}^2 \), there is a sequence \( (\phi^{(n)})_{n \geq 0} \) of processes in \( S \) such that \( \| \phi - \phi^{(n)} \|_{\mathbb{H}^2} \to 0 \) as \( n \to \infty \).

The proof of this result is reported in the Complements section 6.5.

6.1.2 Stochastic integrals of processes in \( \mathbb{H}^2 \)

We now consider a process \( \phi \in \mathbb{H}^2 \), and we intend to define the stochastic integral \( I_t(\phi) \) for every \( t \geq 0 \) by using the density of simple processes.

a. From Proposition 6.15, there is a sequence \( (\phi^{(n)})_{n \geq 0} \) which approximates \( \phi \) in the sense that \( \| \phi - \phi^{(n)} \|_{\mathbb{H}^2} \to 0 \) as \( n \to \infty \). We next observe from (6.3) that, for every \( t \geq 0 \):

\[
\| I_t^0 (\phi^{(n)}) - I_t^0 (\phi^{(m)}) \|_{L^2} = \mathbb{E} \left[ \int_0^t |\phi^{(n)}_s - \phi^{(m)}_s|^2 ds \right]
\]

converges to zero as \( n, m \to \infty \). This shows that the sequence \( (I_t^0 (\phi^{(n)}))_{n \geq 0} \) is a Cauchy sequence in \( L^2 \), and therefore

\( I_t^0 (\phi^{(n)}) \to I_t(\phi) \) in \( L^2 \) for some random variable \( I_t(\phi) \).

b. We next show that the limit \( I_t(\phi) \) does not depend on the choice of the approximating sequence \( (\phi^{(n)})_n \). Indeed, for another approximating sequence \( (\psi^{(n)})_n \) of \( \phi \), we have

\[
\| I_t^0 (\psi^{(n)}) - I_t (\phi) \|_{L^2} \leq \| I_t^0 (\psi^{(n)}) - I_t^0 (\phi^{(n)}) \|_{L^2} + \| I_t^0 (\phi^{(n)}) - I_t (\phi) \|_{L^2}
\]

\[
\to 0 \text{ as } n \to \infty.
\]

c. We finally show that the uniquely defined limit \( I_t(\phi) \) satisfies the analogue properties of (6.2)-(6.3):

\[
\mathbb{E} [I_t(\phi) | \mathcal{F}_s] = I_s(\phi) \quad \text{for } 0 \leq s \leq t \text{ and } \phi \in \mathbb{H}^2, \quad (6.4)
\]
\[
\mathbb{E} [I_t(\phi)^2] = \mathbb{E} \left[ \int_0^t |\phi_t|^2 dt \right] \quad \text{for } t \geq 0 \text{ and } \phi \in \mathbb{H}^2. \quad (6.5)
\]
To see that (6.4) holds, we directly compute with \((\phi^{(n)})_n\) an approximating sequence of \(\phi\) in \(H^2\) that
\[
\|\mathbb{E}[I_t(\phi)|\mathcal{F}_s] - I_s(\phi)\|_{L^2} \leq \|\mathbb{E}[I_t(\phi)|\mathcal{F}_s] - \mathbb{E}[I^0_t(\phi^{(n)})|\mathcal{F}_s]\|_{L^2} \\
+ \|\mathbb{E}[I^0_t(\phi^{(n)})|\mathcal{F}_s] - I_s(\phi)\|_{L^2} = \|\mathbb{E}[I_t(\phi)|\mathcal{F}_s] - \mathbb{E}[I^0_t(\phi^{(n)})|\mathcal{F}_s]\|_{L^2} \\
+ \|I^0_s(\phi^{(n)}) - I_s(\phi)\|_{L^2}
\]
by (6.2). By the Jensen inequality and the law of iterated expectations, this provides
\[
\|\mathbb{E}[I_t(\phi)|\mathcal{F}_s] - I_s(\phi)\|_{L^2} \leq \|I_t(\phi) - I^0_t(\phi^{(n)})\|_{L^2} + \|I^0_s(\phi^{(n)}) - I_s(\phi)\|_{L^2}
\]
which implies (6.4) by sending \(n\) to infinity.

As for (6.5), it follows from the \(H^2\)-convergence of \(\phi^{(n)}\) towards \(\phi\) and the \(L^2\)-convergence of \(I^0(\phi^{(n)})\) towards \(I(\phi)\), together with (6.3), that:
\[
\mathbb{E}\left[\int_0^t |\phi_s|^2 ds \right] = \lim_{n \to \infty} \mathbb{E}\left[\int_0^t |\phi^{(n)}_s|^2 ds \right] = \lim_{n \to \infty} \mathbb{E}\left[I^0(\phi^{(n)})^2\right] = \mathbb{E}\left[I(\phi)^2\right].
\]

Summarizing the above construction, we have proved the following.

**Theorem 6.13** For \(\phi \in H^2\) and \(t \geq 0\), the stochastic integral denoted
\[
I_t(\phi) := \int_0^t \phi_s dW_s
\]
is the unique limit in \(L^2\) of the sequence \((I^0_t(\phi^{(n)}))_n\) for every choice of an approximating sequence \((\phi^{(n)})_n\) of \(\phi\) in \(H^2\). Moreover, \(\{I_t(\phi), t \geq 0\}\) is a square integrable martingale satisfying the Itô isometry
\[
\mathbb{E}\left[\left(\int_0^t \phi_s dW_s\right)^2\right] = \mathbb{E}\left[\int_0^t |\phi_s|^2 ds\right], \quad t \geq 0.
\]

Observe that, since the Brownian motion has continuous trajectories, the stochastic integral of a simple process \(I^0(\phi)\) also has continuous sample paths
by the definition in (6.1). It is not clear whether this property is inherited by the extension $I$ of $I_0$ to $H^2$. The following result shows that it is indeed the case, and provides the useful Doob’s maximal inequality for martingales (6.6).

**Proposition 6.16** Let $\phi$ be a process in $H^2$, and set $M_t := \int_0^t \phi_s \cdot dW_s$, $t \geq 0$. Then $M$ has continuous sample paths a.s. and

$$
E \left[ \sup_{s \leq t} M_s \right] \leq 4E \left[ |M_t|^2 \right] = E \left[ \int_0^t |\phi_s|^2 ds \right].
$$

(6.6)

**Proof.** 1. Set $M^*_t := \sup_{s \leq t} M_s$. In this step, we prove that the maximal inequality (6.6) holds if $M$ is known to be continuous. Consider the sequence of stopping times:

$$
\theta_n := \inf \{ t \geq 0 : M^*_t \geq n \}
$$

Observing that $X^*$ is a non-decreasing process and $M_t = M^*_t$ on $\{dM^*_t > 0\}$, it follows that

$$
d(M^*_t)^2 = 2M^*_t dM^*_t = 2M_t dM^*_t = 2 [d(M_t M^*_t) - M^*_t dM_t].
$$

Then

$$
(M^*_{t \land \theta_n})^2 = 2M_{t \land \theta_n} M^*_{t \land \theta_n} - 2 \int_0^{t \land \theta_n} M^*_s dM_s
$$

Since $0 \leq M^* \leq n$ on $[0, \theta_n]$, this implies that

$$
E \left[ (M^*_{t \land \theta_n})^2 \right] = 2E \left[ M_{t \land \theta_n} M^*_{t \land \theta_n} \right] \leq 2E \left[ M^2_{t \land \theta_n} \right]^{1/2} E \left[ (M^*_{t \land \theta_n})^2 \right]^{1/2}
$$

by the Cauchy-Schwartz inequality. Hence

$$
E \left[ (M^*_{t \land \theta_n})^2 \right] \leq 4E \left[ M^2_{t \land \theta_n} \right] = E \left[ \int_0^{t \land \theta_n} |\phi_s|^2 ds \right] \leq E \left[ \int_0^t |\phi_s|^2 ds \right]
$$

and we obtain the maximal inequality by sending $n$ to infinity and using the monotone convergence theorem.
2. We now show that the process $M$ is continuous. By definition of the stochastic integral, $M_t$ is the $L^2$–limit of $M^n_t := I^0_t(\phi^n)$ for some sequence $(\phi^n)_n$ of simple processes converging to $\phi$ in $\mathbb{H}^2$. By definition of the stochastic integral of simple integrands in (6.1), notice that the process $\{M^n_t - M^m_t = I^0_t(\phi^n) - I^0_t(\phi^m), t \geq 0\}$ is a.s. continuous. We then deduce from the first step that:

$$\mathbb{E}\left[\sup_{s \leq t} |M^n_s - M^m_s|\right] \leq \mathbb{E}\left[\sup_{s \leq t} (M^n_s - M^m_s)\right] + \mathbb{E}\left[\sup_{s \leq t} (M^m_s - M^n_s)\right] \leq 8 \mathbb{E}\left[\int_0^t |\phi^n_s - \phi^m_s|^2 ds\right].$$

This shows that the sequence $(M^n)_n$ is a Cauchy sequence in the Banach space of continuous processes endowed with the norm $\mathbb{E}\left[\sup_{[0,T]} |X_s|\right]$. Then $M^n$ converges towards a continuous process $\bar{M}$ in the sense of this norm. We know however that $M^n \to M := I(\phi)$ in $\mathbb{H}^2$. By passing to subequences we may deduce that $\bar{M} = M$ is continuous.

3. From the previous step, we know that $M$ is continuous. We are then in the context of step 1, and we may deduce that the maximal inequality (6.6) holds for $M$.

**Exercise 6.12** Let $f : [0, T] \to \mathbb{R}^d$ be a deterministic function with $\int_0^T |f(t)|^2 dt < \infty$. Prove that

$$\int_0^T f(t) \cdot dW_t \text{ has distribution } \mathcal{N}\left(0, \int_0^T |f(t)|^2 dt\right).$$

**Hint:** Use the closeness of the Gaussian space.

### 6.1.3 Stochastic integration beyond $\mathbb{H}^2$

Our next task is to extend the stochastic integration to integrands $\phi$ such that

$\phi$ is $\mathbb{F}$ – adapted and $\int_0^T |\phi_t|^2 dt < \infty \ \mathbb{P}$ – a.s. for every $T > 0$. (6.7)
To do this, we consider the sequence of stopping times

$$\tau_n := \inf \left\{ t > 0 : \int_0^t |\phi_u|^2 du \geq n \right\}.$$ 

Clearly, \((\tau_n)_n\) is a non-decreasing sequence of stopping times and

$$\tau_n \to \infty \ P \text{-a.s. when } n \to \infty.$$ 

For fixed \(n > 0\), the stopped process \(\phi^n_t := \phi_t 1_{t \leq \tau_n}\) is in \(H^2\). Then the stochastic integral \(I_t(\phi^n)\) is well-defined by Theorem 6.13. Since \((\tau_n)_n\) is a non-decreasing and \(P[\tau_n \geq t \text{ for some } n \geq 1] = 1\), it follows that the limit

$$I_t(\phi) := (\text{a.s.}) \lim_{n \to \infty} I_t(\phi^n) \quad (6.8)$$

exists (in fact \(I_t(\phi) = I_t(\phi^n)\) for \(n\) sufficiently large, a.s.).

**Remark 6.10** The above extension of the stochastic integral to integrands satisfying (6.7) does not imply that \(I_t(\phi)\) satisfies properties (6.4) and (6.5). This issue will be further developed in the next subsection. The maximal inequality (6.6) is also lost since the stochastic integral is not known to be in \(L^2\). However the continuity property of the stochastic integral is conserved because it is a pathwise property which is consistent with the pathwise definition of (6.8).

As a consequence of this remark, when the integrand \(\phi\) satisfies (6.7) but is not in \(H^2\), the stochastic integral fails to be a martingale, in general. This leads us to the notion of local martingale.

**Definition 6.17** An \(\mathbb{F}\)-adapted process \(M = \{M_t, t \geq 0\}\) is a local martingale if there exists a sequence of stopping times \((\tau_n)_{n \geq 0}\) (called a localizing sequence) such that \(\tau_n \to \infty \ P\text{-a.s. as } n \to \infty\), and the stopped process \(M^{\tau_n} = \{M_{t \wedge \tau_n}, t \geq 0\}\) is a martingale for every \(n \geq 0\).

