Rose-Anne Dana
Monique Jeanblanc

Financial Markets in Continuous Time

$$dY_t = \frac{\partial f}{\partial t}\ dt + \frac{\partial f}{\partial x}\ dX(t) + \frac{1}{2}\ \frac{\partial^2 f}{\partial x^2}\ \sigma^2(t)dt$$

Springer Finance

Springer Finance

Springer Finance is a programme of books aimed at students, academics and practitioners working on increasingly technical approaches to the analysis of financial markets. It aims to cover a variety of topics, not only mathematical finance but foreign exchanges, term structure, risk management, portfolio theory, equity derivatives, and financial economics.

Ammann M., Credit Risk Valuation: Methods, Models, and Application (2001)

Back K., A Course in Derivative Securities: Introduction to Theory and Computation (2005)

Barucci E., Financial Markets Theory. Equilibrium, Efficiency and Information (2003)

Bielecki T.R. and Rutkowski M., Credit Risk: Modeling, Valuation and Hedging (2002)

Bingham N.H. and Kiesel R., Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives (1998, 2nd ed. 2004)

Brigo D. and Mercurio F., Interest Rate Models: Theory and Practice (2001, 2nd ed. 2006)

Buff R., Uncertain Volatility Models-Theory and Application (2002)

Carmona R.A. and Tehranchi M.R., Interest Rate Models: an Infinite Dimensional Stochastic Analysis Perspective (2006)

Dana R-A. and Jeanblanc M., Financial Markets in Continuous Time (2003)

Deboeck G. and Kohonen T. (Editors), Visual Explorations in Finance with Self-Organizing Maps (1998)

Delbaen F. and Schachermayer W., The Mathematics of Arbitrage (2005)

Elliott R.J. and Kopp P.E., Mathematics of Financial Markets (1999, 2nd ed. 2005)

Fengler M.R., Semiparametric Modeling of Implied Volatility (2005)

Geman H., Madan D., Pliska S.R. and Vorst T. (Editors), Mathematical Finance–Bachelier Congress 2000 (2001)

Gundlach M., Lehrbass F. (Editors), CreditRisk$^+$ in the Banking Industry (2004)

Jondeau E., Financial Modeling Under Non-Gaussian Distributions (2007)

Kellerhals B.P., Asset Pricing (2004)

Külpmann M., Irrational Exuberance Reconsidered (2004)

Kwok Y.-K., Mathematical Models of Financial Derivatives (1998)

Malliavin P. and Thalmaier A., Stochastic Calculus of Variations in Mathematical Finance (2005)

Meucci A., Risk and Asset Allocation (2005)

Pelsser A., Efficient Methods for Valuing Interest Rate Derivatives (2000)

Prigent J.-L., Weak Convergence of Financial Markets (2003)

Schmid B., Credit Risk Pricing Models (2004)

Shreve S.E., Stochastic Calculus for Finance I (2004)

Shreve S.E., Stochastic Calculus for Finance II (2004)

Yor M., Exponential Functionals of Brownian Motion and Related Processes (2001)

Zagst R., Interest-Rate Management (2002)

Zhu Y.-L., Wu X., Chern I.-L., Derivative Securities and Difference Methods (2004)

Ziegler A., Incomplete Information and Heterogeneous Beliefs in Continuous-time Finance (2003)

Ziegler A., A Game Theory Analysis of Options (2004)

Rose-Anne Dana · Monique Jeanblanc

Financial Markets in Continuous Time

Translated by Anna Kennedy

Springer

Rose-Anne Dana
Université Paris IX (Dauphine)
CEREMADE
Place de Lattre de Tassigny
75775 Paris Cedex 16, France
E-mail: dana@ceremade.dauphine.fr

Monique Jeanblanc
Université d'Evry
Département de Mathématiques
Rue du Père Jarlan
91025 Evry, France
E-mail: monique.jeanblanc@univ-evry.fr

Translator
Anna Kennedy
E-mail: anna-k.kennedy@db.com

Preface

In modern financial practice, asset prices are modelled by means of stochastic processes. Continuous-time stochastic calculus thus plays a central role in financial modelling. The approach has its roots in the foundational work of Black, Scholes and Merton. Asset prices are further assumed to be rationalizable, that is, determined by the equality of supply and demand in some market. This approach has its roots in the work of Arrow, Debreu and McKenzie on general equilibrium.

This book is aimed at graduate students in mathematics or finance. Its objective is to develop in continuous time the valuation of asset prices and the theory of the equilibrium of financial markets in the complete market case (the theory of optimal portfolio and consumption choice being considered as part of equilibrium theory).

Firstly, various models with a finite number of states and dates are reviewed, in order to make the book accessible to masters students and to provide the economic foundations of the subject.

Four chapters are then concerned with the valuation of asset prices: one chapter is devoted to the Black–Scholes formula and its extensions, another to the yield curve and the valuation of interest rate products, another to the problems linked to market incompletion, and a final chapter covers exotic options.

Three chapters deal with "equilibrium theory". One chapter studies the problem of the optimal choice of portfolio and consumption for a representative agent in the complete market case. Another brings together a number of results from the theory of general equilibrium and the theory of equilibrium in financial markets, in a discrete framework. A third chapter deals with the

Radner equilibrium in continuous time in the complete market case, and its financial applications.

Appendices provide a basic presentation of Brownian motion and of numerical solutions to partial differential equations.

<div align="right">

ROSE–ANNE DANA
MONIQUE JEANBLANC

</div>

Paris,
October 2002

Contents

1

The Discrete Case

In this first chapter, we bring together results concerning both the valuation of financial assets and equilibrium models, in a discrete framework: there are two dates, and the asset prices only take a finite number of values. We have chosen to introduce in the context of very simple models, concepts that will be developed further on in the book, in the hope of easing the reader's task.

1.1 A Model with Two Dates and Two States of the World

Here we study a financial market with two dates, time 0 and time 1, in the very simple case of two possible *states of the world* at time 1. Obviously, this situation is not very realistic. It is a textbook case, which will allow us to draw out concepts (such as hedging portfolios, arbitrage and the risk-neutral measure), which will be useful for dealing with more sophisticated models, describing more realistic situations.

1.1.1 The Model

The financial market that we are studying is made up of one stock, and one riskless investment (such as a savings account).

At time 0 (today), the stock is worth S euros. At time 1 (tomorrow, or in six months' time), the stock will be worth either S_u euros or S_d euros with $S_d < S_u$, depending on whether its price goes up or down. The outcome is not known at time 0. We usually say that the stock is worth S_u or S_d depending on the "state of the world".

The riskless investment has a rate of return equal to r $(r > 0)$: one euro invested today will yield $1 + r$ euros at time 1 (whatever the state of the world). This is why the investment is called riskless.

We now consider a *call option* (an option to buy). A call option is a financial instrument: the buyer of the option pays the seller an amount q (the

premium) at time 0, in return for the right, but not the obligation, to buy the stock at time 1, and at a price K (the *exercise price* or *strike*), which is set when the contract is signed at time 0. At the time when the buyer decides whether or not to buy the stock, he knows its price, which here is either S_u or S_d. If the price of the stock at time 1 is greater than K, the option holder buys the stock at the agreed price K and immediately sells it on, so making a gain; otherwise, he does not buy it.

A *put option* (an option to sell) gives the right to its buyer to sell a stock at a price K, which is agreed upon when the contract is signed.

The profit linked to a call is unlimited, and the losses are limited to q. For a put, the profit is limited, and the losses are unlimited.

The *valuation* of an option consists in determining the price q of the option under normal market conditions.

1.1.2 Hedging Portfolio, Value of the Option

Call Options

First, we consider the case where $S_d \leq K \leq S_u$. The other two cases are not as interesting: if $K < S_d$, the option holder will gain at least $S_d - K$ whatever the state of the world, and the seller will always make a loss (and the opposite is true when $S_u < K$).

Suppose then that $S_d \leq K \leq S_u$. We will see how to build a "portfolio" with the same payoff as the option at time 1. A *portfolio* is made up of a pair (α, β), where α is the amount, in euros, invested in the riskless asset, and β is the number of stocks the investor holds (α and β can be of any sign: one can sell stocks one does not hold[1], and borrow money). If (α, β) is the portfolio held at time 0, its value in euros is $\alpha + \beta S$. At time 1, this same portfolio is worth:

$$\alpha(1+r) + \beta S_u \quad \text{if we are in the first state of the world, the high state}$$
$$\text{(the stock price has risen)},$$

$$\alpha(1+r) + \beta S_d \quad \text{if we are in the second state of the world, the low state.}$$

We say that a portfolio *replicates* the option if it has the same payoff at time 1 as the option, and this whatever the state of the world. In other words, the two following equalities must hold:

$$\alpha(1+r) + \beta S_u = S_u - K$$
$$\alpha(1+r) + \beta S_d = 0 \ .$$

---

[1] A short sale: we can short the stock, or have a short position in the stock. Having a long position means holding the stock.

By solving the linear system above, we can easily obtain a pair (α^*, β^*):

$$\alpha^* = -\frac{S_d(S_u - K)}{(S_u - S_d)(1 + r)}; \qquad \beta^* = \frac{S_u - K}{S_u - S_d} \ .$$

The price of the option is the value at time 0 of the portfolio (α^*, β^*), that is

$$q = \alpha^* + \beta^* S = \frac{S_u - K}{S_u - S_d} \left(S - \frac{S_d}{1 + r}\right) \ . \tag{1.1}$$

This is a "fair price": with the amount q received, the option seller can buy a portfolio (α^*, β^*), which generates the gain $S_u - K$ if prices rise, and which will then cover (or hedge) his losses (we call this a *hedging portfolio*[2]). As to the option buyer, he is not prepared to pay more than q, because otherwise he could use the money to build a portfolio which would yield more than the option, for example using the same β^* and an α that is larger than α^*.

To obtain the option pricing formula without assuming $S_d \le K \le S_u$, we use the same method. We look for a pair (α^*, β^*) such that

$$\alpha^*(1 + r) + \beta^* S_u = \max(0, S_u - K) := C_u \ ,$$
$$\alpha^*(1 + r) + \beta^* S_d = \max(0, S_d - K) := C_d \ .$$

We find $\beta^* = \dfrac{C_u - C_d}{S_u - S_d}$. Notice that $\beta^* \ge 0$. In other words, the hedging portfolio of a call is a long position in the stock.

Moreover,

$$q := \alpha^* + \beta^* S = \frac{1}{1 + r} \left(\pi C_u + (1 - \pi)C_d\right) , \tag{1.2}$$

where

$$\pi := \frac{1}{S_u - S_d} \left((1 + r)S - S_d\right) \ . \tag{1.3}$$

Put Options

Similarly, we can show that the price P of a put option satisfies

$$P := \frac{1}{1 + r} \left(\pi P_u + (1 - \pi)P_d\right) ,$$

where $P_u = \max(0, K - S_u)$; $P_d = \max(0, K - S_d)$.

Of course, call options (options to buy) and put options (options to sell) can themselves be either bought or sold.

The valuation principle employed here is very general, and can be applied to other contingent claims. The cost of replicating a cash flow of H_u in the high state and of H_d in the low state is $\frac{1}{1+r}(\pi H_u + (1 - \pi)H_d)$.

---

[2] The hedging portfolio covers the losses whatever the state of the world.

1.1.3 The Risk-Neutral Measure, Put–Call Parity

The Risk-Neutral Probability Measure

Let us comment on the formulae in (1.2) and (1.3).

If $S_d < (1+r)S < S_u$, then $\pi \in]0,1[$. We can interpret (1.2) in terms of "neutrality with respect to risk". Equation (1.3) can be written

$$(1+r)S \;=\; \pi S_u + (1-\pi)S_d \;. \qquad (1.4)$$

The left-hand side of (1.4) is the gain obtained by putting S euros into a riskless investment, the right-hand side is the expected gain attained by buying a stock at a price of S euros, if the probability of the high state of world occurring is π, and if the low state of the world has probability $(1-\pi)$. Equality (1.4) translates the fact we are in a model that is *"neutral with respect to risk"*: the investor would be indifferent to the choice between the two possibilities for investment (the riskless one and the risky one) as his (expected) gain remains the same. It is "as if" there were a probability π attached to the states of the world, and under which the investor were neutral with respect to risk.

Proposition 1.1.1. *The price of a contingent claim (for example an option) is the discounted value of the expected gain with respect to the "risk-neutral" probability measure.*

Proof. For a call option, the realized gain is equal to C_u or to C_d, depending on the state of the world. As the present value of 1 euro at time 1 is $\dfrac{1}{1+r}$ euros at time 0, so the present values of the realized gains are $\dfrac{1}{1+r}C_u$ and $\dfrac{1}{1+r}C_d$. The fair price of the option being given by (1.2), the result follows.   □

There is another interpretation of this result: let S_1 be the price of the asset at time 1, and let P be the risk-neutral probability measure defined by $P(S_1 = S_u) = \pi$, $P(S_1 = S_d) = 1 - \pi$. The price of a call option is the expectation, under this probability measure, of $(S_1 - K)^+/(1+r)$. Similarly, we can show that the price of a put is the expectation under P of $(K - S_1)^+/(1+r)$.

Put–Call Parity

It is obvious that we have $(S_1 - K)^+ - (K - S_1)^+ = S_1 - K$. Hence, taking present values and expectations under the risk-neutral measure, and noticing also that the expectation of $S_1/(1+r)$ is equal to S (property (1.4)), we obtain

$$C = P + S - K/(1+r) \qquad (1.5)$$

where C is the price of the call and P is that of the put. This formula, which we will later generalize, is known as the "put–call parity".

Remark 1.1.2. It would be tempting to model the situation by introducing the probability of the event "the price goes up". However, the proof above shows that this probability does not come into the valuation formulae.

1.1.4 No Arbitrage Opportunities

An arbitrage opportunity occurs if, with an initial capital that is strictly negative, an agent can obtain a positive level of wealth at time 1, or if, with an amount capital that is initially zero, an agent can obtain a level of wealth that is positive and not identically zero. We generally make the assumption that no such opportunities exist.

Let us first show that there are no arbitrage opportunities (NAO) if and only if $S_d < (1+r)S < S_u$.

If $S_d < (1+r)S < S_u$, there exists $\pi \in]0,1[$ such that $(1+r)S = \pi S_u + (1-\pi)S_d$. Suppose that (α, β) satisfies

$$\alpha(1+r) + \beta S_u \geq 0, \qquad \alpha(1+r) + \beta S_d \geq 0$$

with at least one strict inequality. Then, multiplying the first inequality by π and the second by $1 - \pi$, and by summing the two, we obtain $\alpha + \beta S > 0$. Similarly, if we have simply

$$\alpha(1+r) + \beta S_u \geq 0, \qquad \alpha(1+r) + \beta S_d \geq 0$$

then we deduce that $\alpha + \beta S \geq 0$. In neither case do we have an arbitrage opportunity.

Conversely, if $(1+r)S \leq S_d$, then the agent can, at time 0, borrow S at a rate of r, and buy the stock at price S. At time 1, he sells the stock for S_u or S_d, and repays his loan with $(1+r)S$. So he has made a gain of at least $S_d - (1+r)S \geq 0$. It is easy to apply an analogous reasoning to the case $S_u < (1+r)S$.

We can justify the option valuation formula using the assumption of no arbitrage opportunities. Let us assume that $S_d \leq K \leq S_u$. If the price of the option is $\overline{q} > q$, where q is defined as in (1.2), then there is an arbitrage opportunity:

- at time 0, we sell the option (even if we do not actually own it) at price $\overline{q}$. With q, we can build a hedging portfolio (α^*, β^*) as described previously, and we invest the remaining money $\overline{q} - q$ at a rate of r. We have:

$$\overline{q} = \alpha^* + (\overline{q} - q) + \beta^* S .$$

  The initial investment is zero.

- at time 1:
  - if the price of the stock is S_u: the option is exercised by the buyer. We buy the stock at price S_u and hand it over to the option buyer as agreed, at a price of K; the portfolio $(\alpha^* + (\bar{q} - q), \beta^*)$ is worth $S_u - K + (\bar{q} - q)(1 + r)$, and our final wealth is $K - S_u + [S_u - K + (\bar{q} - q)(1 + r)] = (\bar{q} - q)(1 + r)$, and is strictly positive,
  - if the price of the stock is S_d: the option buyer does not exercise his right, and we are left with the portfolio, which is worth

$$(\bar{q} - q)(1 + r) > 0 .$$

Hence, we have strictly positive wealth in each state of the world with an initial funding of zero, that is, an arbitrage opportunity.

We can reason analogously in the case $\bar{q} < q$.

We will come back to the concept of no arbitrage opportunities repeatedly throughout this book.

Exercise 1.1.3. Show by reasoning in terms of no arbitrage opportunities that:

- the put–call parity formula holds,
- the price of a call is a decreasing function of the strike price,
- the price of a call is a convex function of the strike price.

We can turn to Cox–Rubinstein [72] for further consequences of no arbitrage opportunities.

1.1.5 The Risk Attached to an Option

In this section, we assume that investors believe that the stock will rise with probability p. The calculations here are carried out under this probability measure.

Risk Linked to the Underlying

The rate of return on the stock is by definition $R = \dfrac{S_1 - S}{S}$. Its expectation is

$$m_S = \frac{pS_u + (1 - p)S_d}{S} - 1 ,$$

where p is the probability of being in state of the world u.

The risk of the stock is usually measured by the variance of the rate of return of its price:

$$v_S^2 = p\left(\frac{S_u - S}{S} - m_S\right)^2 + (1 - p)\left(\frac{S_d - S}{S} - m_S\right)^2 ,$$

i.e.,

$$v_S = \frac{S_u - S_d}{S} (p(1-p))^{1/2} .$$

We say that v_S is the *volatility* of the asset.

Risk Linked to the Option

Let C be a call on the stock. The delta (Δ) of the option is the number of shares of the asset that are needed to replicate the option (it is the β of the hedging portfolio given in (1.2)), i.e., $\Delta = \dfrac{C_u - C_d}{S_u - S_d}$. This represents the sensitivity of C to the price S of the underlying asset.

The *elasticity* Ω of the option is equal to $\dfrac{C_u - C_d}{C} \Big/ \dfrac{S_u - S_d}{S}$, i.e., $\Omega = \dfrac{S}{C}\Delta$ where C is the price of the option. We denote by m_C the expectation of the rate of return on the option. The risk of the option is measured by the variance of the rate of return on the option:

$$m_C = \frac{pC_u + (1-p)C_d}{C} - 1$$

$$v_C = \{p(1-p)\}^{1/2} \frac{C_u - C_d}{C} .$$

We have that $v_C = \Omega v_S$: the risk of the call is equal to the product of the elasticity of the option by the volatility of the underlying asset. The greater the volatility of the underlying asset, the greater is the risk attached to the call.

Proposition 1.1.4. *The volatility of an option is greater than the volatility of the underlying asset:*

$$v_C \ge v_S .$$

The excess rate of return of the call is greater than the excess rate return of the asset:

$$m_C - r \ge m_S - r .$$

Notice that this last property makes it worthwhile to purchase a call.

Proof. First, we show that $\Omega \ge 1$.

We have seen how $C = \dfrac{\pi C_u + (1-\pi)C_d}{1+r}$ where $\pi = \dfrac{(1+r)S - S_d}{S_u - S_d}$. Thus

$$(1+r)(S(C_u - C_d) - C(S_u - S_d)) + (S_u C_d - S_d C_u) = 0 .$$

Using the relation $C_u = (S_u - K)^+$ and the equivalent formula for C_d, we check that $S_u C_d - S_d C_u \le 0$, and hence that $\Omega \ge 1$.

We would like to establish a relationship between m_C and m_S. To do this, we use the hedging portfolio (α, β), which satisfies

$$\begin{cases} S_u\beta + (1+r)\alpha = C_u \\ S_d\beta + (1+r)\alpha = C_d , \end{cases}$$

as well as the equality $C = \alpha + S\beta$. We then obtain (using $\beta = \Delta$)

$$S_u\Delta - C_u = (1+r)(S\Delta - C)$$

$$S_d\Delta - C_d = (1+r)(S\Delta - C) ,$$

and hence

$$p(S_u\Delta - C_u) + (1-p)(S_d\Delta - C_d) = (1+r)(S\Delta - C) .$$

Rearranging terms,

$$m_S S\Delta - m_C C = r(S\Delta - C)$$

where

$$m_C - r = \Omega(m_S - r) .$$

The excess rate of return on the call is equal to Ω, the elasticity of the option, multiplied by the excess rate of return on the asset (with $\Omega \geq 1$). $\square$

In Chap. 3, we will study these concepts in continuous time.

1.1.6 Incomplete Markets

A Finite Number of States of the World

When the asset takes the value s_j at time 1 in state of the world j with $j = 1, \ldots, k$, for $k > 2$, it is no longer possible to replicate the option, as we obtain k equations $(k > 2)$ with 2 unknowns. We consider contingent claims that are of the form $H = (h_1, h_2, \ldots, h_k)$, where h_j corresponds to the payoff in state of the world j. This contingent claim is replicable if there exists a pair (α, θ) such that $\alpha(1+r) + \theta S_1 = H$, that is such that

$$\alpha(1+r) + \theta s_j = h_j ; \forall j .$$

In this case, the price of the contingent claim H is the initial value $h = \alpha + \theta S$ of the replicating portfolio.

The set $\mathcal{P}$ of risk-neutral probability measures is by definition the set of probability measures Q that assign strictly positive probability to each state of the world, and satisfy

$$E_Q(S_1) = S(1+r) .$$

The set of risk-neutral probabilities $(q_1, q_2, \ldots, q_k)$ is determined by

$$q_j > 0 \quad \text{for } j = 1, 2, \ldots, k$$

$$\sum_{j=1}^{k} q_j = 1$$

$$\sum_{j=1}^{k} q_j s_j = (1+r)S .$$

The price range associated with the contingent claim H is defined by

$$\Big] \inf_{Q \in \mathcal{Q}} E_Q(\tilde{H}), \sup_{Q \in \mathcal{Q}} E_Q(\tilde{H}) \Big[,$$

where $\tilde{H}$ is the discounted value of H, i.e., $\tilde{H} = H/(1+r)$ in our model. We will come back to the price range later. In the meantime, we note that if the market is incomplete, and if H is replicable, then the value of $E_Q(H/(1+r))$ does not depend on the choice of risk-neutral measure Q. Indeed, if there exists (α, θ) such that

$$\alpha(1+r) + \theta s_j = h_j, \ \forall j$$

then for any choice of risk-neutral probability measure $(q_j, 1 \leq j \leq k)$, we have

$$E_Q(H) = \sum_{j=1}^{k} q_j h_j = \sum_{j=1}^{k} q_j (\alpha(1+r) + \theta s_j) = \alpha(1+r) + \theta(1+r)S .$$

A Continuum of States of the World

Let $(\Omega, \mathcal{A}, Q)$ be a given probability space. Let S be the price of the asset at time 0. Suppose that there exist two numbers S_d and S_u such that the price at time 1 is a random variable S_1 taking values in $[S_d, S_u]$, and with a density f that is strictly positive on $[S_d, S_u]$. Suppose moreover that $S_d < (1+r)S < S_u$. Let $\mathcal{P}$ be the set of risk-neutral probability measures, that is, the set of probability measures P such that $E_P \left(\dfrac{S_1}{1+r} \right) = S$ (condition (1.4)). We need these probability measures to be equivalent to Q. In other words, we need S_1 to admit under P (or Q) a density function that is strictly positive on $[S_d, S_u]$.

Proposition 1.1.5. *For any convex function g (for example $g(x) = (x - K)^+$), we have*

$$\sup_{P \in \mathcal{P}} E_P \left(\frac{g(S_1)}{1+r} \right) = \frac{g(S_u)}{1+r} \frac{S(1+r) - S_d}{S_u - S_d} + \frac{g(S_d)}{1+r} \frac{S_u - S(1+r)}{S_u - S_d} .$$

If g is of class C^1, we have

$$\inf_{P \in \mathcal{P}} E_P \left(\frac{g(S_1)}{1+r} \right) = \frac{g\big((1+r)S\big)}{1+r} .$$

Proof. Let g be a convex function. Let μ and ν be the slope and y-intersect of the line that goes through the points with coordinates $(S_d, g(S_d))$ and $(S_u, g(S_u))$. We then have:

$$\begin{cases} \forall x \in [S_d, S_u], \ g(x) & \leq \ \mu x + \nu \\ g(S_d) & = \ \mu S_d + \nu \\ g(S_u) & = \ \mu S_u + \nu , \end{cases}$$

and hence, for all $P \in \mathcal{P}$,

$$E_P(g(S_1)) \leq \mu E_P(S_1) + \nu = \mu S(1+r) + \nu .$$

As $\mu = \frac{g(S_u)-g(S_d)}{S_u - S_d}$ and $\nu = g(S_d) - S_d \frac{g(S_u)-g(S_d)}{S_u - S_d}$, we obtain an upper bound.

Let P^* be the probability measure such that

$$\begin{cases} P^*(S_1 = S_u) & = \ p \\ P^*(S_1 = S_d) & = \ 1 - p \\ E_{P^*}(S_1) & = \ S(1+r) . \end{cases}$$

The last condition above determines p (equal to the π appearing in formula (1.3)):

$$p = \frac{S(1+r) - S_d}{S_u - S_d}, \qquad 1 - p = \frac{S_u - S(1+r)}{S_u - S_d} .$$

We have $E_{P^*}(g(S_1)) = \mu S(1+r) + \nu$. The supremum is attained under P^*. We notice that this probability measure does not belong to $\mathcal{P}$, as it does not correspond to the case where S_1 has a strictly positive density function. However we can approach P^* with a sequence of probability measures P_n belonging to $\mathcal{P}$, in the sense that $E_{P^*}(g(S_1)) = \lim E_{P_n}(g(S_1))$.