**Proposition 6.17** Let \(\phi\) be a process satisfying (6.7). Then, for every \(T > 0\), the process \(\{I_t(\phi), 0 \leq t \leq T\}\) is a local martingale.
Proof. The above defined sequence $\tau_n$ is easily shown to be a localizing sequence. The result is then a direct consequence of the martingale property of the stochastic integral of a process in $\mathbb{H}^2$.

An example of local martingale which fails to be a martingale will be given in the next subsection.

### 6.2 Itô’s formula

The purpose of this section is to prove the Itô formula for the change of variable. Given a smooth function $f(t, x)$, we will denote by $f_t$, $f_x$ and $f_{xx}$, the partial gradients with respect to $t$, to $x$, and the partial Hessian matrix with respect to $x$.

**Theorem 6.14** Let $f : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}$ be $C^{1,2}([0,T],\mathbb{R}^d)$. Then, with probability 1, we have:

$$f(T, W_T) = f(0, 0) + \int_0^T f_x(t, W_t) \cdot dW_t + \int_0^T \left( f_t + \frac{1}{2} \text{Tr}[f_{xx}] \right)(t, W_t) dt$$

for every $T \geq 0$.

**Proof.** We first fix $T > 0$, and we show that the above Itô’s formula holds with probability 1. By possibly adding a constant to $f$ we may assume that $f(0,0) = 0$. Let $t^n_i := i2^{-n}$ for $i \geq 1$, and $n(T) := \lfloor T2^n \rfloor$, so that $t^n_{n(T)} \leq T < t^n_{n(T)+1}$.

1.a We first decompose

$$f\left(t^n_{n(T)+1}, W^n_{t^n_{n(T)+1}}\right) = \sum_{t^n_i \leq T} \left[ f\left(t^n_i, W^n_{t^n_i+1}\right) - f\left(t^n_i, W^n_{t^n_i+1}\right) \right]$$

$$+ \sum_{t^n_i \leq T} \left[ f\left(t^n_i, W^n_{t^n_i+1}\right) - f\left(t^n_i, W^n_{t^n_i+1}\right) \right].$$
By a Taylor expansion, this provides:

\[ I^n_T(f_x) := \sum_{t^n_i \leq T} f_x(t^n_i, W^n_{t^n_{i+1}}) \cdot (W^n_{t^n_{i+1}} - W^n_{t^n_i}) \]

\[ = f\left(t^n_{n(T)+1}, W^n_{n(T)+1}\right) - \sum_{t^n_i \leq T} f_t\left(t^n_i, W^n_{t^n_{i+1}}\right) 2^{-n} \]

\[ - \frac{1}{2} \sum_{t^n_i \leq T} \text{Tr} \left[ (f_{xx}\left(t^n_i, \xi^n_i\right) - f_{xx}\left(t^n_i, W^n_{t^n_i}\right)) \Delta^n_i W \Delta^n_i W^T \right] \]

\[ - \frac{1}{2} \sum_{t^n_i \leq T} \text{Tr} \left[ f_{xx}\left(t^n_i, W^n_{t^n_i}\right) \Delta^n_i W \Delta^n_i W^T \right], \quad (6.9) \]

where \( \Delta^n_i W := W^n_{t^n_{i+1}} - W^n_{t^n_i} \), \( \tau^n_i \) is a random variable with values in \([t^n_i, t^n_{i+1}]\), and \( \xi^n_i = \lambda^n_i W^n_{t^n_i} + (1 - \lambda^n_i) W^n_{t^n_{i+1}} \) for some random variable \( \lambda^n_i \) with values in \([0, 1]\).

**1.b** Since a.e. sample path of the Brownian motion is continuous, and therefore uniformly continuous on the compact interval \([0, T + 1]\), it follows that

\[ f\left(t^n_{n(T)+1}, W^n_{n(T)+1}\right) \rightarrow f(T, W_T), \quad \mathbb{P} - \text{a.s.} \]

\[ \sum_{t^n_i \leq T} f_t\left(t^n_i, W^n_{t^n_{i+1}}\right) 2^{-n} \rightarrow \int_0^T f_t(t, W_t) dt \quad \mathbb{P} - \text{a.s.} \]

and with Proposition 5.14:

\[ \sum_{t^n_i \leq T} \text{Tr} \left[ (f_{xx}\left(t^n_i, \xi^n_i\right) - f_{xx}\left(t^n_i, W^n_{t^n_i}\right)) \Delta^n_i W \Delta^n_i W^T \right] \rightarrow 0 \quad \mathbb{P} - \text{a.s.} \]

Finally, from Proposition 5.14 and its multi-dimensional extension in Remark 5.8, the last term in the decomposition (6.9):

\[ \sum_{t^n_i \leq T} \text{Tr} \left[ f_{xx}\left(t^n_i, W^n_{t^n_i}\right) \Delta^n_i W \Delta^n_i W^T \right] \rightarrow \int_0^T \text{Tr} \left[ f_{xx}(t, W_t) \right] dt \quad \mathbb{P} - \text{a.s.} \]

**1.c** In order to complete the proof of Itô’s formula for fixed \( T > 0 \), it remains to prove that

\[ I^n_T(f_x) \rightarrow \int_0^T f_x(t, W_t) \cdot dW_t \quad \mathbb{P} - \text{a.s.} \quad \text{along some subsequence(6.10)} \]
Notice that $I_T^n(f_x) = I_T^0(\phi^{(n)})$ where $\phi^{(n)}$ is the simple process defined by
\[
\phi^{(n)}_t = \sum_{t_i^n \leq t} f_x(t_i^n, W_{t_i^n}) 1_{[t_i^n, t_{i+1}^n)}(t), \quad t \geq 0.
\]
Since $\phi^{(n)} \to \phi$ in $\mathbb{H}^2$ with $\phi_t := f_x(t, W_t)$, it follows from the definition of the stochastic integral in Theorem 6.13 that $I_T^n(f_x) \to I_T(\phi)$ in $L^2$, and (6.10) follows from the fact that the $L^2$ convergence implies the a.s. convergence along some subsequence.

2. From the first step, we have the existence of subsets $N_t \subset \mathcal{F}$ for every $t \geq 0$ such that $\mathbb{P}[N_t] = 0$ and the Itô’s formula holds on $N_t^c$, the complement of $N_t$. Of course, this implies that the Itô’s formula holds on the complement of the set $N := \bigcup_{t \geq 0} N_t$. But this does not complete the proof of the theorem as this set is a non-countable union of zero measure sets, and is therefore not known to have zero measure. We therefore appeal to the continuity of the Brownian motion and the stochastic integral, see Proposition 6.16. By usual approximation along rational numbers, it is easy to see that, with probability 1, the Itô formula holds for every $T \geq 0$. ♦

Remark 6.11 Since, with probability 1, the Itô formula holds for every $T \geq 0$, it follows that the Itô’s formula holds when the deterministic time $T$ is replaced by a random time $\tau$.

6.2.1 Extension to Itô processes

We next provide Itô’s formula for a general Itô process, that is an $\mathbb{R}^n$-valued process defined by:
\[
X_t := X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \quad t \geq 0,
\]
where $\mu$ and $\sigma$ are adapted processes with values in $\mathbb{R}^n$ and $\mathcal{M}_\mathbb{R}(n, d)$, respectively and satisfying
\[
\int_0^T |\mu_s| ds + \int_0^T |\sigma_s|^2 ds < \infty \quad \text{a.s.}
\]
Observe that stochastic integration with respect to the Itô process $X$ reduces to the stochastic integration with respect to $W$: for any $\mathbb{F}$–adapted $\mathbb{R}^n$–valued process $\phi$ with

$$\int_0^T |\sigma_t^T \phi_t|^2 dt + \int_0^T |\phi_t \cdot \mu_t| dt < \infty, \text{a.s.}$$

then

$$\int_0^T \phi_t \cdot dX_t = \int_0^T \phi_t \cdot dt + \int_0^T \phi_t \cdot \sigma_t dW_t,$$

$$= \int_0^T \phi_t \cdot dt + \int_0^T \sigma_t^T \phi_t \cdot dW_t.$$ 

**Theorem 6.15** Let $f : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}$ be $C^{1,2}([0,T], \mathbb{R}^n)$. Then, with probability 1, we have:

$$f(T, X_T) = f(0, 0) + \int_0^T f_x(t, X_t) \cdot dX_t + \int_0^T \left( f_t(t, X_t) + \frac{1}{2} \text{Tr}[f_{xx}(t, X_t)\sigma_t^T\sigma_t^T] \right) dt$$

for every $T \geq 0$.

**Proof.** Let $\theta_N := \inf \left\{ t : \max \left( |X_t - X_0|, \int_0^t \sigma_s^2 ds, \int_0^t |\mu_s| ds \right) \geq N \right\}$. Obviously, $\theta_N \rightarrow \infty$ a.s. when $N \rightarrow \infty$, and it is sufficient to prove Itô’s formula on $[0, \theta_N]$, since any $t \geq 0$ can be reached by sending $N$ to infinity. In view of this, we may assume without loss of generality that $X$, $\int_0^t \mu_s ds$, $\int_0^t \sigma_s^2 ds$ are bounded and that $f$ has compact support. We next consider an approximation of the integrals defining $X_t$ by step functions which are constant on intervals of time $[t_{i-1}, t_i)$ for $i = 1, \ldots, n$, and we denote by $X_n$ the resulting simple approximating process. Notice that Itô’s formula holds true for $X^n$ on each interval $[t_{i-1}, t_i)$ as a direct consequence of Theorem 6.14. The proof of the theorem is then concluded by sending $n$ to infinity, and using as much as needed the dominated convergence theorem. 

**Exercise 6.13** Let $W$ be a Brownian motion and consider the process

$$X_t := X_0 + bt + \sigma W_t, \quad t \geq 0,$$
where $b$ is a vector in $\mathbb{R}^d$ and $\sigma$ is an $(d \times d)$-matrix. Let $f$ be a $C^{1,2}(\mathbb{R}_+, \mathbb{R}^d)$ function. Show that

$$df(t, X_t) = \frac{\partial f}{\partial t}(t, X_t)dt + \frac{\partial f}{\partial x}(t, X_t) \cdot dX_t + \frac{1}{2} \text{Tr} \left[ \frac{\partial^2 f}{\partial x \partial x^T}(t, X_t)\sigma \sigma^T \right].$$

We conclude this section by providing an example of local martingale which fails to be a martingale, although positive and uniformly integrable. This shows the limitation of the discrete-time result of Lemma 3.1.

**Example 6.14** Let $W$ be a Brownian motion in $\mathbb{R}^d$.

- In the one-dimensional case $d=1$, we have $\mathbb{P}[|W_t| > 0, \forall t > 0] = 0$, see also Proposition 5.12.
- When $d \geq 2$, the situation is drastically different as it can be proved that $\mathbb{P}[|W_t| > 0, \text{ for every } t > 0] = 1$, see e.g. Karatzas and Shreve [16] Proposition 3.22 p161. In words, this means that the Brownian motion never returns to the origin $\mathbb{P}$-a.s. Then, for a fixed $t_0 > 0$, the process

$$X_t := |W_{t_0+t}|^{-1}, \quad t \geq 0,$$

is well-defined, and it follows from Itô’s formula that

$$dX_t = X_t^3 \left( \frac{1}{2} (3 - d)dt - W_t \cdot dW_t \right).$$

- We now consider the special case $d = 3$. By the previous Proposition 6.17, it follows from Itô’s formula that $X$ is a local martingale. However, by the scaling property of the Brownian motion, we have

$$\mathbb{E}[X_t] = \sqrt{\frac{t_0}{t+t_0}} \mathbb{E}[X_0], \quad t \geq 0,$$

so that $X$ has a non-constant expectation and can not be a martingale.

- Passing to the polar coordinates, we calculate directly that

$$\mathbb{E} \left[ X_t^2 \right] = (2\pi (t_0 + t))^{-3/2} \int |x|^{-2} e^{-|x|^2/(2(t_0+t))} dx$$

$$= (2\pi (t_0 + t))^{-3/2} \int r^{-2} e^{-r^2/(2(t_0+t))} 4\pi r^2 dr$$

$$= \frac{1}{t_0 + t} \leq \frac{1}{t_0} \text{ for every } t \geq 0.$$
This shows that \( \sup_{t \geq 0} \mathbb{E} [X_T^2] < \infty \). In particular, \( X \) is uniformly integrable.

### 6.3 The Feynman-Kac representation formula

In this section, we consider the following linear partial differential equation

\[
\frac{\partial v}{\partial t} + b \cdot \frac{\partial v}{\partial x} + \frac{1}{2} \text{Tr} \left[ \frac{\partial^2 v}{\partial x \partial x^T} \sigma \sigma^T \right] - kv + f = 0, \quad t < T, \ x \in \mathbb{R}^d \tag{6.11}
\]

\[
v(T, .) = g \tag{6.12}
\]

where \( g \) is a given function from \( \mathbb{R}^d \) to \( \mathbb{R} \), \( k \) and \( f \) are functions from \( [0, T] \times \mathbb{R}^d \) to \( \mathbb{R} \), \( b \) is a constant vector in \( \mathbb{R}^d \), and \( \sigma \) is a constant \( d \times d \) matrix. This is the so-called Dirichlet problem. The following result provides a stochastic representation for the solution of this purely deterministic problem.

**Theorem 6.16** Assume that the function \( k \) is uniformly bounded from below, and \( f \) has polynomial growth in \( x \) uniformly in \( t \). Let \( v \) be a \( C^{1,2} \) \(([0, T), \mathbb{R}^d)\) solution of (6.11) with polynomial growth in \( x \) uniformly in \( t \). Then

\[
v(t, x) = \mathbb{E} \left[ \int_t^T \beta_s^{tx} f(s, X_s^{tx}) ds + \beta_T^{tx} g(X_T^{tx}) \right], \quad t \leq T, \ x \in \mathbb{R}^d,
\]

where \( X_s^{tx} := x + b(s - t) + \sigma(W_s - W_t) \) and \( \beta_s^{tx} := e^{-\int_t^s k(u, X_u^{tx}) du} \) for \( t \leq s \leq T \).

**Proof.** We first introduce the sequence of stopping times

\[
\tau_n := T \wedge \inf \{ s > t : |X_s^{tx} - x| \geq n \},
\]

and we observe that \( \tau_n \rightarrow T \mathbb{P}-a.s. \) Since \( v \) is smooth, it follows from Itô's formula that for \( t \leq s < T \):

\[
d \left( \beta_s^{tx} v(s, X_s^{tx}) \right) = \beta_s^{tx} \left( -kv + \frac{\partial v}{\partial t} + b \cdot \frac{\partial v}{\partial x} + \frac{1}{2} \text{Tr} \left[ \frac{\partial^2 v}{\partial x \partial x^T} \sigma \sigma^T \right] \right) (s, X_s^{tx}) ds
\]

\[
+ \beta_s^{tx} \frac{\partial v}{\partial x} (s, X_s^{tx}) \sigma dW_s
\]

\[
= \beta_s^{tx} \left( -f(s, X_s^{tx}) ds + \frac{\partial v}{\partial x} (s, X_s^{tx}) \sigma dW_s \right).
\]

98
by the PDE satisfied by $v$ in (6.11). Then:

$$
\mathbb{E} \left[ \beta_{\tau_n}^{t,x} v(\tau_n, X_{\tau_n}^{t,x}) \right] - v(t, x) \\
= \mathbb{E} \left[ \int_t^{\tau_n} \beta_s^{t,x} \left( -f(s, X_s) ds + \frac{\partial v}{\partial x}(s, X_s^{1,x}) \sigma dW_s \right) \right].
$$

Now observe that the integrands in the stochastic integral is bounded by definition of the stopping time $\tau_n$ and the smoothness of $v$. Then the stochastic integral has zero mean, and we deduce that

$$
v(t, x) = \mathbb{E} \left[ \int_t^{\tau_n} \beta_s^{t,x} f(s, X_s^{t,x}) ds + \beta_{\tau_n}^{t,x} v(\tau_n, X_{\tau_n}^{t,x}) \right].
$$

(6.13)

Since $\tau_n \to T$ and the Brownian motion has continuous sample paths $\mathbb{P}$–a.s. it follows from the continuity of $v$ that, $\mathbb{P}$–a.s.

$$
\int_t^{\tau_n} \beta_s^{t,x} f(s, X_s^{t,x}) ds + \beta_{\tau_n}^{t,x} v(\tau_n, X_{\tau_n}^{t,x})
\xrightarrow{n \to \infty} \int_t^{T} \beta_s^{t,x} f(s, X_s^{t,x}) ds + \beta_{T}^{t,x} v(T, X_T^{t,x})
= \int_t^{T} \beta_s^{t,x} f(s, X_s^{t,x}) ds + \beta_{T}^{t,x} g(X_T^{t,x})
$$

(6.14)

by the terminal condition satisfied by $v$ in (6.11). Moreover, since $k$ is bounded from below and the functions $f$ and $v$ have polynomial growth in $x$ uniformly in $t$, we have

$$
\left| \int_t^{\tau_n} \beta_s^{t,x} f(s, X_s^{t,x}) ds + \beta_{\tau_n}^{t,x} v(\tau_n, X_{\tau_n}^{t,x}) \right| \leq C \left( 1 + \sup_{t \leq T} |W_t| \right).
$$

By Proposition 5.8, the latter bound is integrable, and we deduce from the dominated convergence theorem that the convergence in (6.14) holds in $L^1(\mathbb{P})$, proving the required result by taking limits in (6.13). ♦

The above Feynman-Kac representation formula can be extended to the case where the coefficients $b$ and $\sigma$ of the partial differential equation (6.11) are functions of $(t, x)$. In this case, the process $X$ in the representation formula is defined as the solution of a stochastic differential equation.
We also observe that the Feynman-Kac representation formula has an important numerical implication. Indeed it opens the door to the use of Monte Carlo methods in order to obtain a numerical approximation of the solution of the partial differential equation (6.11). For sake of simplicity, we provide the main idea in the case \( f = k = 0 \). Let \((X^{(1)}, \ldots, X^{(k)})\) be an iid sample drawn in the distribution of \( X_{T}^{t,x} \), and compute the mean:

\[
\hat{v}_k(t,x) := \frac{1}{k} \sum_{i=1}^{k} g(X^{(i)}).
\]

By the Law of Large Numbers, it follows that \( \hat{v}_k(t,x) \to v(t,x) \) \( \mathbb{P} \)-a.s. Moreover the error estimate is provided by the Central Limit Theorem:

\[
\sqrt{k} (\hat{v}_k(t,x) - v(t,x)) \xrightarrow{k \to \infty} \mathcal{N} \left( 0, \text{Var} \left[ g \left( X_{T}^{t,x} \right) \right] \right)
\]

in distribution, and is remarkably independent of the dimension \( d \) of the variable \( X \)!

### 6.4 The Cameron-Martin change of measure

Let \( N \) be a Gaussian random variable with mean zero and unit variance. The corresponding probability density function is

\[
\frac{\partial}{\partial x} \mathbb{P}[N \leq x] = f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}.
\]

For any constant \( a \in \mathbb{R} \), the random variable \( N + a \) is Gaussian with mean \( a \) and unit variance with probability density function

\[
\frac{\partial}{\partial x} \mathbb{P}[N + a \leq x] = f(x-a) = f(x)e^{ax-a^2/2}, \quad x \in \mathbb{R}.
\]

Then, for every (at least bounded) function \( \psi \), we have

\[
\mathbb{E} [\psi(N + a)] = \int \psi(x)f(x)e^{ax-a^2/2}dx = \mathbb{E} \left[ e^{aN-a^2/2} \psi(N) \right].
\]

This easy result can be translated in terms of a change of measure. Indeed, since the random variable \( e^{aN-a^2/2} \) is positive and integrates to 1, we may
introduce the equivalent measure $Q := e^{aN - \frac{a^2}{2}} \cdot \mathbb{P}$. Then, the above equality says that the $Q$--distribution of $N$ coincides with the $\mathbb{P}$--distribution of $N + a$, i.e.

under $Q$, $N - a$ is distributed as $\mathcal{N}(0, 1)$.

The purpose of this section is to extend this result to a Brownian motion $W$ in $\mathbb{R}^d$. Let $h : [0, T] \rightarrow \mathbb{R}^d$ be a deterministic function in $\mathbb{L}^2$, i.e. $\int_0^T |h(t)|^2 dt < \infty$. From Theorem 6.13, the stochastic integral

$$N := \int_0^T h(t) \cdot dW_t$$

is well-defined as the $\mathbb{L}^2$--limit of the stochastic integral of some $\mathbb{H}^2$--approximating simple function. In particular, since the space of Gaussian random variables is closed, it follows that

$N$ is distributed as $\mathcal{N}\left(0, \int_0^T |h(t)|^2 dt\right)$,

and we may define an equivalent probability measure $Q$ by:

$$\frac{dQ}{d\mathbb{P}} := e^{\int_0^T h(t) \cdot dW_t - \frac{1}{2} \int_0^T |h(t)|^2 dt}.$$  \hspace{1cm} (6.15)

**Theorem 6.17** For a Brownian motion $W$ in $\mathbb{R}^d$, let $Q$ be the probability measure equivalent to $\mathbb{P}$ defined by (6.15). Then, the process

$$B_t := W_t - \int_0^t h(u) \, du, \quad t \in [0, T],$$

is a Brownian motion under $Q$.

**Proof.** We first observe that $B_0 = 0$ and $B$ has a.s. continuous sample paths. It remains to prove that, for $0 \leq s < t$, $B_t - B_s$ is independent of $\mathcal{F}_s$ and distributed as a centered Gaussian with variance $t - s$. To do this, we compute the $Q$--Laplace transform

$$E^Q \left[ e^{\lambda (W_t - W_s)} \mid \mathcal{F}_s \right] = E \left[ e^{\int_s^t h(u) \cdot dW_u - \frac{1}{2} \int_s^t |h(u)|^2 du} e^{\lambda (W_t - W_s)} \mid \mathcal{F}_s \right]$$

$$= e^{-\frac{1}{2} \int_s^t |h(u)|^2 du} E \left[ e^{\int_s^t (h(u) + \lambda) \cdot dW_u} \mid \mathcal{F}_s \right]$$

101
Since the random variable \( \int_s^t (h(u) + \lambda) \cdot dW_u \) is a centered Gaussian with variance \( \int_s^t |h(u) + \lambda|^2 du \), independent of \( \mathcal{F}_s \), this provides:

\[
\mathbb{E}^\mathbb{Q} \left[ e^\lambda (W_t - W_s) \mid \mathcal{F}_s \right] = e^{-\frac{1}{2} \int_s^t |h(u)|^2 du} e^{\frac{1}{2} \int_s^t |h(u) + \lambda|^2 du} \\
= e^{\frac{1}{2} \lambda^2 (t-s) + \lambda \int_s^t h(u) du}.
\]

This shows that \( W_t - W_s \) is independent of \( \mathcal{F}_s \) and is distributed as a Gaussian with mean \( \int_s^t h(u) du \) and variance \( t - s \), i.e. \( B_t - B_s \) is independent of \( \mathcal{F}_s \) and is distributed as a centered Gaussian with variance \( t - s \).

Let us observe that the above result can be extended to the case where \( h \) is an adapted stochastic process satisfying some convenient integrability conditions. This is the so-called Girsanov theorem. In the context of these lectures, we shall only need the Gaussian framework of the Cameron-Martin formula.

### 6.5 Complement: density of simple processes in \( \mathbb{H}^2 \)

The proof of Proposition 6.15 is a consequence of the following Lemmas. Throughout this section, \( t^n_i := i2^{-n}, i \geq 0 \) is the sequence of dyadic numbers.

**Lemma 6.4** Let \( \phi \) be a bounded \( \mathbb{F} \)-adapted process with continuous sample paths. Then \( \phi \) can be approximated by a sequence of simple processes in \( \mathbb{H}^2 \).

**Proof.** Define the sequence

\[
\phi^{(n)}_t := \phi_0 1_{(0)}(t) + \sum_{t^n_i \leq T} \phi_{t^n_i} 1_{(t^n_i, t^n_{i+1})}(t), \quad t \leq T.
\]

Then, \( \phi^{(n)} \) is a simple process for each \( n \geq 1 \). By the dominated convergence theorem, \( \mathbb{E} \left[ \int_0^T |\phi^{(n)}_t - \phi_t|^2 \right] \to 0 \) as \( n \to \infty \).
Lemma 6.5 Let $\phi$ be a bounded $\mathbb{F}-$progressively measurable process. Then $\phi$ can be approximated by a sequence of simple processes in $\mathbb{H}^2$.