Similarly, we can obtain a lower bound

$$\inf_{P \in \mathcal{P}} E_P \left(\frac{g(S_1)}{1+r} \right) = \frac{g(S(1+r))}{1+r} .$$

Indeed, if γ and δ are the slope and the y-intersect of the tangent to the curve $y = g(x)$ at the point with coordinates $(S(1+r), g(S(1+r)))$, then

$$g(x) \geq \gamma x + \delta, \quad \gamma S(1+r) + \delta = g(S(1+r)) .$$

Hence $E_P(g(S_1)) \geq E_P(\gamma S_1 + \delta) = g(S(1+r))$, and the minimum is attained by the Dirac measure at $S(1+r)$.   □

This result can be interpreted in terms of volatility. If S_1 takes values in $[S_d, S_u]$ and has expectation $S(1+r)$, then its variance is bounded below by 0 (this value is attained when $S_1 = S(1+r)$), and achieves a maximum when S_1 takes only the extreme values S_d and S_u.

As we remarked earlier, if there does not exist a portfolio that replicates the option, we cannot assign the option a unique price. We define the selling price of the option as the minimal expenditure enabling the seller to hedge himself: it is the smallest amount of money to be invested in a portfolio (α, β) with final value greater than the value of the option $g(S_1)$. Hence the selling price is

$$\inf_{(\alpha,\beta)\in\mathcal{A}} (\alpha + \beta S)$$

with $\mathcal{A} = \{(\alpha,\beta)|\ \alpha(1+r) + \beta x \geq g(x),\ \forall x \in [S_d, S_u]\}$. We have

$$\inf_{(\alpha,\beta)\in\mathcal{A}} (\alpha + \beta S) = \sup_{P\in\mathcal{P}} E_P\left(\frac{g(S_1)}{1+r}\right).$$

Indeed, by definition of $\mathcal{A}$, we have $\alpha(1+r) + \beta S_1 \geq g(S_1)$, and hence

$$\inf_{(\alpha,\beta)\in\mathcal{A}} (\alpha + \beta S) \geq \sup_{P\in\mathcal{P}} E_P\left(\frac{g(S_1)}{1+r}\right).$$

Moreover, using the pair (μ,ν) from the previous section, we can check that $\left(\dfrac{\nu}{1+r}, \mu\right)$ is in $\mathcal{A}$:

$$\inf_{(\alpha,\beta)\in\mathcal{A}} (\alpha + \beta S) \leq \mu S + \frac{\nu}{1+r} = \sup_{P\in\mathcal{P}} E_P\left(\frac{g(S_1)}{1+r}\right).$$

The two problems,

$$\sup_{P\in\mathcal{P}} E_P\left(\frac{g(S_1)}{1+r}\right) \quad \text{and} \quad \inf_{(\alpha,\beta)\in\mathcal{A}} (\alpha + \beta S)$$

are called "dual problems".

We define the buying price of an option as the maximum amount that can be borrowed against the option. The buying price of a call is then defined by:

$$\sup_{(\alpha,\beta)\in\mathcal{C}} (\alpha + \beta S)$$

with $\mathcal{C} = \{(\alpha,\beta)|\ \alpha(1+r) + \beta x \leq g(x),\ \forall x \in [S_d, S_u]\}$.

Similarly, we get:

$$\sup_{(\alpha,\beta)\in\mathcal{C}} (\alpha + \beta S) = \inf_{P\in\mathcal{P}} E_P\left(\frac{g(S_1)}{1+r}\right).$$

1.2 A One-Period Model with $(d+1)$ Assets and k States of the World

We now construct a model that is slightly more complex than the previous one. We consider the case of a one-period market with $(d+1)$ assets and k states of the world. Here again, we do not claim to describe the real world (and nor will we at any point of the book). Instead, we aim to draw out concepts with which we can develop acceptable forms of model.

If S^i is the price at time 0 of the i-th asset $(i = 0, \ldots, d)$, then let its value at time 1 in state j be denoted by v_j^i.

A portfolio $(\theta^0, \theta^1, \ldots, \theta^d)$ is made up of θ^i stocks of type i, and therefore its value at time 0 is $\sum_{i=0}^{d} \theta^i S^i$, and its value at time 1 is $\sum_{i=0}^{d} \theta^i v_j^i$ if we are in the j-th state of the world.

Notation 1.2.1. The column vector S has components S^i, and the column vector θ has components θ^i.

Let V be the matrix of prices at time 1: that is the $(k \times (d+1))$-matrix whose i-th column is made up of the prices of the i-th asset at time 1, that is $(v_j^i, 1 \leq j \leq k)$.

We use matrix notation: $\theta \cdot S = \sum_{i=0}^{d} \theta^i S^i$ is the scalar product of θ and S, and $V\theta$ denotes the $\mathbb{R}^k$-vector with components $(V\theta)_j = \sum_{i=0}^{d} \theta^i v_j^i$.

We write V^T for the matrix transpose of V, and S^T for the vector transpose of S.

A riskless asset is an asset worth $(1+r)$ at time 1 whatever the state of the world, and worth 1 at time 0. The rate of interest r is used as both a lending rate and as a borrowing rate for the sake of simplicity. Thus, $\dfrac{1}{1+r}$ is the price that must be paid at time 0 in order to hold one euro at time 1 in all states of the world.

Notation 1.2.2. $\mathbb{R}_+^k$ denotes the set of vectors of $\mathbb{R}^k$ that have non-negative components. $\mathbb{R}_{++}^k$ denotes the set of vectors of $\mathbb{R}^k$ that have strictly positive components. Δ^{k-1} refers to the unit simplex in $\mathbb{R}^k$:

$$\Delta^{k-1} = \left\{ \lambda \in \mathbb{R}_+^k \mid \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

Let z and z' be two vectors in $\mathbb{R}^k$. We write $z \geq z'$ to express $z_i \geq z_i'$ for all i.

Exercise 1.2.3. Show that if V is a $k \times (d+1)$-matrix, then there is an equivalence between the statements:

(i)   The rank of the mapping associated with V is k.
(ii) The linear mapping associated with V is surjective, and the one associated with V^T is injective.

1.2.1 No Arbitrage Opportunities

We now introduce the concept of an arbitrage opportunity, which was touched upon earlier.

The Assumption of No Arbitrage Opportunities

Definition 1.2.4. *An arbitrage is a portfolio $\theta = (\theta^0, \theta^1, \ldots, \theta^d)$ with a non-positive initial value $S \cdot \theta = \sum_{i=0}^{d} \theta^i S^i$ and a non-negative value $V\theta$ at time 1, with at least one strict inequality. In other words, either $S \cdot \theta < 0$ and $V\theta \geq 0$, or $S \cdot \theta = 0$ and $V\theta \geq 0$ with a strict inequality in at least one state of the world.*

*We say that there are **no arbitrage opportunities** when there is no arbitrage. That is to say, the following conditions must hold:*

(i) $V\theta = 0$ implies $S \cdot \theta = 0$,

(ii) $V\theta \geq 0$, $V\theta \neq 0$ implies $S \cdot \theta > 0$.

Indeed, in the first case, if we had $V\theta = 0$ and $S \cdot \theta < 0$ (or $S \cdot \theta > 0$), then the portfolio θ (or $-\theta$) would be an arbitrage. In the second case, if $V\theta \geq 0$, $V\theta \neq 0$ and $S \cdot \theta \leq 0$, then θ would be an arbitrage.

An arbitrage opportunity is a means of obtaining wealth without any initial capital. Obviously an arbitrage opportunity could not exist without being very quickly exploited. We therefore make the following assumption, referred to as the assumption of *no arbitrage opportunity* (NAO).

The NAO Assumption: there exists no arbitrage opportunity.

Using the same notation as before, we recall a result from linear programming:

Lemma 1.2.5 (Farkas' Lemma). *The implication $V\theta \geq 0 \Rightarrow S \cdot \theta \geq 0$ holds if and only if there exists a sequence $(\beta_j)_{j=1}^{k}$ of non-negative numbers such that $S^i = \sum_{j=1}^{k} v_j^i \beta_j, \quad i \in \{0, \ldots, d\}$.*

We remark that the assumption of NAO is a little bit stronger than the assumptions of Farkas' Lemma, as according to the former, if the portfolio's payments are non-negative, and strictly positive in at least one state of the world, then the price of the portfolio is strictly positive. From this we will deduce (with a proof that is in fact simpler than that of Farkas' Lemma) that the β_j are strictly positive.

We recall the Minkowski separation theorem.

Theorem 1.2.6 (The Minkowski Separation Theorem). *Let C_1 and C_2 be two non-empty disjoint convex sets in $\mathbb{R}^k$, where C_1 is closed and C_2 is compact. Then there exists a family $(a_1, \ldots, a_k)$ of non-zero coefficients, and two distinct numbers b_1 and b_2 such that*

$$\forall\, x \in C_1, \,\forall\, y \in C_2, \qquad \sum_{j=1}^{k} a_j x_j \;\leq\; b_1 \;<\; b_2 \;\leq\; \sum_{j=1}^{k} a_j y_j \;.$$

Theorem 1.2.7. *The NAO assumption is equivalent to the existence of a sequence $(\beta_j)_{j=1}^{k}$ of strictly positive numbers, called state prices, such that*

$$S^i \;=\; \sum_{j=1}^{k} v_j^i \,\beta_j; \quad i \in \{0, \ldots, d\} \;. \tag{1.6}$$

Proof. (of Theorem 1.2.7)
Let S^T be the row vector $(S^0, S^1, \ldots, S^d)$ and let U be the vector subspace of $\mathbb{R}^{k+1}$

$$U := \left\{ z \in \mathbb{R}^{k+1} \mid z = \begin{pmatrix} -S^T \\ V \end{pmatrix} x; \; x \in \mathbb{R}^{d+1} \right\}.$$

The assumption of NAO implies that $U \cap \mathbb{R}_+^{k+1} = \{0\}$, so that in particular, $U \cap \Delta^k = \emptyset$. According to Minkowski's theorem, there exists a set of non-zero coefficients $\{\beta_j \,;\, j = 0, \ldots, k\}$ and two numbers b_1 and b_2, such that

$$\sum_{j=0}^{k} \beta_j z_j \;\leq\; b_1 \;<\; b_2 \;\leq\; \sum_{j=0}^{k} \beta_j w_j; \qquad z \in U, \; w \in \Delta^k \;.$$

As $0 \in U$, $b_1 \geq 0$, and hence, by choosing a vector w whose components are all zero except for the j-th, which is equal to 1, we deduce that $\beta_j > 0$, $\forall\, j \in \{0, \ldots, k\}$. Without loss of generality, we can take $\beta_0 = 1$.
Then let β be the vector $(\beta_1, \ldots, \beta_k)^T$. Taking into account the form of the elements of U, we write the inequality $z_0 + \sum_{j=1}^{k} \beta_j z_j \;\leq\; 0$ as $(-S + V^T \beta) \cdot x \leq 0$. Hence $S = V^T \beta$, i.e., $S^i = \sum_{j=1}^{k} \beta_j v_j^i$ with $\beta_j > 0$, $j \in \{1, \ldots, k\}$.

The proof of the converse is trivial.   □

The vector β is called a state price vector: β_j corresponds to the price at time 0 of a product that is worth 1 at time 1 in state j, and 0 in all the other states. We will come back to this interpretation later.

Probabilistic Interpretation of the State Prices

Until now in this section, we have not used probabilities. We will now give a probabilistic interpretation of the NAO assumption and of Theorem 1.2.7. Introducing probabilities will enable us to study more general models, and to exploit the concept of NAO.

If asset 0 is riskless, then we have

$$v_j^0 = 1 + r, \quad j \in \{1, \ldots, k\},$$

and hence, using (1.6), for $i = 0$:

$$\frac{1}{1+r} = \sum_{j=1}^{k} \beta_j.$$

Let us set $\pi_j = (1+r)\beta_j$. The π_j are positive numbers such that $\sum_{j=1}^{k} \pi_j = 1$. Therefore, they can be interpreted as probabilities on the different states of the world. We have

$$S^i = \frac{1}{1+r} \sum_{j=1}^{k} \pi_j v_j^i \quad i \in \{1, \ldots, d\}.$$

We have thus constructed a probability measure under which the price S^i of the i-th asset is the expectation of its price at time 1, discounted using the riskless rate.

If we construct a portfolio θ, we get:

$$(1+r) \sum_{i=0}^{d} \theta^i S^i = \sum_{j=1}^{k} \pi_j \sum_{i=0}^{d} \theta^i v_j^i,$$

where π is (as in Sect. 1.1) a probability measure that is neutral with respect to risk: a riskless investment with initial value $\sum_{i=0}^{d} \theta^i S^i$ yields $(1+r) \sum_{i=0}^{d} \theta^i S^i$, which is equal to the expectation (under π) of the value of the portfolio at time 1.

The rate of return on asset i in state j is by definition equal to $(v_j^i - S^i)/S^i$. The expectation of the rate of return on i is, under probability measure π, equal to the rate of return on the riskless asset:

$$\sum_{j=1}^{k} \pi_j \frac{v_j^i - S^i}{S^i} = r .$$

Proposition 1.2.8. *Under the assumption of NAO, if asset 0 is riskless, then there exists a probability measure π on the states of the world, under which the price at time 0 of asset i is equal to the expectation of its price at time 1, discounted by the riskless rate:*

$$S^i = \frac{1}{1+r} \sum_{j=1}^{k} \pi_j v_j^i . \tag{1.7}$$

Exercise 1.2.9. Let $V\theta$ be the vector with components $(V\theta)_j = \sum_{i=0}^{d} v_j^i \, \theta^i$,

and let $S \cdot \theta$ denote the scalar product $S \cdot \theta = \sum_{i=0}^{d} S^i \, \theta^i$.

a. Let $z \in \operatorname{Im} V$. Let θ be any vector satisfying $z = V\theta$. Show that, under the assumption of NAO, the mapping $\pi : z \to S \cdot \theta$ does not depend on the choice of θ, and defines a positive linear functional on $\operatorname{Im} V$.

b. Show that π can be extended to a positive linear functional $\bar{\pi}$ on $\mathbb{R}^k$. To do this, show that for all $\hat{z} \notin \operatorname{Im} V$, there exists $\phi(\hat{z}) \in \mathbb{R}$ such that

$$\max \{\pi(z'), \ z' \le \hat{z}, \ z' \in \operatorname{Im} V\} < \phi(\hat{z}) < \min \{\pi(z'), \ z' \ge \hat{z}, \ z' \in \operatorname{Im} V\} .$$

Next show that the mapping $z + \lambda \hat{z} \to \pi(z) + \lambda \phi(\hat{z})$ is linear and positive, and extends π to the space generated by $\operatorname{Im} V$ and $\hat{z}$.

c. Show, using the Riesz representation theorem, that $\bar{\pi}(z) = \beta \cdot z$ with $\beta \in \mathbb{R}_{++}^k$.

d. Thence deduce Theorem 1.2.7

Exercise 1.2.10. Suppose that there are constraints on portfolios, modeled by a closed convex cone $\mathcal{C}$: $\theta \in \mathcal{C}$. For example:

$$\begin{cases} \theta^i \ unconstrained \ for \ 0 \le i \le r \\ \theta^i \ge 0 \qquad\qquad for \ r+1 \le i \le r+p \\ \theta^i \le 0 \qquad\qquad for \ r+p+1 \le i \le d. \end{cases}$$

Adapt the definition of NAO to the restriction to $\mathcal{C}$.

1. Suppose that there are $k \ge 4$ states of the world, and 4 assets. Asset 0 is riskless, and the rate of interest is r. The other assets are risky, and their returns are given by a matrix V. Suppose that the constraints are $\theta^2 \ge 0$

and $\theta^3 \leq 0$. Let $\tilde{V} = \begin{bmatrix} \ldots \ldots V \ldots \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$.

Show that NAO with restrictions on portfolios can be expressed as
(i) $\tilde{V}\theta = 0 \implies S \cdot \theta = 0$.
(ii) $\tilde{V}\theta \in \mathbb{R}_+^k$, $\tilde{V}\theta \neq 0 \implies S \cdot \theta > 0$.

Hence deduce that there exists a probability measure π such that

$$S^1 = \frac{1}{1+r} \sum_{j=1}^{k} v_j^1 \pi_j, \ S^2 \geq \frac{1}{1+r} \sum_{j=1}^{k} v_j^2 \pi_j \text{ and } S^3 \leq \frac{1}{1+r} \sum_{j=1}^{k} v_j^3 \pi_j \ .$$

2. For $\bar{\theta} \in \mathcal{C}$, write $N_{\mathcal{C}}(\bar{\theta}) = \{p \in \mathbb{R}^{d+1} | p^T(\theta - \bar{\theta}) \leq 0, \forall \theta \in \mathcal{C}\}$. Show, by generalizing the proof of Theorem 1.2.7, that NAO restricted to $\mathcal{C}$, is equivalent to the existence of $\beta \in \mathbb{R}_{++}^k$ such that $-S + V^T\beta \in N_{\mathcal{C}}(0)$.

3. Recover the results of 1.

Exercise 1.2.11.

1. Suppose that there are 2 states of the world, and 2 assets, one riskless (the rate of interest is taken to be r) and the other a stock worth either S_u or S_d at time 1. Suppose that the risky asset has purchase price S_0 and selling price $S_0' \leq S_0$. We use the notation $\theta = \theta^+ - \theta^-$ for the amount of stock held, and θ^0 for the amount of riskless asset held. The cost of this portfolio is then $\theta^0 + \theta^+ S_0 - \theta^- S_0'$, and it pays

$$\begin{cases} (1+r)\theta^0 + (\theta^+ - \theta^-)S_u \text{ in the high state, after an up-move} \\ (1+r)\theta^0 + (\theta^+ - \theta^-)S_d \text{ in the low state, after a down-move} \end{cases}$$

Show, using Farkas' Lemma, that there is NAO if and only if there exists at least one probability measure π such that

$$S_0' \leq \frac{S_u \pi}{1+r} + \frac{S_d(1-\pi)}{1+r} \leq S_0 \ .$$

The reader can introduce the matrix:

$$\begin{bmatrix} (1+r) & S_u & -S_u \\ (1+r) & S_d & -S_d \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2. Suppose that there are d assets, with an injective gains matrix V. Suppose that the cost $\phi(\theta)$, $\phi : \mathbb{R}^d \to \mathbb{R}$ of a portfolio $\theta \in \mathbb{R}^d$ is a sublinear function, that is, one satisfying

$$\begin{cases} \phi(\theta_1 + \theta_2) \le \phi(\theta_1) + \phi(\theta_1) & \forall (\theta_1, \theta_2) \in \mathbb{R}^{2d} \\ \phi(t\theta) = t\phi(\theta) & \forall\, t \ge 0 \,. \end{cases}$$

Notice that in particular, we have $\phi(0) = 0$ and $-\phi(-\theta) \le \phi(\theta)$. Let $U := \{(z_1, z_2) \in \mathbb{R}^k \times \mathbb{R} \mid \exists \theta \text{ such that } z_1 \le -\phi(\theta) \text{ and } z_2 = V\theta\}$. Show that U is a convex cone.

We say that there is NAO if $V\theta = 0 \Rightarrow \phi(\theta) = 0$, and $V\theta \in \mathbb{R}^k_+$, $V\theta \ne 0 \Rightarrow \phi(\theta) > 0$. Show that under the assumption of NAO, $U \cap \mathbb{R}^{k+1}_+ = \{0\}$. Show, by adapting the proof of Theorem 1.2.7, that NAO is equivalent to the existence of a strictly positive β such that

$$-\phi(-\theta) \le \beta^T V\theta \le \phi(\theta).$$

Hence recover the results of the first question.

1.2.2 Complete Markets

Definition and Characterization

Definition 1.2.12. *A market is complete if, for any vector w of $\mathbb{R}^k$, we can find a portfolio θ such that $V\theta = w$; that is to say, there exists θ such that*

$$\sum_{i=0}^{d} \theta^i\, v^i_j \;=\; w_j, \quad j \in \{1, \dots, k\} \,.$$

A market is complete if we can choose a portfolio at time 0 in such a way as to attain any given vector of wealth at time 1.

Proposition 1.2.13. *A market is complete if and only if the matrix V is of rank k.*

Proof. Matrix V has rank k if and only if the mapping associated with V is surjective; the equation $V\theta = w$ then has at least one solution. □

Economic Interpretation of State Prices

In a complete market, for any $j \in \{1, \dots, k\}$, there exists a portfolio θ_j such that the payoff of θ_j satisfies $V\theta_j = (\delta_{1,j}, \dots, \delta_{k,j})^T$, with $\delta_{i,j} = 0$ when $i \ne j$, and $\delta_{j,j} = 1$ (the asset is then called an Arrow–Debreu asset). In an arbitrage-free market, the initial value of θ_j is $S \cdot \theta_j = \beta^T V\theta_j = \beta_j$. Therefore we can interpret β_j as the price to be paid at time 0 in order to have one euro at time 1 in state j and nothing in the other states of the world. Hence the terminology "state price".

Moreover, we note that if there exists β such that $V^T\beta = S$, then, as the mapping associated with matrix V^T is injective, the vector β is unique.

The Risk-Neutral Probability Measure

In a complete market, there necessarily exists a riskless portfolio, that is a portfolio θ such that $(V\theta)_j = a$ for all $j \in \{1, \ldots, k\}$. The initial value of this portfolio is taken to be V_0. The rate of return on the portfolio is $(a - V_0)/V_0$, and will be denoted by r. Without loss of generality, we can assume asset 0 to be riskless, and we can normalize its price so that it is 1 at time 0, its value at time 1 being $1 + r$. If there exists a probability measure π satisfying $V^T \pi = (1+r)S$, then it is unique. We then call it the "risk-neutral measure".

1.2.3 Valuation by Arbitrage in the Case of a Complete Market

Let z be a vector of $\mathbb{R}^k$. Under the assumption of NAO, if there exists a portfolio $\theta = (\theta^0, \theta^1, \ldots, \theta^d)$ taking the value z at time 1, i.e., such that

$$\sum_{i=0}^{d} \theta^i \, v_j^i \; = \; z_j \, ,$$

then we say that z is replicable. The value of the portfolio at time 0 is $z_0 = \sum_{i=0}^{d} \theta^i S^i$, and this value does not depend on the hedging portfolio chosen. Indeed, suppose that there exist two portfolios θ and $\widetilde{\theta}$ such that $V\theta = V\widetilde{\theta}$ and $S \cdot \theta > S \cdot \widetilde{\theta}$. The portfolio $\widetilde{\theta} - \theta$ is an arbitrage opportunity. In the complete market framework, there always exists a hedging portfolio.

Proposition 1.2.14. *In a complete and arbitrage-free market, the initial value of the payoff $z \in \mathbb{R}^k$, delivered at time 1, is given by*

$$\frac{1}{1+r} \sum_{j=1}^{k} \pi_j \, z_j \; = \; \sum_{j=1}^{k} \beta_j \, z_j \, .$$

Remark 1.2.15. The initial value of z is a linear function of z.

Proof. (of Proposition 1.2.14)

The value of any hedging portfolio is $z_0 = \sum_{i=0}^{d} \theta^i \, S^i$. It is enough to use (1.6) or (1.7) and write

$$z_0 \; = \; \sum_{i=0}^{d} \theta^i \sum_{j=1}^{k} \beta_j \, v_j^i \; = \; \beta \cdot z = \; \frac{1}{1+r} \sum_{j=1}^{k} \pi_j z_j \, .$$

$\square$

Remark 1.2.16. The expression above has a two-fold advantage. It does not depend on the portfolio, and it can be interpreted, as follows: the price at time 0 of the replicating portfolio $(z_j; \; j = 1, \ldots, k)$ is the discounted expectation under π of its value at time 1.

In the case of an option on the i-th asset, we have $z_j = \sup(v_j^i - K, 0)$, and hence we get the arbitrage price

$$\frac{1}{1+r} \sum_{j=1}^{k} \pi_j \sup(v_j^i - K, 0) .$$

1.2.4 Incomplete Markets: the Arbitrage Interval

Generally speaking, it is not possible to valuate a product by arbitrage in an incomplete market. If z is not replicable, then we can define an *arbitrage interval*. We associate with any portfolio θ, its corresponding initial value $\theta \cdot S$.

We define the selling price of z as the smallest amount of wealth that can be invested in a portfolio θ in such a way that the final value of this portfolio is greater than z. In the following, we suppose that there is a riskless asset. The super-replication price is then

$$\overline{S}(z) := \inf\{\theta \cdot S \mid (V\theta)_j \geq z_j; \, \forall j\} .$$

We define the purchase price of z as the maximal amount of money that can be borrowed against z, i.e.,

$$\underline{S}(z) := \sup\{\theta \cdot S \mid (V\theta)_j \leq z_j; \, \forall j\} .$$

First we note that $\overline{S}(z)$ is well-defined. Indeed, let us take $\tilde{\theta}$ to be an element of the non-empty set $\{\theta \mid V\theta \geq z\}$. The set $\{\theta \mid S \cdot \theta \leq S \cdot \tilde{\theta} \text{ and } V\theta \geq z\}$ is a compact set (from the NAO condition), on which the function $S \cdot \theta$ attains its minimum.

Moreover, we can easily show that if $\overline{S}(z) \neq \underline{S}(z)$ and if the price $S(z)$ of the contingent asset z satisfies $S(z) \geq \overline{S}(z)$ or $S(z) \leq \underline{S}(z)$, then an arbitrage occurs if we use strategies that include this new asset. If $\overline{S}(z) \neq \underline{S}(z)$ and if the price $S(z)$ of the contingent asset z satisfies $\underline{S}(z) < S(z) < \overline{S}(z)$, then there is NAO when we use strategies that include this new asset. Let us show that, indeed, if a portfolio (θ_z, θ) satisfies $\theta_z z + V\theta \geq 0$ and $\theta_z z + V\theta \neq 0$, then $\theta_z S(z) + S \cdot \theta > 0$.