Proof. Notice that the process
\[ Y_t^{(k)} := k \int_{0 \vee (t-k-1)}^{t \wedge T} X_s ds, \quad t \geq 0, \]
is progressively measurable as the difference of two adapted continuous processes, see Proposition 5.5, and satisfies
\[ \| Y^{(k)} - X \|_{\mathbb{H}^2} \to 0 \quad \text{as} \quad k \to \infty, \tag{6.16} \]
by the dominated convergence theorem. For each $k \geq 1$, we can find by Lemma 6.4 a sequence $(Y^{(k,n)})_{n \geq 0}$ of simple processes such that $\| Y^{(k,n)} - Y^{(k)} \|_{\mathbb{H}^2} \to 0$ as $n \to \infty$. Then, for each $k \geq 0$, we can find $n_k$ such that
the process $X^{(k)} := Y^{(k,n_k)}$ satisfies $\| X^{(k)} - X \|_{\mathbb{H}^2} \to 0$ as $k \to \infty$.

\[ \diamond \]

Lemma 6.6 Let $\phi$ be a bounded measurable and $\mathbb{F}-$adapted process. Then $\phi$ can be approximated by a sequence of simple processes in $\mathbb{H}^2$.

Proof. In the present setting, the process $Y^{(k)}$, defined in the proof of the previous Lemma 6.5, is measurable but is not known to be adapted. For each $\varepsilon > 0$, there is an integer $k \geq 1$ such that $Y^\varepsilon := Y^{(k)}$ satisfies $\| Y^\varepsilon - X \|_{\mathbb{H}^2} \leq \varepsilon$. Then, with $X_t = X_0$ for $t \leq 0$:
\[ \| X - X_{-h} \|_{\mathbb{H}^2} \leq \| X - Y^\varepsilon \|_{\mathbb{H}^2} + \| Y^\varepsilon - Y^\varepsilon_{,-h} \|_{\mathbb{H}^2} + \| Y^\varepsilon_{,-h} - X_{-h} \|_{\mathbb{H}^2} \leq 2\varepsilon + \| Y^\varepsilon - Y^\varepsilon_{,-h} \|_{\mathbb{H}^2}. \]

By the continuity of $Y^\varepsilon$, this implies that
\[ \limsup_{h \searrow 0} \| X - X_{-h} \|_{\mathbb{H}^2} \leq 4\varepsilon^2. \tag{6.17} \]
We now introduce
\[ \phi_n(t) := 1_{\{0\}}(t) + \sum_{i \geq 1} t^n 1_{(t^n_{i-1}, t^n_i]}, \]
and
\[ X^{(n,s)}_t := X_{\phi_n(t-s) + s}, \quad t \geq 0, s \in (0, 1]. \]
Clearly \( X^{n,s} \) is a simple adapted process, and
\[
\mathbb{E} \left[ \int_0^T \int_0^1 |X^{(n,s)}_t - X_t|^2 ds dt \right] = 2^n \mathbb{E} \left[ \int_0^T \int_0^{2^{-n}} |X_t - X_{t-h}|^2 dh dt \right] = 2^n \int_0^{2^{-n}} \mathbb{E} \left[ \int_0^T |X_t - X_{t-h}|^2 dt \right] dh \leq \max_{0 \leq h \leq 2^{-n}} \mathbb{E} \left[ \int_0^T |X_t - X_{t-h}|^2 dt \right]
\]
which converges to zero as \( n \to \infty \) by (6.17). Hence
\[ X^{(n,s)}(\omega) \to X_t(\omega) \text{ for almost every } (s, t, \omega) \in [0, 1] \times [0, T] \times \Omega, \]
and the required result follows from the dominated convergence theorem.
\[ \diamond \]

**Lemma 6.7** The set of simple processes \( S \) is dense in \( \mathbb{H}^2 \).

**Proof.** We only have to extend Lemma 6.6 to the case where \( \phi \) is not necessarily bounded. This is easily achieved by applying Lemma 6.6 to the bounded process \( \phi \wedge n \), for each \( n \geq 1 \), and passing to the limit as \( n \to \infty \).
\[ \diamond \]
Chapter 7

Financial markets in continuous time

Let \( T \) be a finite horizon, and \((\Omega, \mathcal{F}, P)\) be a complete probability space supporting a Brownian motion \( W = \{(W^1_t, \ldots, W^d_t), 0 \leq t \leq T\} \) with values in \( \mathbb{R}^d \). We denote by \( \mathbb{F} = \{\mathcal{F}_t, 0 \leq t \leq T\} \) the (continuous) canonical completed filtration of \( W \), i.e. the canonical filtration augmented by zero measure sets. For simplicity, we shall always assume that \( \mathcal{F} = \mathcal{F}_T \).

The financial market consists in \( d + 1 \) assets:

(i) The first asset \( S^0 \) is non-risky, and is defined by

\[
S^0_t = \exp \left( \int_0^t r(u) du \right), \quad 0 \leq t \leq T,
\]

where \( r : [0, T] \rightarrow \mathbb{R}_+ \) is a deterministic instantaneous interest rate function satisfying

\[
\int_0^T r(t) dt < \infty.
\]

(ii) The \( d \) remaining assets \( S^i, i = 1, \ldots, d \), are risky assets with price processes:

\[
S^i_t = S^i_0 \exp \left( \int_0^t \left( b^i(u) - \frac{1}{2} \sum_{j=1}^d |\sigma^{ij}(t)|^2 \right) du + \sum_{j=1}^d \int_0^t \sigma^{ij}(t) dW^j_t \right), \quad t \geq 0,
\]

where \( b^i \) and \( \sigma^{ij} \) are the drift and diffusion coefficients, respectively.
for $1 \leq i \leq d$. Here $b = (b^1, \ldots, b^d)$ and $\sigma = (\sigma^{ij})$ are deterministic functions on $[0, T]$ with values respectively in $\mathbb{R}^d$ and $\mathcal{M}_\mathbb{R}(d, d)$, the set of $d \times d$ matrices with real coefficients. In order for the above stochastic integral to be well-defined, we assume that

$$\int_0^T (|b(t)| + |\sigma(t)|^2) \, dt < \infty.$$  

By a direct application of Itô’s formula, it follows that:

$$dS_t^i = S_t^i \left( b^i(t) \, dt + \sum_{j=1}^d \sigma^{ij}(t) \, dW_t^j \right),$$

or in vector notations:

$$dS_t = \text{diag}[S_t] (b(t) \, dt + \sigma(t) \, dW_t) .$$

where, for $z \in \mathbb{R}^n$, $\text{diag}[z]$ denotes the diagonal matrix with diagonal elements $z^i$. We assume that the $\mathcal{M}_\mathbb{R}(d, d)$—matrix $\sigma(t)$ is invertible for every $t \in [0, T]$, and we introduce the function

$$\lambda^0(t) := \sigma(t)^{-1} (b(t) - r(t)), \quad 0 \leq t \leq T,$$

called the risk premium. We shall frequently make use of the discounted process

$$\tilde{S}_t := \frac{S_t}{S^0} = S_t \exp \left( - \int_0^t r(u) \, du \right),$$

where $S = (S^1, \ldots, S^d)$. Using the above matrix notations, the dynamics of the process $\tilde{S}$ are given by in

$$d\tilde{S}_t = \text{diag}[\tilde{S}_t] \{(b(t) - r(t) \mathbf{1}) \, dt + \sigma(t) \, dW_t \} = \text{diag}[\tilde{S}_t] \sigma(t) (\lambda^0(t) \, dt + dW_t)$$

where $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^d$. 

106
7.1 Portfolio and wealth process

A portfolio strategy is an \( F - \)adapted process \( \pi = \{\pi_t, \ 0 \leq t \leq T\} \) with values in \( \mathbb{R}^d \). For \( 1 \leq i \leq n \) and \( 0 \leq t \leq T \), \( \pi^i_t \) is the amount (in Euros) invested in the risky asset \( S^i \).

As in the discrete-time financial market, we shall consider self-financing strategies. Let \( X_t \) denote the wealth process at time \( t \) induced by the portfolio strategy \( \pi \). Then, the amount invested in the non-risky asset is \( X_t - \sum_{i=1}^n \pi^i_t = X_t - \pi_t \cdot 1 \).

Under the self-financing condition, the dynamics of the wealth process is given by

\[
dX_t = \sum_{i=1}^n \frac{\pi^i_t}{S^i_t} dS^i_t + \frac{X_t - \pi_t \cdot 1}{S^0_t} dS^0_t
\]

Let \( \tilde{X} \) be the discounted wealth process

\[
\tilde{X}_t := X_t \exp \left( -\int_0^t r(u)du \right), \quad 0 \leq t \leq T.
\]

Then, by an immediate application of Itô’s formula, we see that

\[
d\tilde{X}_t = \tilde{\pi}_t \cdot \text{diag}[\tilde{S}_t]^{-1} d\tilde{S}_t
= \tilde{\pi}_t \cdot \sigma(t) \left( \lambda^0(t) dt + dW_t \right), \quad 0 \leq t \leq T.
\]

We still have to place further technical conditions on \( \pi \), at least for the above wealth process to be well-defined as a stochastic integral. Before this, let us observe that, assuming that the risk premium function satisfies:

\[
\int_0^T |\lambda^0(t)|^2 dt < \infty,
\]

it follows from the Cameron-Martin change of measure theorem that the process

\[
W^0_t := W_t - \int_0^t \lambda^0(u) du, \quad 0 \leq t \leq T,
\]
is a Brownian motion under the equivalent probability measure
\[ \mathbb{P}^0 := Z_T^0 \cdot \mathbb{P} \quad \text{on} \quad \mathcal{F}_T, \] (7.4)
where
\[ Z_T^0 := \exp \left( -\int_0^T \lambda^0(u) \cdot dW_u - \frac{1}{2} \int_0^T |\lambda^0(u)|^2 du \right). \] (7.5)

In terms of the \( \mathbb{P}^0 \) Brownian motion \( W^0 \), the discounted wealth process induced by an initial capital \( x \) and a portfolio strategy \( \pi \) can be written in
\[ \tilde{X}_t^x,\pi = \tilde{X}_0 + \int_0^t \tilde{\pi}_u \cdot \sigma(u) dW^0_u, \quad 0 \leq t \leq T. \] (7.6)

**Definition 7.18** An admissible portfolio process \( \pi = \{\pi_t, t \in [0,T]\} \) is an \( \mathbb{F} \)-adapted process such that the stochastic integral (7.6) is well-defined, and the corresponding discounted wealth process satisfies the finite credit line condition

\[ \inf_{0 \leq t \leq T} \tilde{X}_t^x,\pi \geq -C \]

for some constant \( C > 0 \) which may depend on \( \pi \). The collection of all admissible portfolio processes will be denoted by \( \mathcal{A} \).

By the previous chapter, the existence of the stochastic integral in (7.6) is guaranteed for portfolio processes \( \pi \) satisfying
\[ \int_0^T |\sigma(t)^T \pi_t|^2 dt < \infty \quad \mathbb{P} \text{-a.s.} \]

The finite credit line condition is needed in order to exclude any arbitrage opportunity, and will be justified in the subsequent subsection.

### 7.2 Admissible portfolios and no-arbitrage

Similarly to the discrete-time framework, we introduce the following no-arbitrage condition.
Definition 7.19 We say that the financial market contains no arbitrage opportunities if for any admissible portfolio process \( \pi \in \mathcal{A} \),

\[
X^{0,\pi}_T \geq 0 \quad \mathbb{P} - a.s. \quad \text{implies} \quad X^{0,\pi}_T = 0 \quad \mathbb{P} - a.s.
\]

The purpose of this section is to show that the financial market described above contains no arbitrage opportunities. Our first observation is that \( \mathbb{P}^0 \) is a risk neutral measure, or an equivalent martingale measure, for the price process \( S \), i.e.

the process \( \{ \tilde{S}_t, \, 0 \leq t \leq T \} \) is a \( \mathbb{P}^0 \) – martingale. (7.7)

Indeed, in terms of the \( \mathbb{P}^0 \)–Brownian motion \( W^0 \), the discounted price process of the risky asset is given by

\[
\tilde{S}_t^i = S_0^i \exp \left( \int_0^t \sigma_i(s) \cdot dW_t^0 - \frac{1}{2} \int_0^t |\sigma_i(s)|^2 \, ds \right),
\]

where \( \sigma_i = (\sigma_{i1}, \ldots, \sigma_{id}) \). This is clearly an (exponential) martingale under \( \mathbb{P}^0 \).