- If $\theta_z = 0$, it follows from NAO.

- If $\theta_z < 0$, we have $V \frac{\theta}{-\theta_z} \geq z$, so that $S \cdot \frac{\theta}{-\theta_z} \geq \overline{S}(z) > S(z)$, and hence $\theta_z S(z) + S \cdot \theta > 0$.

- If $\theta_z > 0$, we have $z \geq V \frac{\theta}{-\theta_z}$, so that $S \cdot \frac{\theta}{-\theta_z} \leq \underline{S}(z) < S(z)$, and hence $\theta_z S(z) + S \cdot \theta > 0$.

In addition, we check that $\theta_z z + V\theta = 0$ implies $\theta_z S(z) + S \cdot \theta = 0$. Indeed, as $V\frac{\theta}{-\theta_z} = z$, so $\underline{S}(z) = \bar{S}(z) = S(z) = S \cdot \frac{\theta}{-\theta_z}$, and hence $\theta_z S(z) + S \cdot \theta = 0$.

Therefore, there is NAO when we use strategies that include the new asset. Finally, $\overline{S}(z)$ is sublinear: it satisfies

$$\overline{S}(z + z') \le \overline{S}(z) + \overline{S}(z') \quad \text{and} \quad \overline{S}(az) = a\overline{S}(z) \quad \forall a \in \mathbb{R}_+ .$$

Moreover, $-\overline{S}(-z) = \underline{S}(z)$.

Let us now show that

$$\overline{S}(z) = \max\{\beta^T z \mid \beta \ge 0, \ V^T\beta = S\} .$$

Indeed, for any θ such that $V\theta \ge z$, and any $\beta \ge 0$ such that $V^T\beta = S$, we have $S \cdot \theta = \beta^T V\theta \ge \beta^T z$. Hence

$$\min\{S \cdot \theta \mid V\theta \ge z\} \ \ge \ \max\{\beta^T z \mid \beta \ge 0, \ V^T\beta = S\} .$$

In addition, if $\bar{\theta}$ is a solution to $\min_{\{V\theta \ge z\}} S \cdot \theta$, then there exists a Lagrange multiplier[3] $\bar{\beta} \ge 0$ such that $S = V^T\bar{\beta}$ and $\bar{\beta}^T(V\bar{\theta} - z) = 0$. Hence

$$\overline{S}(z) = S \cdot \bar{\theta} = \bar{\beta}^T z \ \le \ \max\{\beta^T z \mid \beta \ge 0, \ V^T\beta = S\} .$$

The required equality follows.

If there is a riskless asset, we can normalize $\bar{\beta}$, and hence

$$\overline{S}(z) = \max\left\{ \frac{E_\pi(z)}{1+r} \ \middle| \ V^T\pi = (1+r)S \right\} .$$

The expression above represents the maximum of the expectation across all the probability measures under which discounted prices are martingales. In this way, we have generalized the results of Sect. 1.1.6.

Exercise 1.2.17. Arbitrage bounds in the presence of portfolio constraints.

We use the notation of Exercise 1.2.10, and restrict ourselves to portfolios belonging to $\mathcal{C}$. Let

$$\overline{S}(z) := \inf\{\theta \cdot S \mid \theta \in \mathcal{C}, V\theta \ge z\} .$$

(If there exists no $\theta \in \mathcal{C}$ such that $V\theta \ge z$, we set $\overline{S}(z) := \infty$).

1. Show that $\overline{S}(z)$ is well-defined and sublinear.

---

[3] See annex.

2. Show that $\overline{S}(z) = \max\{\beta^T z \mid \beta \geq 0, -S + V^T \beta \in N_{\mathcal{C}}(0)\}$. (Recall that if $\overline{\theta}$ minimizes $\theta \cdot S$ under the constraints $\theta \in \mathcal{C}$ and $V\theta \leq z$, then there exists $\beta \geq 0$ and $v \in N_{\mathcal{C}}(\overline{\theta})$ such that $S = \beta^T V - v$ and $\beta^T(V\overline{\theta} - z) = 0$.)

Exercise 1.2.18. Arbitrage bounds in the case of transaction costs.

We use the assumptions and notation of Exercise 1.2.11. Suppose that there is a riskless asset. In addition, for any z there exists θ such that $V\theta \geq z$, and we define

$$\overline{S}(z) := \inf\{\phi(\theta) \mid V\theta \geq z\} \ .$$

1. Show that $\overline{S}(z)$ is well-defined, and sublinear.

2. Show that $\overline{S}(z) = \max\{\beta^T z \mid \beta \geq 0, \phi(\theta) \geq \beta^T V\theta, \ \forall \theta\}$.

3. Calculate the purchase price for a call with strike K, where the rest of the data is as in as in question 1 of Exercise 1.2.11. (First consider the case $S_0 \geq \frac{S_u}{1+r}$, and next the case $S_0 < \frac{S_u}{1+r}$).

1.3 Optimal Consumption and Portfolio Choice in a One-Agent Model

The two models introduced previously were purely financial. We now consider a very simple economy, which has a single good for consumption, taken as the numéraire, and a single economic agent. This agent has known resources $R_0 > 0$ a time 0, and his resources at time 1 given by $R_j > 0$ in state of the world j.

In order to modify his future revenue, the agent can buy a portfolio of assets at time 0, on condition that he does not run into debt. We assume that the $(d + 1)$ assets have the same characteristics as in the previous section.

The agent consumes: c_0 is the amount of his consumption at time 0; c_j that of his consumption at time 1 in state of the world j.

The agent constructs a portfolio θ. The set of consumption–portfolio pairs that are compatible with the agent's revenue, is defined by the following inequalities:

$$
\begin{aligned}
\text{(i)} \quad & R_0 \ \geq \ c_0 + \sum_{i=0}^{d} \theta^i \, S^i \\
\text{(ii)} \quad & R_j \ \geq \ c_j - \sum_{i=0}^{d} \theta^i \, v_j^i, \quad j \in \{1, \dots, k\} \ .
\end{aligned}
\tag{1.8}
$$

The first constraint states that money invested in the portfolio comes from the portion of revenue that has not been consumed, and the second, that consumption at time 1 is covered by his resources and by the portfolio.

The set of consumption strategies that are compatible with the agent's revenue is then:

$$B(S) := \left\{ c \in \mathbb{R}_+^{k+1}; \; \exists\, \theta \in \mathbb{R}^{d+1}, \text{ satisfying } (1.8) \right\}.$$

The agent has "preferences" on $\mathbb{R}_+^{k+1}$, that is to say, a preorder (a reflexive and transitive binary relation), written $\succeq$, which is complete (any two elements of $\mathbb{R}_+^{k+1}$ can be compared). We say that $u : \mathbb{R}_+^{k+1} \to \mathbb{R}$ is a utility function that represents the preorder of preferences if $u(c) \geq u(c')$ is equivalent to $c \succeq c'$. Historically, the concept of a utility function came before that of a preorder of preferences. Utility functions have long been part of the basis of economic theory ("marginalist" theory). Later, much work sought to give foundations to utility theory, by taking the preorders as a starting point.

We assume here that the investor's preferences are represented by a function u from $\mathbb{R}_+^{k+1}$ into $\mathbb{R}$, which is strictly increasing with respect to each of its variables, strictly concave and differentiable. We suppose that the agent maximizes his utility under budgetary constraints (1.8). The derivative u' is called the marginal utility. We assume moreover that $\dfrac{\partial u}{\partial c_i}(c_0, \ldots, c_i, \ldots, c_k) \to \infty$ when $c_i \to 0$. This condition means that the agent has a strong aversion to consuming nothing at time 0 or at time 1 in one of the states of the world.

1.3.1 The Maximization Problem

Let u be a utility function. We say that $c^* \in B(S)$ is an optimal consumption if

$$u(c^*) = \max \left\{ u(c); \; c \in B(S) \right\}.$$

Existence of an Optimal Consumption

Proposition 1.3.1. *There is an optimal solution if and only is S satisfies the NAO assumption. The optimal solution is strictly positive.*

Proof. Suppose that there exists an optimal solution (c_0^*, c_1^*) financed by θ^*, and an arbitrage θ^a. We then have $S \cdot \theta^a \leq 0$ and $V\theta^a \geq 0$ where at least one of the inequalities is strict. It is true that the consumption $(c_0^* - S \cdot \theta^a, c_1^* + V\theta^a) \in B(S)$ (an associated portfolio is $\theta^* + \theta^a$). Using the property of an arbitrage strategy, $c_0^* - S \cdot \theta^a \geq c_0^*$, $c_1^* + V\theta^a \geq c_1^*$ with at least one strict inequality. As $u(c)$ is strictly increasing, this contradicts the optimality of (c_0^*, c_1^*).

Conversely, let us show that under the assumption of NAO, if V is injective, then there exists an optimal solution. A more general result will be proved in Chap. 6. Let us show that the set

$$\left\{ \theta \in \mathbb{R}^{d+1}; \ \exists \, c \in \mathbb{R}_+^{k+1}, \text{ satisfying } (1.8) \right\}$$

is bounded. Suppose that, on the contrary, there exists a sequence (c_n, θ_n) satisfying (1.8), and such that $\|\theta_n\| \to \infty$, and let $\hat{\theta}$ be a limit point of the sequence $\frac{\theta_n}{\|\theta_n\|}$. We have

$$\frac{S \cdot \theta_n}{\|\theta_n\|} + \frac{c_{0n}}{\|\theta_n\|} \leq \frac{R_0}{\|\theta_n\|} \ ,$$

$$\frac{c_{jn}}{\|\theta_n\|} \leq \frac{R_j}{\|\theta_n\|} + \frac{V\theta_n}{\|\theta_n\|} \ ,$$

for all n, and hence $S \cdot \hat{\theta} \leq 0$ and $V\hat{\theta} \geq 0$. By the NAO assumption, $V\hat{\theta} = 0$, and $\hat{\theta} = 0$, so contradicting the fact that $\|\hat{\theta}\| = 1$. We deduce that $B(S)$ is closed and bounded, and thus compact, and hence that an optimal solution c^* does exist. Let us now show that c^* is strictly positive.

As the function u is strictly increasing, the budget constraints (1.8) are binding. Hence there exists θ^* such that

$$\begin{cases} c_0^* + \displaystyle\sum_{i=0}^d \theta^{i*} S^i - R_0 \ = \ 0 \\[2mm] c_j^* - \displaystyle\sum_{i=0}^d \theta^{i*} v_j^i - R_j \ = \ 0, \quad j \in \{1, \ldots, k\} \ . \end{cases}$$

Let ε satisfy $c_0^* + \varepsilon S \cdot \theta^* > 0$ and $c_j^* - \varepsilon (V\theta^*)_j > 0$ for any $j \in \{1, \ldots, k\}$. The consumption $(c_0, c_1, \ldots, c_k)$ where $c_0 = \varepsilon S \cdot \theta^* + c_0^*$ and $c_j = c_j^* + \varepsilon (V\theta^*)_j$ for any $j \in \{1, \ldots, k\}$ is in $B(S)$ (an associated portfolio is $(1 - \varepsilon)\theta^*$). As u is concave,

$$u(c) - u(c^*) \geq \varepsilon \left(S \cdot \theta^* \frac{\partial u}{\partial c_0}(c) - \sum_{j=1}^k (V\theta^*)_j \frac{\partial u}{\partial c_j}(c) \right) \ .$$

For ε small enough, if $c_0^* = 0$ or if $c_j^* = 0$ for $j \in \{1, \ldots, k\}$, the last expression above is strictly positive: since if $c_0^* = 0$ (respectively $c_j^* = 0$), $S \cdot \theta^* = R_0 > 0$ (respectively $(V\theta^*)_j < 0$), and when $\varepsilon \to 0$, $\frac{\partial u}{\partial c_0}(c) \to \infty$ (respectively $\frac{\partial u}{\partial c_j}(c) \to \infty$). This contradicts the optimality of c^*. □

Remark 1.3.2. It is important to take note of the conditions under which this proposition holds. In the first part of the proof, we used the fact that u is **strictly** increasing with respect to all of its variables. In the second part, we used the non-negativity of consumption. The following exercises provide very simple counterexamples to the statement of the proposition when these conditions are no longer satisfied.

Exercise 1.3.3.

1. Consider an economy in which there are two dates, 0 and 1. At time 1, there are two possible states of the world. At time 0, an agent holding one euro, can buy a portfolio made up of two assets whose payoffs are represented by the payment vectors $[1, 0]$ and $[0, 1]$ respectively, and whose prices are $S^1 = 1$, $S^2 = 0$. Further assume that the agent consumes c_0. At time 1, in addition to the payment vector of his portfolio, the agent receives $[1, 2]$ and consumes (c_1, c_2). Suppose that the agent has utility function

$$u(c_0, c_1, c_2) = c_0 + \min\{c_1, c_2\} .$$

   Show that the agent's consumption–portfolio problem admits a solution (notice that the maximum utility that the agent can achieve, is 2). Is the solution unique? Show that the financial market admits an arbitrage. Comment on these results.