We next observe that the discounted wealth process is a local martingale:

\[
\tilde{X}^{x,\pi} \quad \text{is a local } \mathbb{P}^0 - \text{martingale for every } (x, \pi) \in \mathbb{R} \times \mathcal{A}, \quad (7.8)
\]
as a stochastic integral with respect to the \( \mathbb{P}^0 \)–Brownian motion \( W^0 \).

Before stating the main result of this section, we recall the situation in the finite discrete-time framework of the first chapters of these notes. Assuming that the set \( \mathcal{M}(S) \) of risk neutral measures is not empty, the discounted wealth process is a local martingale under any fixed risk neutral measure \( \mathbb{Q} \). Then, whenever the final wealth is non-negative, it follows from Lemme 3.1 that the discounted wealth process is a martingale under \( \mathbb{Q} \). We then conclude immediately that the final wealth must coincide with the zero random variable by taking expected values under \( \mathbb{Q} \).

In the present context, the non-negativity of the terminal wealth does not ensure that the \( \mathbb{P}^0 \)–local martingale \( \tilde{X}^{0,\pi} \) is a martingale, recall the example
developed in Example 6.14. Observe that the latter example also shows that the finite credit line condition on admissible portfolios does not guarantee that the discounted wealth process is a $\mathbb{P}^0$-martingale.

In the continuous-time framework, one has to rule out the so-called class of doubling strategies which are naturally excluded in finite discrete-time. Doubling strategies are well-known in Casinos. Suppose that a player starts playing in the Casino without any initial funds. So he borrows some capital $x_0 > 0$ (with zero interest rate, say) and invests it on some game whose payoff is $2x_0$ with some positive probability $p$, and 0 with probability $1 - p$. At the end of this first round, the player’s wealth is $X_1 = x_0$ with probability $p$, and $X_1 = -x_0$ with probability $1 - p$. If $X_1 = -x_0$, then the player borrows again the amount $x_1 = 2x_0$, i.e. doubles his debt, and plays again. At the end of this second round, the player’s wealth, conditionally on $X_1 = -x_0$, is again either $X_2 = x_0 = x_0$ with probability $p$, and $X_2 = -2x_0$ with probability $1 - p$. And so on... If the player is allowed to play an infinite number of rounds, then he would perform this doubling strategy up to the first moment that his wealth is positive, which happens at some finite time with probability 1.

In conclusion, this strategy allows to perform some positive gain without any initial investment! the drawback of this strategy is, however, that the wealth process can be arbitrarily large negative. So it is only feasible if there is no restriction on borrowing, which is not realistic.

Since the continuous-time analogue of the discrete-time random walk is the Brownian motion, see Chapter 4.2, the above doubling strategies provide a good intuition for our continuous-time model. The following result from stochastic integration is due to Dudley [10], and is stated for completeness in order to show how things can get even worse in continuous-time.

**Theorem 7.18** Let $W$ be a standard one-dimensional Brownian motion. For $T > 0$, and $\xi \in \mathbb{L}^0(T)$, there exists a progressively measurable process $\phi$
satisfying
\[ \xi = \int_0^T \phi_t dW_t \quad \text{and} \quad \int_0^T |\phi_t|^2 dt < \infty \quad \mathbb{P} - \text{a.s.} \]

We now turn to the main technical ingredient in order to exclude arbitrage opportunities.

**Lemma 7.8** Let \( M = \{M_t, 0 \leq t \leq T\} \) be a local martingale bounded from below by some constant \( m \), i.e. \( M_t \geq m \) for all \( t \in [0, T] \) a.s. Then \( M \) is a supermartingale.

**Proof.** Let \((T_n)_n\) be a localizing sequence of stopping times for the local martingale \( M \), i.e. \( T_n \to \infty \) a.s. and \( \{M_{t \wedge T_n}, 0 \leq t \leq T\} \) is a martingale for every \( n \). Then:
\[
\mathbb{E}[M_{t \wedge T_n} | \mathcal{F}_u] = M_{u \wedge T_n}, \quad 0 \leq u \leq t \leq T
\]
for every fixed \( n \). We next send \( n \) to infinity. By the lower bound on \( M \), we can use Fatou’s lemma, and we deduce that:
\[
\mathbb{E}[M_t | \mathcal{F}_u] \leq M_u, \quad 0 \leq u \leq t \leq T,
\]
which is the required inequality. \( \diamond \)

**Exercise 7.14** Show that the conclusion of Lemma 7.8 holds true under the weaker condition that the local martingale \( M \) is bounded from below by a martingale.

Given an admissible portfolio process \( \pi \in \mathcal{A} \) and an initial capital \( x \in \mathbb{R} \), the corresponding wealth process \( \tilde{X}^{x, \pi} \) is a \( \mathbb{P}^0 \)-local martingale which is bounded from below by some constant. It follows from Lemma 7.8 that
\[
\tilde{X}^{x, \pi} \quad \text{is a} \quad \mathbb{P}^0 - \text{supermartingale for every} \quad (x, \pi) \in \mathbb{R} \times \mathcal{A}. \quad (7.9)
\]
A first consequence of this supermartingale property concerns the no-arbitrage condition in our continuous-time financial market.
Theorem 7.19 The continuous-time financial market described above contains no arbitrage opportunities.

Proof. For $\pi \in A$, the corresponding discounted wealth process $\tilde{X}^{0,\pi}$ is a $\mathbb{P}^0$-super-martingale. Then $\mathbb{E}^{\mathbb{P}^0}[\tilde{X}^{0,\pi}_T] \leq 0$. Recall that $\mathbb{P}^0$ is equivalent to $\mathbb{P}$ and $S^0$ is strictly positive. Then, this inequality shows that, whenever $X^{0,\pi}_T \geq 0 \mathbb{P}-a.s.$, or equivalently $\tilde{X}^{0,\pi}_T \geq 0 \mathbb{P}^0-a.s.$, we have $\tilde{X}^{0,\pi}_T = 0 \mathbb{P}^0-a.s.$ and therefore $X^{0,\pi}_T = 0 \mathbb{P}-a.s.$

7.3 The Merton consumption/investment problem

In this section, we study the problem of an investor who has access to the financial market with one risk-free and $N$ risky assets, described earlier in this chapter, and wants to optimize her consumption and investment policy.

Let $\theta = (\theta^0, \ldots, \theta^N)$ denote the amounts of assets held by the agent. The agent is allowed to consume continuously at rate $c_t$ from the bank account, so that her portfolio value $X_t$ satisfies

$$X_t = \sum_{i=0}^{N} \theta^i_t S^i_t = x + \sum_{i=0}^{N} \int_{0}^{t} \theta^i_s dS^i_s - \int_{0}^{t} c_s ds$$

Throughout this section, we impose the positivity constraint on the portfolio: $X_t \geq 0, t \in [0, T]$, where $T$ is the time horizon of the problem (retirement, house purchase etc.)

The agent wants to maximize the functional

$$E \left[ \int_{0}^{T} e^{-\rho t} u(c_t) dt + e^{-\rho T} U(X_T) \right]$$

over all admissible consumption/investment policies (those for which the integrals exist and the wealth remains nonnegative). This problem was first studied by R. Meron in his seminal paper [18] and is known as the Merton consumption/investment problem.
The first step is to reduce the dimension of the problem by introducing portfolio fractions

\[ \omega^i_t = \frac{\theta^i_t S^i_t}{\sum_{j=0}^{N} \theta^j_t S^j_t} \] if \( X_t \neq 0 \) and 0 otherwise, \( i = 1, \ldots, N. \)

The fractions allow to get rid of the dependence on \( S \) in the portfolio dynamics which becomes

\[ dX_t = X_t \{ \omega^T_t \sigma (\lambda dt + dW_t) + r dt \} - c_t dt. \]

Note that the positivity constraint is essential to be able to work with portfolio fractions.

### 7.3.1 The Hamilton-Jacobi-Bellman equation

We want to solve the optimization problem going backwards, from \( T \) to 0, similarly to how we solved the American option pricing problem on a binomial tree in discrete time. We mainly give heuristic explanations without proofs, but the optimality of the strategies will be shown below in particular examples.

The first step is to introduce the value function:

\[ v(t, x) = \sup_{\omega, c} E \left[ \int_t^T e^{-\rho(s-t)} u(c_s) ds + e^{-\rho(T-t)} U(X_T) | X_t = x \right]. \]

Observe that \( v(0, x) \) gives the solution to our optimization problem and \( v(T, x) = U(x) \) produces the terminal condition. We now need to link the value functions for different values of \( t \) via a kind of “infinitesimal” backward induction. The basic idea is that if we know the value function at date \( t + h \), we can use this information to compute the value function at date \( t \). The relationship between the two is called the **dynamic programming principle**:

\[ v(t, x) = \sup_{\omega, c} E \left[ \int_t^{t+h} e^{-\rho(s-t)} u(c_s) ds + e^{-\rho h} v(t + h, X_{t+h}) | X_t = x \right]. \]
Suppose that $v$ is a $C^{1,2}$ function. Then we can apply the Itô formula (Theorem 6.15) obtaining

$$
0 = \sup_{\omega, c} E \left[ \int_t^{t+h} e^{-\rho(s-t)} \left\{ u(c_s) - c_s \frac{\partial v}{\partial x} - \rho v(s, X_s) + \frac{\partial v}{\partial t} + X_s \frac{\partial v}{\partial x} (r + \omega_s^T \sigma \lambda) + \frac{X_s^2}{2} \frac{\partial^2 v}{\partial x^2} \omega_s^T \sigma \sigma^T \omega_s \right\} ds + \int_t^{t+h} e^{-\rho(s-t)} \frac{\partial v}{\partial x} X_s \omega_s^T \sigma dW_s \right].
$$

Supposing (this reasoning is completely heuristic) that the stochastic integral is a true martingale, dividing everything by $h$ and making $h$ go to zero, we obtain the Hamilton-Jacobi-Bellman equation:

$$
0 = \sup_{\omega, c} \left\{ u(c) - c \frac{\partial v}{\partial x} - \rho v(t, x) + \frac{\partial v}{\partial t} + x \frac{\partial v}{\partial x} (r + \omega^T \sigma \lambda) + \frac{x^2}{2} \frac{\partial^2 v}{\partial x^2} \omega^T \sigma \sigma^T \omega \right\},
$$

from which the optimal investment policy can be immediately computed (if $v$ is known):

$$
\omega = -\frac{\frac{\partial v}{\partial x}}{x \frac{\partial^2 v}{\partial x^2}} \sigma^{-1} \lambda.
$$

We therefore recover the mutual fund theorem in the continuous-time framework: every agent invests in the same portfolio $\sigma^{-1} \lambda$, and only the fraction $-\frac{\partial v}{x \frac{\partial^2 v}{\partial x^2}}$ depends on the agents’ preferences. Note that the optimal portfolio is defined in terms of constant proportions: this is the so-called fixed mix strategy, where the weights of the assets remain constant, but the actual amounts of each asset that the agent holds may change continuously.

In the following we compute the optimal consumption/investment strategy in particular cases, using the HJB equation to guess the optimal solution, but giving a rigorous proof of optimality.

### 7.3.2 Pure investment problem with power utility

We suppose $u \equiv 0$ and $U = \frac{x^2}{2}$, $\gamma \in (0, 1)$. Since $U$ is increasing, it is not optimal to consume, and therefore $c_t = 0$ for all $t \leq T$. Moreover, the effect
of \( \rho \) reduces to multiplying the terminal utility by a constant, which means that we can set \( \rho = 0 \) without loss of generality. With these assumptions, the HJB equation simplifies to

\[
0 = \sup_\omega \left\{ \frac{\partial v}{\partial t} + x \frac{\partial v}{\partial x} (r + \omega^T \sigma \lambda) + \frac{x^2}{2} \frac{\partial^2 v}{\partial x^2} \omega^T \sigma \sigma^T \omega \right\}.
\]

From the form of the terminal condition, we make an “educated guess” that

\[
v(t, x) = x^{\gamma} f(t)
\]

for some function \( t \). This substitution reduces the HJB equation to

\[
\frac{\partial f}{\partial t} + \gamma f(t) \sup_\omega \left\{ r + \omega^T \sigma \lambda + (\gamma - 1) \omega^T \sigma \sigma^T \omega \right\},
\]

which is solved by

\[
\omega^* = \frac{\sigma^{-1} \lambda}{1 - \gamma},
\]

\[
f(t) = e^{\varepsilon(T-t)}, \quad \varepsilon = \frac{\gamma \| \lambda \|^2}{2(1 - \gamma)} + r \gamma.
\]

The optimal investment strategy in Merton’s model with power utility therefore consists in keeping a constant proportion of one’s wealth in each asset, including the bank account.

It remains to check the optimality of the strategy \( \omega^* \), that is, to show that \( v(0, x) = E[U(X^* T)] \) and that \( v(0, x) \geq E[U(X^ T)] \) for every admissible strategy \( \omega \). Let \( \omega \) be one such strategy. Since \( v \) satisfies the HJB equation, Itô formula yields

\[
v(0, x) + \int_0^T \frac{\partial v}{\partial x} X_t^\omega \omega_t^T \sigma dW_t \geq U(X^* T),
\]

\[
v(0, x) + \int_0^t \frac{\partial v}{\partial x} X_s^\omega \omega_s^T \sigma dW_s \geq v(0, X^* t) \geq 0 \quad \forall t \leq T.
\]

The second line shows that the stochastic integral is a local martingale bounded from below and thus a supermartingale. Taking the expectation
of the first line, we then get
\[ v(0, x) \geq v(0, x) + E \left[ \int_0^T \frac{\partial v}{\partial x} X_t^T \sigma dW_t \right] \geq E[U(X_T^*)]. \]

For the optimal strategy, we have equality in the HJB equation, which implies
\[ v(0, x) + \int_0^T (X_t^*)^T \omega_t^T \sigma dW_t = U(X_T^*) \]

Moreover, using the explicit form of \( X_t^* \), it is easy to check that
\[ E \left[ \int_0^T (X_t^*)^2 \gamma dt \right] < \infty, \]
which guarantees that
\[ \int_0^T \frac{\partial v}{\partial x} (X_t^*)^T \omega_t^T \sigma dW_t \]
has zero expectation.

### 7.3.3 Investment-consumption problem in infinite horizon

In this example, we set \( T = \infty \), \( u(x) = \frac{x^{\gamma}}{\gamma} \) and \( U = 0 \). In this case, the value function does not depend on time, and the HJB equation becomes
\[ 0 = \sup_{\omega, c} \left\{ u(c) - c \frac{\partial v}{\partial x} - \rho v(x) + x \frac{\partial v}{\partial x} (r + \omega^T \sigma \lambda) + \frac{x^2}{2} \frac{\partial^2 v}{\partial x^2} \omega^T \sigma \sigma^T \omega \right\}. \]

The optimal consumption satisfies
\[ c^* = \arg \sup_c \left\{ \frac{c^\gamma}{\gamma} - c \frac{\partial v}{\partial x} \right\} \quad \Rightarrow \quad c^* = \left( \frac{\partial v}{\partial x} \right)^{\frac{1}{1-\gamma}} \]
and
\[ u(c^*) - c^* \frac{\partial v}{\partial x} = \frac{1 - \gamma}{\gamma} \left( \frac{\partial v}{\partial x} \right)^{\frac{1}{1-\gamma}}. \]
Once again, we guess the solution: $v(x) = \frac{Kx^\gamma}{\gamma}$, and substituting this into the equation, we see, as before, that the optimal investment policy is a constant proportion strategy:

$$\omega^* = \frac{\sigma^{-1}\lambda}{1 - \gamma}.$$ 

After some more computations, we find the constant $K$ and the optimal consumption policy:

$$K = (\alpha^*)^{\gamma^{-1}} \quad \text{and} \quad c^* = \alpha^* x,$$

where

$$\alpha^* = \frac{1}{1 - \gamma} \left\{ \rho - \gamma r - \frac{\gamma \|\lambda\|^2}{2(1 - \gamma)} \right\}.$$ 

The agent must therefore consume a constant fraction of his/her current wealth. The portfolio value corresponding to the optimal investment and consumption policy is easily computed:

$$X_t^* = x \exp \left\{ \left( \frac{r - \rho}{1 - \gamma} + \frac{\|\lambda\|^2}{2(1 - \gamma)} \right) t + \frac{\lambda^Tw_t}{1 - \gamma} \right\}.$$ 

To verify that the strategy we have found is indeed optimal, we first show, by the same arguments as in the previous example, that for any consumption-investment pair $(c, \omega)$,

$$v(x) \geq E \left[ \int_0^T e^{-\rho t} u(c_t) dt + e^{-\rho T} v(X_T^\omega) \right] \geq E \int_0^T e^{-\rho t} u(c_t) dt,$$

and by making $T$ go to $\infty$,

$$v(x) \geq E \int_0^\infty e^{-\rho t} u(c_t) dt.$$ 

For the optimal strategy, we get

$$v(x) = E \left[ \int_0^T e^{-\rho t} u(c_t^*) dt + e^{-\rho T} v(X_T^{\omega^*}) \right],$$

and in order to conclude, it remains to prove that

$$\lim_{T \to \infty} E[e^{-\rho T}v(X_T^{\omega^*})] = 0.$$ 

Using the explicit form of $X_T^{\omega^*}$ it is easy to check that the latter condition holds if and only if $\alpha^* > 0$. 

117
7.3.4 The martingale approach

In this section we study a different approach to the problem considered in section 7.3.2:

\[
\sup_{\omega \in \mathcal{A}} E[U(X_\omega^T)],
\]

where \( \mathcal{A} \) is the set of strategies corresponding to admissible portfolios. We denote by

\[
Z_T = \exp\left(-\lambda^T W_T - \frac{\|\lambda\|^2 T}{2}\right)
\]

the density of the martingale measure. From the positivity condition, it follows that every admissible portfolio satisfies

\[
E[e^{-rT} Z_T X_T^\omega] \leq x.
\]

Similarly to how this was done in discrete time, we decompose the portfolio optimization problem into a problem of finding the optimal contingent claim under the budget constraint, and the problem of replicating this claim with an admissible portfolio. The optimal claim problem is

\[
\sup_X E[U(X)] \quad \text{subj to} \quad E[e^{-rT} Z_T X_T] \leq x.
\]

The Lagrangian of this problem is

\[
E[U(X_T) - y(e^{-rT} Z_T X_T - x)].
\]

Introduce the Fenchel transform of \( U \):

\[
V(y) = \sup_x \{U(x) - xy\}, \quad \text{with maximizer given by} \quad x^* = (U')^{-1}(y) := I(y).
\]

Then the optimization problem is (formally) solved by

\[
X_T^* = I(y e^{-rT} Z_T),
\]

where \( y \) is found from the condition

\[
E[e^{-rT} Z_T I(y e^{-rT} Z_T)] = x.
\]
In the case of power utility,

\[
U(x) = \frac{x^\gamma}{\gamma} \quad \Rightarrow \quad I(y) = y^{\frac{1}{\gamma}} \quad V(y) = \frac{1 - \gamma}{\gamma} y^{\frac{\gamma}{1 - \gamma}}
\]

and the candidate optimal claim is given by

\[
X^*_T = xe^{rT} \frac{Z_T^{-\frac{1}{\gamma}}}{E[Z_T^{-\frac{1}{\gamma}}]} = xe^{rT} \exp \left( \frac{\lambda^T W_T}{1 - \gamma} + \frac{1 - 2\gamma}{(1 - \gamma)^2} \| \lambda \|^2 T \right),
\]

which is easily shown to be the terminal value of a portfolio corresponding to the strategy

\[
\omega = \frac{\sigma^{-1} \lambda}{1 - \gamma}.
\]

It remains to check the optimality of the claim. For any admissible strategy \( \omega \) and any \( y \),

\[
E[U(X^*_T)] \leq E[V(ye^{-rT}Z_T)] + X^*_T ye^{-rT} Z_T \leq E[V(ye^{-rT}Z_T)] + xy.
\]

On the other hand, for the optimal claim \( X^* \) it is easy to see that there exists an \( y^* \) for which

\[
E[U(X^*_T)] = E[V(y^* e^{-rT} Z_T)] + xy^*.
\]
Chapter 8

The Black-Scholes valuation theory

8.1 Pricing and hedging Vanilla options

In this section, we consider a special class of contingent claims defined by the payoff at some maturity $T > 0$:

$$G := g(S_T) \text{ for some } g : \mathbb{R}_+^d \rightarrow \mathbb{R}. $$

We shall refer to such contingent claims as Vanilla options. The following result provides an explicit perfect replication strategy for such options by continuous investment on the financial market, and provides as a by-product their no-arbitrage price.

Theorem 8.20 (i) Suppose that the function $g$ is bounded. Then, the no arbitrage price of the Vanilla option $G$ at time $0$ is given by

$$p_0(G) = v(0, S_0) \text{ where } v(t, s) := \mathbb{E}^{\mathbb{P}_0} \left[ e^{-\int_t^T r(u)du} g(S_T) \bigg| S_t = s \right].$$

(ii) Assume further that the market price $p_0$ is a continuous map from $L^1(\mathbb{P}^0)$ to $\mathbb{R}$. Then (i) holds whenever $g(S_T) \in L^1(\mathbb{P}^0)$. 

120
(iii) The above function \( v \in C^{1,2} ([0,T), \mathbb{R}^d) \). If \( g \) is bounded from below, then

\[
\hat{\pi}_{t} := \text{diag} [S_{t} \frac{\partial v}{\partial s} (t, S_{t})], \quad 0 \leq t \leq T,
\]
defines an admissible portfolio strategy in \( \hat{\pi} \in \mathcal{A} \) which perfectly replicates \( G \), i.e. \( X_{T}^{(v(0,S_{0})),\hat{\pi}} = g(S_{T}) \).

**Proof.** Part (ii) is a direct consequence of (i) which follows by considering the approximating contingent claims defined by the truncated payoff function \( g^{n} := g \wedge n \vee (-n) \), and using the continuity assumption. So we only have to prove (i) and (iii).

We introduce the super-hedging problem

\[
V_{0}(G) := \inf \{ x : X_{T}^{x,\pi} \geq G \text{ a.s. for some } \pi \in \mathcal{A} \},
\]
and we argue that the market price \( p_{0}(G) \) lies inside the no-arbitrage bounds:

\[
-V_{0}(-G) \leq p_{0}(G) \leq V_{0}(G).
\]

(8.1)

1. For an arbitrary \( x > V_{0}(G) \), we have \( \hat{X}_{T}^{x,\pi} \geq \hat{G} \text{ a.s. for some } \pi \in \mathcal{A} \). Since \( \hat{X}_{T}^{x,\pi} \) is a \( \mathbb{P} \)-supermartingale, we deduce that \( x \geq \mathbb{E}^{\mathbb{P}^{0}} [\hat{G}] \), and therefore \( V_{0}(G) \geq \mathbb{E}^{\mathbb{P}^{0}} [\hat{G}] \).

2. In order to prove the reverse inequality \( V_{0}(G) \geq \mathbb{E}^{\mathbb{P}^{0}} [\hat{G}] \), we show in Step 3 below that \( G \) can be perfectly hedged starting from the initial capital \( \mathbb{E}^{\mathbb{P}^{0}} [\hat{G}] \). Since \( -G \) satisfies the same conditions as \( G \), the conclusions of Step 1 and the first part of this step hold also for \( -G \). The required results then follow from the no-arbitrage bounds (8.1).