2. The data here is the same as that of the previous question, except that the agent's utility function is given by

$$u(c_0, c_1, c_2) = -(c_0 - 1)^2 - (c_1 - 1)^2 - (c_2 - 2)^2 .$$

   Show that the agent's consumption–portfolio problem admits a solution. Comment on the result.

3. We no longer assume the consumption to be positive. At time 0, an agent holding one euro, can buy an asset, whose payment vector is $[1, 1]$, and whose price is $S^1 = 1$. At time 1, in addition to the payment vector of his portfolio, the agent receives $[1, 2]$. We assume that his utility function is

$$u(c_0, c_1, c_2) = c_0 + c_1 + c_2 .$$

   Show that the financial market does not admit arbitrage, and that nevertheless, the agent's consumption–portfolio problem does not admit a solution.

Asset Valuation Formula

As c^* is strictly positive, it follows from the method of Lagrange multipliers, that a necessary and sufficient condition for c^* to be optimal, is for there to exist $\theta^* \in \mathbb{R}^{d+1}$ and $\lambda^* \in \mathbb{R}_+^{k+1}$ such that

$$\begin{cases} \dfrac{\partial u}{\partial c_0}(c^*) - \lambda_0^* = 0 \\[2mm] \dfrac{\partial u}{\partial c_j}(c^*) - \lambda_j^* = 0 , \qquad j \in \{1, \ldots, k\} , \end{cases} \tag{1.9.i}$$

$$\lambda_0^* S^i - \sum_{j=1}^{k} \lambda_j^* v_j^i \;=\; 0 \,, \qquad i \in \{0, \dots, d\} \,, \tag{1.9.ii}$$

$$\begin{cases} \lambda_0^* \left(c_0^* + \sum_{i=0}^{d} \theta^{i*} S^i - R_0 \right) \;=\; 0 \\[2mm] \lambda_j^* \left(c_j^* - \sum_{i=0}^{d} \theta^{i*} v_j^i - R_j \right) \;=\; 0 \,, \qquad j \in \{1, \dots, k\} \,. \end{cases} \tag{1.9.iii}$$

The assumption that u is strictly increasing, implies that its derivatives are strictly positive. Hence, from (1.9.i), we have $\lambda^* \in \mathbb{R}_{++}^{k+1}$ and we can write expression (1.9.iii) as

$$\begin{cases} c_0^* + \sum_{i=0}^{d} \theta^{i*} S^i - R_0 \;=\; 0 \\[2mm] c_j^* - \sum_{i=0}^{d} \theta^{i*} v_j^i - R_j \;=\; 0, \quad j \in \{1, \dots, k\} \,. \end{cases} \tag{1.9.iv}$$

Defining β_j as

$$\beta_j \;=\; \frac{\lambda_j^*}{\lambda_0^*} \;=\; \frac{\partial u/\partial c_j}{\partial u/\partial c_0} \, (c^*) \,, \tag{1.10}$$

the β_j are strictly positive, and, using (1.9.ii), we obtain a formula for evaluating the price of the assets:

$$S^i \;=\; \sum_{j=1}^{k} \beta_j \, v_j^i \,. \tag{1.11}$$

The interest rate is given by the expression:

$$1 + r = \frac{\partial u/\partial c_0(c^*)}{\displaystyle\sum_{j=1}^{k} \partial u/\partial c_j(c^*)} \,. \tag{1.12}$$

Finally, eliminating θ^{i*} from the equations in (1.9.iv), we get

$$c_0^* + \sum_{j=1}^{k} \beta_j c_j^* \;=\; R_0 + \sum_{j=1}^{k} \beta_j R_j \,.$$

The Complete Market Case

In the case of a complete market with no arbitrage, the optimization problem under constraints, defined by (1.8), takes a simpler form. As the market is complete, there exists a unique β such that $S = V^T\beta$. Let us define the inequality

$$c_0 + \sum_{j=1}^{k} \beta_j c_j \leq R_0 + \sum_{j=1}^{k} \beta_j R_j . \tag{1.13}$$

This is the budgetary constraint placed on an agent who buys a consumption of c_j at a contingent price of β_j.

If the market is complete

$$B(S) = \left\{ c \in \mathbb{R}_+^{k+1} \text{ satisfying } (1.13) \right\} .$$

Indeed, if $c \in B(S)$, using (1.9.iv) to eliminate θ, we can show that c satisfies (1.13).

Conversely, let c satisfy (1.13). If the market is complete, there exists θ such that

$$c_j - \sum_{i=0}^{d} \theta^i v_j^i - R_j = 0 \;\; \text{for all} \;\; j \in \{1,\ldots,k\} .$$

Using (1.9.iv) and (1.13), we can show that (1.8) is satisfied, and hence that $c \in B(S)$.

Thus we are brought back to a maximization problem under a single budgetary constraint $c_0 + \sum_{j=1}^{k} c_j \beta_j \leq \sum_{j=1}^{k} R_j \beta_j + R_0$. Formula (1.10) then follows trivially. We observe that the price in state j is proportional to the marginal utility of consumption in state j.

As we showed previously, if there is a riskless asset, the β_j can be interpreted in terms of risk-neutral probabilities $\beta_j = \frac{\pi_j}{1+r}$. The risk-neutral probability of state j is therefore proportional to the marginal utility of consumption in state j. We note that by using risk-neutral probabilities, we can write constraint (1.12), if asset 0 is riskless, as

$$c_0 + \sum_{j=1}^{k} \frac{c_j}{1+r} \pi_j \leq R_0 + \sum_{j=1}^{k} \frac{R_j}{1+r} \pi_j .$$

The consumption at time 0, plus the value, discounted by the risk-free return, of the expectation with respect to π of consumption at time 1, is less than or equal to the revenue at time 0, plus the discounted expectation of the revenue at time 1. This formulation of the constraint will be used in continuous time, in Chaps. 4 and 8, as it allows us to transform a path-wise constraint into a constraint on an expected value.

The Incomplete Market Case

We suppose that there is a riskless asset. We write $\mathcal{P}$ for the set of probability measures π satisfying $V^T\pi = (1+r)S$. If $c \in B(S)$ and $\pi \in \mathcal{P}$, we have

$$c_0 + \sum_{j=1}^{k} \frac{c_j}{1+r}\,\pi_j \;\leq\; R_0 + \sum_{j=1}^{k} \frac{R_j}{1+r}\,\pi_j\;.$$

We use the notation
$$V(\pi) = \max u(c)$$
where the maximum is taken over the c that satisfy the constraint

$$c_0 + \sum_{j=1}^{k} \frac{c_j}{1+r}\,\pi_j \;\leq\; R_0 + \sum_{j=1}^{k} \frac{R_j}{1+r}\,\pi_j\;.$$

Thus we obtain
$$u(c^*) \leq \min_{\pi \in \mathcal{P}} V(\pi)\;.$$

As the corresponding necessary and sufficient first order conditions are satisfied, it follows from (1.10) and (1.11) that $u(c^*) = V(\beta(1+r))$ where β is defined as in (1.10). Therefore we have:

$$u(c^*) = \min_{\pi \in \mathcal{P}} V(\pi)\;.$$

We refer to such a π as a "minimax" probability measure.

1.3.2 An Equilibrium Model with a Representative Agent

We take as given the endowments $(R_0, \ldots, R_k)$ of the agent, and the asset prices S. A pair $((R_0, \ldots, R_k), S)$ is an equilibrium if the optimal solution to the agent's consumption–portfolio problem is $((R_0, \ldots, R_k), 0_{d+1})$. In other words, at price S, the agent does not carry out any transactions. Let z be a contingent claim, and let $S(z)$ be its price. We say that the claim is valued at equilibrium if, when it is introduced into the financial markets in equilibrium, the optimal demand θ_z for the claim is zero. In other words, writing R for the agent's random endowments at time 1 and C for his consumption vector at time 1, the optimal solution to the problem

$$\max u(c_0, c) \quad \text{under the constraints}$$
$$c_0 + \theta \cdot S + \theta_z S(z) \leq R_0$$
$$C \leq R + V\theta + \theta_z z$$

is given by $(R_0, \ldots, R_k)$ and by the associated portfolio $(0_{d+1}, 0_z)$.

Proposition 1.3.4. *If $((R_0, \ldots, R_k), S)$ is an equilibrium, then the interest rate and the asset prices are given by:*

$$1 + r = \frac{\partial u / \partial c_0 (R_0, \ldots, R_k)}{\sum\limits_{j=1}^{k} \partial u / \partial c_j (R_0, \ldots, R_k)}$$

and

$$S^i = \sum_{j=1}^{k} \frac{\partial u / \partial c_j (R_0, \ldots, R_k)}{\partial u / \partial c_0 (R_0, \ldots, R_k)} v_j^i \quad \text{for all } i = \{1, \ldots, d\} \ .$$

The equilibrium price of a contingent claim $z \in \mathbb{R}^k$ is:

$$S(z) = \sum_{j=1}^{k} \frac{\partial u / \partial c_j (R_0, \ldots, R_k)}{\partial u / \partial c_0 (R_0, \ldots, R_k)} z_j \ .$$

Proof. The first part of the proposition follows from (1.10) and (1.11) with $(c^*) = (R_0, \ldots, R_k)$. To prove the second part, we suppose that the agent's consumption–portfolio problem

$$\max u(c_0, c) \text{ under the constraints}$$
$$c_0 + \theta \cdot S + \theta_z S(z) \leq R_0$$
$$C \leq R + V\theta + \theta_z z$$

has for optimal solution $((R_0, \ldots, R_k), 0_{d+1}, 0_z)$. According to the Kuhn–Tucker theorem, there exists $\tilde{\lambda} \in \mathbb{R}_+^{k+1}$ such that

i) $\frac{\partial u}{\partial c_0}(R_0, \ldots, R_k) - \tilde{\lambda}_0 = 0$

ii) $\frac{\partial u}{\partial c_j}(R_0, \ldots, R_k) - \tilde{\lambda}_j = 0 , \qquad j \in \{1, \ldots, k\}$

iii) $\tilde{\lambda}_0 S^i = \sum_{j=1}^{k} \tilde{\lambda}_j v_j^i , \qquad i \in \{0, \ldots, d\}$

iv) $\tilde{\lambda}_0 S(z) = \sum_{j=1}^{k} \tilde{\lambda}_j z_j.$

The equilibrium price of a contingent asset follows trivially from these formulae. $\qquad\square$

Exercise 1.3.5. Consider an economy with two dates 0 and 1. At time 1, there are two states of the world. At time 0, an agent holding one euro can buy a portfolio made up of two assets with respective payment vectors $[1, 1]$ and $[2, 0]$. Assume moreover that he consumes c_0. At time 1, in addition to the payment vector of his portfolio, the agent receives $[1, 2]$ and consumes (c_1, c_2). Suppose the agent has utility function $u(c_0, c_1, c_2) = \log(c_0) + \frac{1}{2}(\log(c_1) + \log(c_2))$. Suppose that the agent's optimal strategy is to buy nothing. What are the assets' equilibrium prices? What is the risk-neutral probability measure?

1.3.3 The Von Neumann–Morgenstern Model, Risk Aversion

First of all, we present the theory for decisions taken over one period. In the interests of simplicity, we assume here that there is only a single consumption good.

Let $\mathcal{P}$ be the set of probability measures on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. In particular, if there are only a finite number of states, if state j occurs with probability μ_j, and if consumption C at time 1 is a random variable taking values c_j, then the probability law μ_C of C with $\mu_C = \sum_{j=1}^{k} \mu_j \delta_{c_j}$ is an element of $\mathcal{P}$.

We make the assumption that only the consequences of random events (that is possible cash flows and their probabilities) are taken into account. This assumption comes down to supposing that the agent's preferences are not expressed on the positive or zero random variables, but directly on $\mathcal{P}$. We use the notation $\succeq$ for the preorder of the agent's preferences, which is assumed to be complete. We say that $u : \mathcal{P} \to \mathbb{R}$ is a utility function that represents the preorder of preferences if $u(\mu) \geq u(\mu')$ is equivalent to $\mu \succeq \mu'$.

We say that the utility is Von Neumann–Morgenstern if there exists $v : \mathbb{R}_+ \to \mathbb{R}$ such that

$$u(\mu) = \int_0^\infty v(x) \mathrm{d}\mu(x) .$$

In the particular case $\mu_C = \sum_{j=1}^{k} \mu_j \delta_{c_j}$, the VNM utility is written

$$u(\mu_C) = \sum_{j=1}^{k} \mu_j v(c_j) ,$$

so that our criterion is to maximize the expectation of the consumption's utility.

We do not discuss here the abundant literature that establishes axioms on the preorder of preferences on $\mathcal{P}$ in such a way that it admits a VNM representation.