3. It remains to show that \( G \) can be perfectly hedged by means of an admissible portfolio, starting from the initial capital \( \mathbb{E}^{\mathbb{P}^{0}} [\hat{G}] \). Since \( \ln S_{t} \) has a Gaussian distribution, satisfying the heat equation, the function \( v \) satisfies

\[
\frac{\partial v}{\partial t} + rs \cdot \frac{\partial v}{\partial s} + \frac{1}{2} \text{Tr} \left[ \text{diag}[s] \sigma \sigma^{T} \text{diag}[s] \frac{\partial^2 v}{\partial s \partial s^{T}} \right] - rv = 0
\]
for \( t < T, s \in \mathbb{R}_+^d \). We then deduce from Itô’s formula that
\[
e^{-rT}v(T, S_T) = v(0, S_0) + \int_0^T e^{-ru} \tilde{\pi}_u \cdot \sigma dW_u = \tilde{X}_T^{v(0, S_0), \tilde{\pi}}.
\]
Since \( v(T, s) = g(s) \), this shows that \( X_T^{v(0, S_0), \tilde{\pi}} = g(S_T) \). Since \( g(S) \) is square integrable, as a consequence of the polynomial growth condition on \( g \), it follows that \( \tilde{X}_T^{v(0, S_0), \tilde{\pi}} \) is a \( \mathbb{P}^0 \)-square integrable martingale (exercise !), and therefore
\[
\tilde{X}_t^{v(0, S_0), \tilde{\pi}} = e^{-\int_0^t r(u) du} v(t, S_t)
\]
hers thus the boundness of \( g \). Hence \( \tilde{\pi} \) is an admissible portfolio strategy in \( \mathcal{A} \).

**Remark 8.12** In the above proof, we used the fact that the option price \( v(t, s) \) satisfies the linear partial differential equation
\[
\frac{\partial v}{\partial t} + rs \cdot \frac{\partial v}{\partial s} + \frac{1}{2} \text{Tr} \left[ \text{diag}[s] \sigma \sigma^T \text{diag}[s] \frac{\partial^2 v}{\partial s \partial s^T} \right] - rv = 0, \ (t, s) \in [0, T) \times \mathbb{R}_+.
\]
From Theorem 6.16, this partial differential equation provides a characterization of the option price. In the subsequent section, we shall obtain this characterization by means of a heuristic argument which avoids to use the martingale theory.

### 8.2 The PDE approach for the valuation problem

In this section, we derive a formal argument in order to obtain the valuation PDE (8.2). The following steps have been employed by Black and Scholes in their pioneering work [1].

1. Let \( p(t, S_t) \) denote the time–t market price of a contingent claim defined by the payoff \( B = g(S_T) \) for some function \( g : \mathbb{R}_+ \to \mathbb{R} \). Notice that we
are accepting without proof that $p$ is a deterministic function of time and the spot price, this has been in fact proved in the previous section.

2. The holder of the contingent claim completes his portfolio by some investment in the risky assets. At time $t$, he decides to holds $-\Delta^i$ shares of the risky asset $S^i$. Therefore, the total value of the portfolio at time $t$ is

$$P_t := p(t, S_t) - \Delta \cdot S_t, \quad 0 \leq t < T.$$  

3. Considering delta as a constant vector in the time interval $[t, t + dt)$, and assuming that the function $p$ is of class $C^{1,2}$, the variation of the portfolio value is given by:

$$dP_t = \mathcal{L}p(t, S_t)dt + \frac{\partial p}{\partial s}(t, S_t) \cdot dS_t - \Delta \cdot dS_t.$$  

where $\mathcal{L}p = \frac{\partial p}{\partial t} + \frac{1}{2} \text{Tr} [\text{diag}[s] \sigma \sigma^T \text{diag}[s] \frac{\partial^2 p}{\partial s \partial s^T}]$. In particular, by setting

$$\Delta = \frac{\partial p}{\partial s},$$

we obtain a portfolio value with finite quadratic variation

$$dP_t = \mathcal{L}p(t, S_t)dt.$$  

(8.3)

4. The portfolio $P_t$ is non-risky since the variation of its value in the time interval $[t, t + dt)$ is known in advance at time $t$. Then, by the no-arbitrage argument, we must have

$$dP_t = r(t, S_t)P_t dt = r(t, S_t)[p(t, S_t) - \Delta \cdot S_t]$$

$$= r(t, S_t) \left[ p(t, S_t) - \frac{\partial p}{\partial s} \cdot S_t \right]$$  

(8.4)

By equating (8.3) and (8.4), we see that the function $p$ satisfies the PDE

$$\frac{\partial p}{\partial t} + rs \cdot \frac{\partial p}{\partial s} + \frac{1}{2} \text{Tr} \left[ \text{diag}[s] \sigma \sigma^T \text{diag}[s] \frac{\partial^2 p}{\partial s \partial s^T} \right] - rp = 0,$$

which is exactly the PDE obtained in the previous section.
8.3 The Black and Scholes model for European call options

8.3.1 The Black-Scholes formula

In this section, we consider the one-dimensional Black-Scholes model $d = 1$ so that the price process $S$ of the single risky asset is given in terms of the $\mathbb{P}^0$-Brownian motion $W^0$:

$$S_t = S_0 \exp \left[ \left( r - \frac{\sigma^2}{2} \right) t + \sigma W^0_t \right], \quad 0 \leq t \leq T. \quad (8.5)$$

Observe that the random variable $S_t$ is log-normal for every fixed $t$. This is the key-ingredient for the next explicit result.

**Proposition 8.18** Let $G = (S_T - K)^+$ for some $K > 0$. Then the no-arbitrage price of the contingent claim $G$ is given by the so-called Black-Scholes formula:

$$p_0(G) = S_0 N \left( d_+(S_0, \tilde{K}, \sigma^2 T) \right) - \tilde{K} N \left( d_-(S_0, \tilde{K}, \sigma^2 T) \right), \quad (8.6)$$

where

$$\tilde{K} := Ke^{-rT}, \quad d_\pm(s, k, v) := \frac{\ln(s/k)}{\sqrt{v}} \pm \frac{1}{2} \sqrt{v}, \quad (8.7)$$

and the optimal hedging strategy is given by

$$\pi_t = S_t N \left( d_+(s_t, \tilde{K}, \sigma^2 (T - t)) \right), \quad 0 \leq t \leq T. \quad (8.8)$$

**Proof.** This formula can be derived by various methods. One can just calculate directly the expected value by exploiting the explicit probability density function of the random variable $S_T$. One can also guess a solution for the valuation PDE corresponding to the call option. We shall present another method which relies on the technique of change of measure and reduces considerably the computational effort. We first decompose

$$p_0(G) = \mathbb{E}^{\mathbb{P}^0} \left[ \tilde{S}_T \mathbf{1}_{\{\tilde{S}_T \geq \tilde{K} \}} \right] - \tilde{K} \mathbb{E}^{\mathbb{P}^0} \left[ \tilde{S}_T \geq \tilde{K} \right] \quad (8.9)$$
where as usual, the *tilde* notation corresponds to discounting, i.e. multiplication by $e^{-rT}$ in the present context.

1. The second term is directly computed by exploiting the knowledge of the distribution of $\tilde{S}_T$:

$$
\mathbb{P}^0 \left[ \tilde{S}_T \geq \tilde{K} \right] = \mathbb{P}^0 \left[ \frac{\ln (\tilde{S}_T / S_0) + (\sigma^2 / 2)T}{\sigma \sqrt{T}} \geq \frac{\ln (\tilde{K} / S_0) + (\sigma^2 / 2)T}{\sigma \sqrt{T}} \right]
= 1 - \mathcal{N} \left( \frac{\ln (\tilde{K} / S_0) + (\sigma^2 / 2)T}{\sigma \sqrt{T}} \right)
= \mathcal{N} \left( d_- (S_0, \tilde{K}, \sigma^2 T) \right).
$$

2. As for the first expected value, we define the new measure $\mathbb{P}^1 := Z^1_T \cdot \mathbb{P}^0$ on $\mathcal{F}_T$, where

$$
Z^1_T := \exp \left( \sigma W_T - \frac{\sigma^2}{2} T \right) = \frac{\tilde{S}_T}{S_0}.
$$

By the Girsanov theorem, the process $W^1_t := W^0_t - \sigma_t$, $0 \leq t \leq T$, defines a Brownian motion under $\mathbb{P}^1$, and the random variable

$$
\frac{\ln (\tilde{S}_T / S_0) - (\sigma^2 / 2)T}{\sigma \sqrt{T}}
$$

is distributed as $\mathcal{N}(0, 1)$ under $\mathbb{P}^1$.

We now re-write the first term in (8.9) as

$$
\mathbb{E}^{\mathbb{P}^0} \left[ \tilde{S}_T \mathbf{1}_{\{\tilde{S}_T \geq \tilde{K}\}} \right] = S_0 \mathbb{P}^1 \left[ \tilde{S}_T \geq \tilde{K} \right]
= S_0 \text{Prob} \left[ \mathcal{N}(0, 1) \geq \frac{\ln (\tilde{K} / S_0) - (\sigma^2 / 2)T}{\sigma \sqrt{T}} \right]
= S_0 \mathcal{N} \left( d_+ (S_0, \tilde{K}, \sigma^2 T) \right).
$$

3. The optimal hedging strategy is obtained by directly differentiating the price formula with respect to the underlying risky asset price, see Theorem 8.20. 

\[\Diamond\]
Exercise 8.15 (Black-Scholes model with time-dependent coefficients)
Consider the case where the interest rate is a deterministic function $r(t)$, and the risky asset price process is defined by the time dependent coefficients $b(t)$ and $\sigma(t)$. Show that the European call option price is given by the extended
Black-Scholes formula:
\[ p_0(G) = S_0 \mathcal{N} \left( d_+ (S_0, \tilde{K}, v(T)) \right) - \tilde{K} \mathcal{N} \left( d_- (S_0, \tilde{K}, v(T)) \right) \]
(8.10)
where
\[ \tilde{K} := Ke^{-\int_0^T r(t) dt}, \quad v(T) := \int_0^T \sigma^2(t) dt. \]  
(8.11)

What is the optimal hedging strategy.  

\[ \diamond \]

8.3.2 The Black’s formula

We again assume that the financial market contains one single risky asset with price process defined by the constant coefficients Black-Scholes model. Let \( \{F_t, t \geq 0\} \) be the price process of the forward contract on the risky asset with maturity \( T' > 0 \). Since the interest rates are deterministic, we have
\[ F_t = S_t e^{r(T'-t)} = F_0 e^{-\frac{1}{2} \sigma^2 t + \sigma W_t^0}, \quad 0 \leq t \leq T. \]
In particular, we observe that the process \( \{F_t, t \in [0, T']\} \) is a martingale under the risk neutral measure \( \mathbb{P}^0 \). As we shall see in next chapter, this property is specific to the case of deterministic interest rates, and the corresponding result in a stochastic interest rates framework requires to introduce the so-called forward neutral measure.

We now consider the European call option on the forward contract \( F \) with maturity \( T \in (0, T'] \) and strike price \( K > 0 \). The corresponding payoff at the maturity \( T \) is \( G := (F_T - K)^+ \). By the previous theory, its price at time zero is given by
\[ p_0(G) = \mathbb{E}^{\mathbb{P}^0} \left[ e^{-rT} (F_T - K)^+ \right]. \]
In order to compute explicitly the above expectation, we shall take advantage of the previous computations, and we observe that \( e^{rT} p_0(G) \) corresponds to the Black-Scholes formula for a zero interest rate. Hence:
\[ p_0(G) = e^{-rT} \left( F_0 \mathcal{N} (d_+(F_0, K, \sigma^2 T)) - K \mathcal{N} (d_-(F_0, K, \sigma^2 T)) \right) \]  
(8.12)
This is the so-called Black’s formula.
8.3.3 Option on a dividend paying stock

When the risky asset $S$ pays out some dividend, the previous theory requires some modifications. We shall first consider the case where the risky asset pays a lump sum of dividend at some pre-specified dates, assuming that the process $S$ is defined by the Black-Scholes dynamics between two successive dates of dividend payment. This implies a downward jump of the price process upon the payment of the dividend. We next consider the case where the risky asset pays a continuous dividend defined by some constant rate. The latter case can be viewed as a model simplification for a risky asset composed by a basket of a large number of dividend paying assets.