We say that the agent is *risk averse* if

$$v(E(C)) \geq E(v(C)), \text{ for all } C .$$

Thus an investor prefers a future consumption $E(C)$ with certainty, to a consumption c_1 with probability μ_1, a consumption c_2 with probability $\mu_2, \ldots,$ a consumption c_k with probability μ_k.

If the agent has a preorder on all finite probabilities, the presence of risk aversion is equivalent to the concavity of v. Indeed, if v is concave, then according to Jensen's inequality[4], we have $v(E(C)) \geq E(v(C))$, for all C.

---

[4] See for example Chung [58].

Conversely, we suppose that $v(E(C)) \geq E(v(C))$ for all C. Let $(x,y) \in \mathbb{R}_+^2$. We consider the random variable C worth x with probability α and y with probability $1 - \alpha$. As

$$v(E(C)) = v(\alpha x + (1 - \alpha)y) \geq E(v(C)) = \alpha v(x) + (1 - \alpha)v(y) \;,$$

by letting α, x and y vary, we obtain the concavity of v.

We say that an investor is *risk-neutral*, if v is an affine function. Then

$$v(E(C)) = E(v(C)), \text{ for all } C \;.$$

When an agent is risk averse, we define the *risk premium* $\rho(C)$ linked to the random consumption C: it is the amount the investor is prepared to give up in order to obtain, with certainty, a consumption level equal to $E(C)$. As v is a continuous, strictly increasing and strictly concave function, for all C there exists $\rho(C) \geq 0$ such that

$$v(E(C) - \rho(C)) \;=\; E\left(v(C)\right) \;. \tag{1.14}$$

The amount $E(C) - \rho(C)$ is called the *certainty equivalent* of C, and ρ is called the risk premium. When the investor is risk-neutral, $E[v(C)] = v[E(C)]$, so that $\rho(C) \;=\; 0$. We now assume that there are a finite number of states, that consumption C at time 1 is a random variable taking values c_j with probability μ_j, and that v is of class C^2. Using Taylor's expansion, on the condition that the values c_j taken by the consumption C are close enough to $E(C)$, we get

$$v(c_j) \;\simeq\; v[E(C)] + [c_j - E(C)]v'[E(C)] + \frac{[c_j - E(C)]^2}{2}\, v''[E(C)] \;.$$

Taking expectations on both sides,

$$E[v(C)] \;=\; \sum_{j=1}^{k} \mu_j v(c_j) \;\simeq\; v[E(C)] + v''[E(C)] \frac{\text{Var } C}{2} \;.$$

Expanding the first term of (1.13), using Taylor's expansion once again, we get

$$v[E(C) - \rho(C)] \;\simeq\; v[E(C)] - \rho(C)v'[E(C)]$$

and hence we can evaluate $\rho(C)$:

$$\rho(C) \;\simeq\; -\frac{v''[E(C)]}{2v'[E(C)]} \text{ Var } C \;.$$

The coefficient $I_a(v,x) = -\dfrac{v''(x)}{v'(x}$ is called the *absolute risk aversion co-efficient of v*. Thus the certainty equivalent of C is approximately equal to

$E(C) - \dfrac{I_a(v, E(C))}{2} \operatorname{Var} C$, which as a first approximation justifies the choice of a mean–variance criterion.

Exercise 1.3.6. We denote by $\mathcal{N}(\mu, \sigma^2)$ the normal distribution with mean μ and variance σ^2. Let $C \overset{\text{law}}{=} \mathcal{N}(\mu, \sigma^2)$ and $v(c) = -\mathrm{e}^{-\beta c}$, $\beta > 0$. Show that $I_a(v, x) = \beta$ for any x and that $E(C) - \rho(C) = \mu - \frac{\beta}{2}\sigma^2$. (Here we can use that fact that $C \overset{\text{law}}{=} \mu + \sigma Y$, where $Y \overset{\text{law}}{=} \mathcal{N}(0, 1)$). In this particular case, the certainty equivalent of C is exactly equal to $E(C) - \dfrac{\alpha}{2}\operatorname{Var} C$.

Exercise 1.3.7. Calculate the absolute risk aversion index in the following cases: $v(c) = \frac{c^\gamma}{\gamma}$ with $0 < \gamma < 1, \quad v(c) = \ln c$.

1.3.4 Optimal Choice in the VNM Model

We return to the two date model considered in Sect. 1.3.1, and assume that at time 1, state j occurs with probability $\mu = (\mu_j)_{j=1}^k$. We suppose here that there is a riskless asset, and that the market is complete. The investor has preferences on $\mathbb{R} \times \mathcal{P}$. Let us look at the special case in which the preferences can be represented by a utility function that is "additively separable" with respect to time:

$$u(c_0, C) \; = \; v_0(c_0) + \alpha \, E(v(C)) \; = \; v_0(c_0) + \alpha \sum_{j=1}^k \mu_j \, v(c_j)$$

$$0 < \alpha < 1$$

where α is a discount factor, and where v^0 and v are strictly concave, strictly increasing C^2 functions satisfying

$$\lim_{x \to 0} v_0'(x) = \infty \;, \qquad \lim_{x \to 0} v'(x) = \infty \;,$$

$$\lim_{x \to \infty} v_0'(x) = 0 \;, \qquad \lim_{x \to \infty} v'(x) = 0 \;.$$

Let $I_0 :]0, \infty[\to]0, \infty[$ (respectively $I :]0, \infty[\to]0, \infty[$) be the inverse function of $v^{0\prime}$ (respectively of v'). The functions I_0 and I are continuous and strictly decreasing. Denote the riskless rate by r.

In this special case, formulae (1.11) and (1.12) become:

$$1 + r = \alpha \frac{\sum_{j=1}^k \mu_j v'(c_j^*)}{v_0'(c_0^*)} = \alpha \frac{E(v'(C^*))}{v_0'(c_0^*)} \tag{1.15}$$

$$S^i \; = \; \alpha \frac{\sum_{j=1}^k \mu_j v'(c_j^*) v_j^i}{v_0'(c_0^*)} = \frac{1}{1+r} E(V^i v'(C^*)) \tag{1.16}$$

where V^i and C^* are random variables taking the values v_j^i and c_j^* respectively.

Let us show how to obtain the optimal solution in explicit form. Indeed, the investor solves the following problem $\mathcal{P}$:

$$\max v_0(c_0) + \alpha \sum_{j=1}^{k} \mu_j\, v(c_j) \text{ under the constraint}$$

$$c_0 + \sum_{j=1}^{k} \frac{c_j}{1+r}\, \pi_j \;\leq\; R_0 + \sum_{j=1}^{k} \frac{R_j}{1+r}\, \pi_j\;.$$

Let λ be the Lagrange multiplier associated with the constraint. We have

$$v_0'(c_0^*) = \lambda$$
$$\alpha\mu_j v'(c_j^*) = \lambda\tfrac{\pi_j}{1+r}, \quad \forall\, j = 1,\ldots,k\;,$$

$$(1.17)$$

and hence

$$c_0^* = I_0(\lambda)$$
$$c_j^* = I\left(\frac{\pi_j \lambda}{\mu_j(1+r)\alpha}\right), \quad \forall\, j = 1,\ldots,k\;.$$

$$(1.18)$$

The Lagrange multiplier λ is determined by the budget constraint, and satisfies the following equation:

$$I_0(\lambda) + \frac{1}{1+r}\sum_{j=1}^{k} \pi_j I\left(\frac{\lambda\pi_j}{\mu_j(1+r)\alpha}\right) = R_0 + \frac{1}{1+r}\sum_{j=1}^{k} \pi_j R_j\;. \qquad (1.19)$$

As the function $x \to I_0(x) + \frac{1}{1+r}\sum_{j=1}^{k} \pi_j I\left(\frac{\pi_j x}{\mu_j(1+r)\alpha}\right)$ is a decreasing function from $]0,\infty[$ into itself, equation (1.16) has a unique solution. Once the Lagrange multiplier has been determined, we can deduce the optimal consumption from (1.15), and then finally obtain the optimal portfolio using the relation $V\theta^* = C^* - R$.

Let us show that under the assumption that the agent is in equilibrium, we can give an estimate of the risk-neutral probability.

Replacing c_0^* by R_0 and C^* by R, it follows from (1.14) that

$$\lambda = v_0'(R_0) = \alpha(1+r)E(v'(R))$$

and that

$$\pi_j = \mu_j \frac{v'(R_j)}{E(v'(R))}\;. \qquad (1.20)$$

This expression for the risk-neutral probability does not depend on future consumption. Using Taylor's expansion, we get:

$$v'(R_j) \simeq v'[E(R)] + [R_j - E(R)]v''[E(R)] .$$

Taking expectations,
$$E[v'(R)] \simeq v'[E(R)] .$$

Hence

$$\frac{\pi_j}{\mu_j} \simeq 1 + \frac{[R_j - E(R)]v''[E(R)]}{v'[E(R)]} = 1 + I_a(v, E(R))(E(R) - R_j) .$$

The greater the agent's index of absolute aversion to risk for $E(R)$ and the greater the difference between the average value of his resources and his resources in a given state j, the greater is the risk-neutral probability of state j occurring.

If the investor had a neutral attitude to risk ($v' = $ cst), he would be prepared to pay $\dfrac{\alpha\mu_j}{v_0'(R_0)}$ at time 0 in order to receive 1 euro tomorrow in the state of the world j. If he is risk averse, he is prepared to pay $\dfrac{\alpha\mu_j v'(R_j)}{v_0'(R_0)}$ today so as to receive 1 euro tomorrow in state of the world j.

To summarize, we have used two different approaches:

- the assumption of no arbitrage opportunities enabled us to construct a probability measure under which we are neutral with respect to risk,

- the introduction of a utility function and of exogenous (or subjective) probabilities led us firstly to define the concept of risk aversion, secondly to obtain valuation formula (1.16), and finally to exhibit a risk-neutral probability measure.

We remark on the fact that these two risk-neutral probability measures are equivalent.

Exercise 1.3.8.

1. We assume that the market is complete and that $v_0(c) = v(c) = \log(c)$. Calculate I_0, I and the optimal solution. Carry out the corresponding calculations when $v_0(c) = 0$ and $v(c) = \log(c)$, and similarly obtain I_0, I and the optimal solution for $v_0(c) = v(c) = c^\alpha$, $0 < \alpha < 1$.

2. We consider an economy with two dates, 0 and 1. At time 1, there are two states of the world. At time 0, an agent does not own anything and can buy for a price $[1, 1]$, a portfolio of two assets whose respective payment vectors are $[1, -1]$ and $[2, -2]$. At time 1, in addition to the payment vector of

his portfolio, the agent receives $[1, 1]$ and consumes (c_1, c_2). Suppose that the agent has a VNM utility function, that he attributes the probabilities $(\frac{1}{2}, \frac{1}{2})$ to the two states of the world, and that his utility index is $u(c) = \log c$. Show that the agent's consumption–portfolio problem has a solution, and that nevertheless, the market admits an arbitrage. Comment on these results.

Exercise 1.3.9. We consider an economy with two dates 0 and 1. There are three states of the world at time 1. At time 0, an agent does not own anything, and he can buy a portfolio of three assets that have respective payment vectors $[1, 1, 1]$, $[3, 2, 1]$ and $[1, 2, 6]$, and respective prices $S^1 = 1$, $S^2 = 2$ and $S^3 = 3$. He must not run into debt. The agent does not consume at time 0. At time 1, in addition to the payment vector of his portfolio, the agents receives $[1, 2, 1]$ in the different states, and consumes (c_1, c_2, c_3).

1. Calculate the state prices and the interest rate. Show that the market is complete. Calculate the risk-neutral probability. Show that the set of consumptions that the agent can achieve at time 1 is given by

$$\{c \in \mathbb{R}^3 \mid c_1 + c_2 + c_3 \le 4\}.$$

2. We assume that the agent attributes the probabilities $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ to the different states of the world, and that he has a VNM utility function of index $u(x) = \log(x)$. Calculate his optimal consumption and portfolio. How would the results change if the agent attributed the probabilities $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ to the different states of the world?

3. Suppose that the agent can only buy a portfolio that contains the two first assets. Calculate the interest rate, and characterize the set of risk-neutral probabilities. Find the set of its extrema. Show that the set of consumptions that can be attributed to the agent at time 1 is

$$\{c \in \mathbb{R}^3 \mid c_2 \le 2, \ \frac{1}{2}(c_1 + c_3) \le 1\}.$$

Calculate the optimal consumption and portfolio when the agent attributes probabilities $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ to the states of the world, and has a VNM utility function with index $u(x) = \log(x)$.

4. We suppose that the agent can, without running into debt, purchase a portfolio made up of the three assets, and that in addition he can buy a positive amount of asset 2.
Characterize the set of risk-neutral probabilities. Find the set of its extrema. Calculate the agent's optimal consumption and portfolio when he affects the probabilities $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ to the states of the world, and has a VNM utility function with index $u(x) = \log(x)$.

1.3.5 Equilibrium Models with Complete Financial Markets

The Representative Agent

We now study a simple model in order to illustrate the effect of introducing financial markets into an economy. This model will be further developed in Chap. 6.