**Lump payment of dividends** Consider a European call option with maturity $T > 0$, and suppose that the underlying security pays out a lump of dividend at the pre-specified dates $t_1, \ldots, t_n \in (0, T)$. At each time $t_j$, $j = 1, \ldots, n$, the amount of dividend payment is

$$\delta_j S_{t_j}$$

where $\delta_1, \ldots, \delta_n \in (0, 1)$ are some given constants. In other words, the dividends are defined as known fractions of the security price at the pre-specified dividend payment dates. After the dividend payment, the security price jumps down immediately by the amount of the dividend:

$$S_{t_j} = (1 - \delta_j) S_{t_j^-}, \quad j = 1, \ldots, n.$$

Between two successives dates of dividend payment, we are reduced to the previous situation where the asset pays no dividend. Therefore, the discounted security price process must be a martingale under the risk neutral measure $\mathbb{P}^0$, i.e. in terms of the Brownian motion $W^0$, we have

$$S_t = S_{t_{j-1}} e^{(r - \frac{\sigma^2}{2})(t-t_{j-1}) + \sigma (W_t - W_{t_{j-1}})}, \quad t \in [t_{j-1}, t_j),$$

for $j = 1, \ldots, n$ with $t_0 := 0$. Hence

$$S_T = \hat{S}_0 e^{(r - \frac{\sigma^2}{2})T + \sigma W_T} \quad \text{where} \quad \hat{S}_0 := S_0 \prod_{j=1}^{n}(1 - \delta_j),$$

128
and the no-arbitrage European call option price is given by
\[
\mathbb{E}^{P_0}\left[e^{rT}(S_T - K)^+\right] = \tilde{S}_0 \mathcal{N}\left(d_+(\tilde{S}_0, \tilde{K}, \sigma^2T)\right) - \tilde{K} \mathcal{N}\left(d_-(\tilde{S}_0, \tilde{K}, \sigma^2T)\right),
\]
with \(\tilde{K} = Ke^{-rT}\), i.e. the Black-Scholes formula with modified spot price from \(S_0\) to \(\tilde{S}_0\).

**Continuous dividend payment** We now suppose that the underlying security pays a continuous stream of dividend \(\{\delta S_t, \ t \geq 0\}\) for some given constant rate \(\delta > 0\). This requires to adapt the no-arbitrage condition so as to account for the dividend payment. From the financial viewpoint, the holder of the option can immediately re-invest the dividend paid in cash into the asset at any time \(t \geq 0\). By doing so, the position of the security holder at time \(t\) is
\[
S_t(\delta) := S_t e^{\delta t}, \quad t \geq 0.
\]
In other words, we can reduce the problem the non-dividend paying security case by increasing the value of the security. By the no-arbitrage theory, the discounted process \(\{r^{-rt}S_t(\delta), \ t \geq 0\}\) must be a martingale under the risk neutral measure \(\mathbb{P}^0\):
\[
S_t(\delta) = S_0 e^{\left((r-\delta-\frac{\sigma^2}{2})t + \sigma W^0_t\right)} = S_0 e^{\left(r-\frac{\sigma^2}{2}\right)t + \sigma W^0_t}, \quad t \geq 0,
\]
where \(W^0\) is a Brownian motion under \(\mathbb{P}^0\). By a direct application of Itô’s formula, this provides the expression of the security price process in terms of the Brownian motion \(W^0\):
\[
S_t = S_0 e^{\left((r-\delta-\frac{\sigma^2}{2})t + \sigma W^0_t\right)}, \quad t \geq 0. \tag{8.13}
\]
We are now in a position to provide the call option price in closed form:
\[
\mathbb{E}^{P_0}\left[e^{-rT}(S_T - K)^+\right] = e^{-\delta T} \mathbb{E}^{P_0}\left[e^{-(r-\delta)T}(S_T - K)^+\right]
= e^{-\delta T}\left[S_0 \mathcal{N}\left(d_+(S_0, \tilde{K}(\delta), \sigma^2T)\right) - \tilde{K}(\delta) \mathcal{N}\left(d_-(S_0, \tilde{K}(\delta), \sigma^2T)\right)\right] \tag{8.14}
\]
where

\[ \tilde{K}^{(\delta)} := Ke^{-(r-\delta)T}. \]

### 8.3.4 The practice of the Black-Scholes model

The Black-Scholes model is used all over the industry of derivative securities. Its practical implementation requires the determination of the coefficients involved in the Black-Scholes formula. As we already observed the drift parameter \( \mu \) is not needed for the pricing and hedging purposes. This surprising feature is easily understood by the fact that the perfect replication procedure involves the underlying probability measure only through its zero-measure sets, and therefore the problem is not changed by passage to any equivalent probability measure; by the Girsanov theorem this means that the problem is not modified by changing the drift \( \mu \) in the dynamics of the risky asset.

Since the interest rate is observed, only the volatility parameter \( \sigma \) needs to be determined in order to implement the Black-Scholes formula. After discussing this important issue, we will focus on the different control variables which are carefully scrutinized by derivatives traders in order to account for the departure of real life financial markets from the simple Black-Scholes model.

**Volatility: statistical estimation versus calibration**

1. According to the Black-Scholes model, given the observation of the risky asset prices at times \( t_i := ih, i = 1, \ldots, n \) for some time step \( h > 0 \), the returns

\[ R_{t_i} := \ln \left( \frac{S_{t_i}}{S_{t_{i-1}}} \right) \]

are iid distributed as \( \mathcal{N} \left( \left( \mu - \frac{\sigma^2}{2} \right) h, \sigma^2 h \right) \).

Then the sample variance

\[ \hat{\sigma}^2_n := \frac{1}{n} \sum_{i=1}^{n} (R_{t_i} - \bar{R}_n)^2, \]

where \( \bar{R}_n := \frac{1}{n} \sum_{i=1}^{n} R_{t_i} \),
is the maximum likelihood estimator for the parameter $\sigma^2$. The estimator $\hat{\sigma}_n$ is called the \textit{historical volatility} parameter.

The natural way to implement the Black-Scholes model is to plug the historical volatility into the Black-Scholes formula

$\text{BS} \left( S_t, \sigma, K, T \right) := S_t N \left( d_+ \left( S_t, \tilde{K}, \sigma^2 T \right) \right) - \tilde{K} N \left( d_- \left( S_t, \tilde{K}, \sigma^2 T \right) \right)$

(8.15)

to compute an estimate of the option price, and into the optimal hedge ratio

$\Delta \left( S_t, \sigma, K, T \right) := N \left( d_+ \left( S_t, \tilde{K}, \sigma^2 T \right) \right)$

(8.16)
in order to implement the optimal hedging strategy.

Unfortunately, the options prices estimates provided by this method performs very poorly in terms of fitting the observed data on options prices. Also, the use of the historical volatility for the hedging purpose leads to a very poor hedging strategy, as it can be verified by a back-testing procedure on observed data.

2. This anomaly is of course due to the simplicity of the Black-Scholes model which assumes that the log-returns are gaussian independent random variables. The empirical analysis of financial data reveals that securities prices exhibit fat tails which are by far under-estimated by the gaussian distribution. This is the so-called leptokurtic effect. It is also documented that financial data exhibits an important skewness, i.e. asymmetry of the distribution, which is not allowed by the gaussian distribution.

Many alternative statistical models have suggested in the literature in order to account for the empirical evidence (see e.g. the extensive literature on ARCH models). But none of them is used by the practitioners on (liquid) options markets. The simple and by far imperfect Black-Scholes models is still used allover the financial industry. It is however the statistical estimation procedure that practitioners have gave up very early...

3. On liquid options markets, prices are given to the practitioners and are determined by the confrontation of demand and supply on the market. Therefore, their main concern is to implement the corresponding hedging strategy.
To do this, they use the so-called calibration technique, which in the present context reduce to the calculation of the implied volatility parameter.

It is very easily checked the Black-Scholes formula (8.15) is a on-to-one function of the volatility parameter, see (8.21) below. Then, given the observation of the call option price $C^*_t(K, T)$ on the financial market, there exists a unique parameter $\sigma_{imp}^t$ which equates the observed option price to the corresponding Black-Scholes formula:

$$\text{BS} \left( S_t, \sigma_{imp}^t(K, T), K, T \right) = C^*_t(K, T),$$

(8.17)

provided that $C^*_t$ satisfies the no-arbitrage bounds of Subsection 1.4. This defines, for each time $t \geq 0$, a map $(K, T) \mapsto \sigma_{imp}^t(K, T)$ called the implied volatility surface. For their hedging purpose, the option trader then computes the hedge ratio

$$\Delta_{t}^{imp}(T, K) := \Delta \left( S_t, \sigma_{imp}^t(K, T), K, T \right).$$

If the constant volatility condition were satisfied on the financial data, then the implied volatility surface would be expected to be flat. But this is not the case on real life financial markets. For instance, for a fixed maturity $T$, it is usually observed that that the implied volatility is U-shaped as a function of the strike price. Because of this empirical observation, this curve is called the volatility smile. It is also frequently argued that the smile is not symmetric but skewed in the direction of large strikes.

From the conceptual point of view, this practice of options traders is in contradiction with the basics of the Black-Scholes model: while the Black-Scholes formula is established under the condition that the volatility parameter is constant, the practical use via the implied volatility allows for a stochastic variation of the volatility. In fact, by doing this, the practitioners are determining a wrong volatility parameter out of a wrong formula!

Despite all the criticism against this practice, it is the standard on the derivatives markets, and it does perform by far better than the statistical method. It has been widely extended to more complex derivatives markets as the fixed income derivatives, defaultable securities and related derivatives...
Risk control variables: the Greeks

With the above definition of the implied volatility, all the parameters needed for the implementation of the Black-Scholes model are available. For the purpose of controlling the risk of their position, the practitioners of the options markets various sensitivities, commonly called Greeks, of the Black-Scholes formula to the different variables and parameters of the model. The following picture shows a typical software of an option trader, and the objective of the following discussion is to understand its content.

![Figure 8.2: A typical option trader software](image)

1. **Delta**: This control variable is the most important one as it represents the number of shares to be held at each time in order to perform a perfect (dynamic) hedge of the option. The expression of the Delta is given in (8.16). An interesting observation for the calculation of this control variables and the subsequent ones is that

\[ sN'(d_+(s, k, v)) = kN'(d_-(s, k, v)) , \]
where $N'(x) = (2\pi)^{-1/2}e^{-x^2/2}$.

2. **Gamma**: is defined by

$$\Gamma(S_t, \sigma, K, T) := \frac{\partial^2 BS}{\partial s^2} (S_t, \sigma, K, T)$$

$$= \frac{1}{S_t \sigma \sqrt{T - t}} N' \left( d_+ \left( S_t, \hat{K}, \sigma^2 T \right) \right). \quad (8.18)$$

The interpretation of this risk control coefficient is the following. While the simple Black-Scholes model assumes that the underlying asset price process is continuous, practitioners believe that large movements of the prices, or jumps, are possible. A stress scenario consists in a sudden jump of the underlying asset price. Then the Gamma coefficient represents the change in the hedging strategy induced by such a stress scenario. In other words, if the underlying asset jumps immediately from $S_t$ to $S_t + \xi$, then the option hedger must immediately modify his position in the risky asset by buying $\Gamma_t \xi$ shares (or selling if $\Gamma_t \xi < 0$).

Given this interpretation, a position with a large Gamma is very risky, as it would require a large adjustment in case of a stress scenario.

3. **Rho**: is defined by

$$\rho(S_t, \sigma, K, T) := \frac{\partial BS}{\partial r} (S_t, \sigma, K, T)$$

$$= \hat{K}(T - t) N \left( d_- \left( S_t, \hat{K}, \sigma^2 T \right) \right), \quad (8.19)$$

and represents the sensitivity of the Black-Scholes formula to a change of the instantaneous interest rate.

4. **Theta**: is defined by

$$\theta(S_t, \sigma, K, T) := \frac{\partial BS}{\partial T} (S_t, \sigma, K, T)$$

$$= \frac{1}{2} S_t \sigma \sqrt{T - t} N' \left( d_- \left( S_t, \hat{K}, \sigma^2 T \right) \right), \quad (8.20)$$

134
is also called the time value of the call option. This coefficient isolates the depreciation of the option when time goes on due to the maturity shortening.

5. **Vega**: is one of the most important Greeks (although it is not a Greek letter!), and is defined by

\[ V(S_t, \sigma, K, T) := \frac{\partial BS}{\partial \sigma} (S_t, \sigma, K, T) = S_t \sqrt{T-t} N'(d_1 (S_t, K, \sigma^2 T)) \]  

(8.21)

This control variable provides the exposition of the call option price to the volatility risk. Practitioners are of course aware of the stochastic nature of the volatility process (recall the smile surface above), and are therefore seeking a position with the smallest possible Vega in absolute value.

![Figure 8.3: Representation of the Greeks](image)

Figure 8.3: Representation of the Greeks
Bibliography


[22] Revuz D. and Yor M. (1991), Continuous Martingales and Brownian Motion, Springer Verlag.