Consider an exchange economy with a single consumption good and m economic agents. We suppose that there are $(d + 1)$ assets, with the same characteristics as in the previous sections. We assume the market to be complete.

Agent h has an initial endowment of e_{h0} units of the good at time 0, and knows that he will receive e_{hj} units of the good at time 1 in state of the world j. To modify his future resources, he can, at time 0, buy a portfolio of securities $\theta_h = (\theta_h^0, \dots, \theta_h^d)$ on condition that he does not run into debt.

Given a price S for the assets, we define the agent's budget set as the set of consumption plans which he can carry out with his initial wealth and future income:

$$B_h(S) := \{c \in \mathbb{R}_+^{k+1} \mid \exists \theta \in \mathbb{R}^{d+1},$$
$$e_{h0} \geq c_0 + \theta \cdot S; \ e_{hj} \geq c_j - (V\theta)_j, j \in \{1, \dots, k\}\}.$$

As in Sect. 1.3.3, we suppose that agent h has preferences that are represented by a VNM utility function of the form

$$u_h(c_0, C) = v_{h0}(c_0) + \alpha \sum_{j=1}^{k} \mu_j v_{h1}(c_j) .$$

We suppose here that all the agents have the same discount factor α.

Definition 1.3.10. *The collection $\{\overline{S}, (\overline{c}_h, \overline{\theta}_h); h = 1, \dots, m\}$ is an equilibrium of the economy with financial markets if, given $\overline{S}$*

1. $\overline{c}_h$ maximizes $u_h(c_{h0}, C_h)$ under the constraint $c_h = (c_{h0}, C_h) \in B_h(\overline{S})$,

2. $\sum_{h=1}^{m} \overline{c}_{hj} = \sum_{h=1}^{m} e_{hj} := e_j, \ j \in \{1, \dots, k\}$,

3. $\sum_{h=1}^{m} \overline{\theta}_h = 0$.

In other words, the market in the good clears (equality 2) and the security market also clears (equality 3).

Remark 1.3.11. If v_h^0 and v_h are strictly increasing, and if V is injective, then equality 3 is implied by 1 and 2. Indeed, as the utility functions are strictly increasing, at equilibrium the constraints are binding. Therefore we have $e_{hj} = \overline{c}_{hj} - (V\overline{\theta}_h)_j$ for all h and $j \in \{1, \dots, k\}$. As $\sum_{h=1}^{m} e_{hj} = \sum_{h=1}^{m} \overline{c}_{hj}$ for $j \in \{1, \dots, k\}$, we have $V\left(\sum_{h=1}^{m} \overline{\theta}_h\right) = 0$, which implies $\sum_{h=1}^{m} \overline{\theta}_h = 0$.

We suppose now that an equilibrium exists. We can use the first order necessary and sufficient conditions of the precious section. Hence for all h,

$$S^i = \alpha \sum_{j=1}^{k} \mu_j \frac{v_h'(\overline{c}_{hj})}{v_{h0}'(\overline{c}_{h0})} v_j^i = \sum_{j=1}^{k} \frac{1}{1+r} \mu_j \frac{v_h'(\overline{c}_{hj})}{E(v_h'(\overline{C}_h))} v_j^i , \qquad (1.19)$$

$$\frac{1}{1+r} = \alpha \sum_{j=1}^{k} \mu_j \frac{v_h'(\overline{c}_{hj})}{v_{h0}'(\overline{c}_{h0})} . \qquad (1.20)$$

Under the assumption of complete markets, the equation $S = V^T \beta$ has a unique solution. Under this assumption, the ratios

$$\frac{v_h'(\overline{c}_{hj})}{v_{h0}'(\overline{c}_{h0})} \quad \text{are therefore independent of } h.$$

Let us then consider a fictitious agent, the "representative agent", whose utility is

$$u(c_0, C) := v_0(c_0) + \alpha \sum_{j=1}^{k} \mu_j v(c_j) ,$$

where

$$v_0(c) := \max \left\{ \sum_{h=1}^{m} \frac{v_{h0}(c_h)}{v_{h0}'(\overline{c}_{h0})}; \quad \sum_{h=1}^{m} c_h = c \right\}$$

$$v(c) := \max \left\{ \sum_{h=1}^{m} \frac{v_h(c_h)}{v_{h0}'(\overline{c}_{h0})}; \quad \sum_{h=1}^{m} c_h = c \right\} .$$

Using the first order necessary and sufficient conditions of these new optimization problems, we check that

$$u(e_0, e) = \sum_{h=1}^{m} \frac{v_{0h}(\overline{c}_{h0})}{v_{h0}'(\overline{c}_{h0})} + \alpha \sum_{j=1}^{k} \sum_{h=1}^{m} \frac{v_h(\overline{c}_{hj})}{v_{h0}'(\overline{c}_{h0})} \mu_j ,$$

where e is a random variable taking the value e_j with probability μ_j. Using the first order conditions and the implicit function theorem, we show that v_0 and v are differentiable, that $v_0'(e_0) = 1$ and that

$$v'(e_j) = \frac{v_h'(\overline{c}_{hj})}{v_{h0}'(\overline{c}_{h0})} \quad j = 1, \ldots, k .$$

Hence, (1.20) can be rewritten as

$$\frac{1}{1+r} = \alpha E(v'(e)) \qquad (1.21)$$

$$S^i = \frac{1}{1+r} \sum_{j=1}^{k} \mu_j \frac{v'(e_j)}{E(v'(e))} v_j^i = \frac{1}{1+r} \left[E(V^i) + \text{Cov} \left(\frac{v'(e)}{E(v'(e))}, V^i \right) \right]$$

(1.22)

where V^i is a random variable taking value v_j^i.

The formula above plays a very important role in the financial literature, as it shows that when there is a equilibrium, the price of an asset is only a function of aggregate endowment (and not of each individual's endowment). In the next section, we will look at the relationship more closely.

Exercise 1.3.12. Consider an economy with two dates, two agents, two states of the world, and one good in each state. Suppose that the agents have utility functions $v_{0h}(c) = v_h(c) = \log(c)$, $h = 1, 2$ and assign probabilities $\frac{1}{2}$ to the states of the world. Assume that the agents have endowments $e_{01} = 1$, $e_1 = (1, 3)$ and $e_{02} = 2$, $e_2 = (3, 1)$. Assume that the two assets are traded at time 0: the riskless asset, and an asset that pays 1 in state 1 and 0 in state 2. Find the equilibrium of this economy.

The Capital Asset Pricing Model (CAPM) Formula

This model will be developed in greater generality in Chap. 6.

We suppose that the agents have quadratic utility functions (i.e., $v_h'(c) = -a_h c + b_h$ with $a_h > 0$ for all h), and that at equilibrium the agents have strictly positive consumption. In this particular case, we can easily check that v' is linear and decreasing, that is

$$v'(c) = -ac + b \quad (\text{with } a > 0) .$$

Equation (1.22) then becomes

$$S^i = \frac{1}{1+r} \left[E(V^i) - \frac{a}{E(v'(e))} \text{Cov}(e, V^i) \right] . \qquad (1.23)$$

Formula (1.23) is called the CAPM (Capital Asset Pricing Model) formula. As long as $E(v'(e)) > 0$, the price of asset i is therefore greater than the discounted expectation of returns, if it is negatively correlated with e (i.e., $\text{Cov}(e, V^i) \le 0$): the asset provides a form of insurance.

If we introduce $\rho^i = \frac{V^i}{S^i}$, the return on asset i, and M, such that $e = VM$ (M is called the market portfolio), we can express (1.23) in the form

$$E(\rho^i) - (1 + r) = \frac{a}{E(v'(e))} \text{Cov}(\rho^i, e) = \frac{aS \cdot M}{E(v'(e))} \text{Cov}(\rho_i, \rho_M) ,$$

setting $\rho_M = \frac{e}{S \cdot M}$ (ρ_M is the return on the market portfolio). We then get, in particular:

$$E(\rho_M) - (1+r) = \frac{aS \cdot M}{E(v'(e))} \text{Var } \rho_M .$$

From this we deduce

$$E(\rho^i) - (1+r) = \frac{\text{Cov }(\rho^i, \rho_M)}{\text{Var } \rho_M} \{E(\rho_M) - (1+r)\} . \qquad (1.24)$$

This formula, which links the excess return on an asset to the return on the market portfolio, is called the beta formula, where the β^i coefficient is given by $\frac{\text{Cov}(\rho^i, \rho_M)}{\text{Var } \rho_M}$. We note that β^i is the coefficient of the regression of ρ^i on ρ_M. In valuation models for financial assets (or in the CAPM: Capital Asset Pricing Model), it is interpreted as a sensitivity factor to the risk of asset i. We find that the risk premium for asset i, that is $E(\rho^i) - (1+r)$, is a linear function of its β.

An Approximate CAPM Formula

More generally, taking any utility function for the representative agent, let us suppose that the e_j are close to $E(e)$. We then obtain an approximation of the CAPM formula. Indeed,

$$\frac{v'(e_j)}{E(v'(e))} \simeq 1 + \alpha[E(e) - e_j] ,$$

where α is the representative agent's index of absolute aversion to the risk in $E(e)$. Formula (1.22) then becomes

$$S^i \simeq \frac{E(V^i)}{1+r} - \frac{\alpha}{1+r} \text{Cov }(e, V^i) . \qquad (1.25)$$

Notes

The financial literature in discrete time is extensive, and it would be quite impossible to give a detailed bibliography here. We restrict ourselves to a few books, which provide the basics: Elliott and Van Der Hoek [150], (2004),Huang and Litzenberger [197], (1988), Pliska [301], (1997), Mel'nikov [271], (1999), Shreve [337], (2004), the first part of Shiryaev [336], (1999), Le Roy and Werner [252], (2001), Cvitanic and Zapatero [79], (2002) and the first part of Föllmer and Schied [162], (2004).

The initial formulation and use of the NAO assumption are due to Ross [316, 317], (1976, 1978). Varian [359], (1988) gives a summary of this approach

and Cochrane [61] develops its applications to asset pricing. For the probabilistic aspects, see Bingham and Kiesel [33] (1998) and Björk [34] (1998).

For the axiomatic approach to the Von Neumann–Morgenstern utility, the reader can consult the books of Huang and Litzenberger [197], (1988), Kreps [244], (1990), and Le Roy and Werner [252], (2001), and Föllmer and Schied [162], (2004).

The problem of choosing an optimal consumption and portfolio in incomplete markets, or in the presence of portfolio constraints, was originally studied by He and Pearson [185], (1991). Different solution methods are presented in Pliska [301], (1997) and Mel'nikov [271], (1999).

The options literature goes back to Merton [273, 274], (1973). That too is vast. We have only given the basic definitions here. The reader is referred to Cox–Rubinstein [72], (1985). Wilmott's books [369, 370], (1998, 2001) provide a good introduction to the problem of valuation and hedging. A more detailed study of our own and other references will be given in the chapter on exotic options.

We will study the equilibria of financial markets more thoroughly in Chap. 6.

For optimization in finite dimensions and duality properties, the reader can refer to Rockafellar [312], (1970), Luenberger [260],(1969), Hiriart-Urruty and Lemaréchal [192], (1996), and Florenzano and Le Van [159], (2000).

ANNEX 1

Optimization under Constraints, the Kuhn–Tucker Theorem with Linear Constraints

Let C be an open convex set in $\mathbb{R}^n$.

We consider the following problem denoted $P_{\alpha\beta}$ and formulated for $\alpha = (\alpha_1, \ldots, \alpha_p) \in \mathbb{R}^p$ and $\beta = (\beta_1, \ldots, \beta_q) \in \mathbb{R}^q$, by:

$$\begin{aligned}
&\max f(x), \quad \text{under the constraints} \\
&f_i(x) \le \alpha_i, \quad \forall i = 1, \ldots, p, \\
&g_j(x) = \beta_j, \quad \forall j = 1, \ldots, q, \\
&x \in C
\end{aligned}$$

where the function $f : C \to \mathbb{R}$ is concave and differentiable, and where the functions $f_i : C \to \mathbb{R}, \quad i = 1, \ldots, p$ and $g_j \quad j = 1, \ldots, q$ are affine. We call f the **objective function**. We write K for the **admissible set**

$$K = \{x \in C \mid f_i(x) \le \alpha_i, \quad \forall i = 1, \ldots, p, \quad g_j(x) = \beta_j, \quad \forall j = 1 \ldots q\} \, .$$

Theorem *Let $\bar{x} \in K$. Then $\bar{x}$ is a solution to $P_{\alpha\beta}$ if and only if there exists $(\bar{\lambda}, \bar{\mu}) \in \mathbb{R}_+^p \times \mathbb{R}^q$ such that*

1. $\nabla f(\bar{x}) = \sum_{i=1}^p \bar{\lambda}_i \nabla f_i(\bar{x}) + \sum_{j=1}^q \bar{\mu}_j \nabla g_j(\bar{x}),$
2. $\bar{\lambda}_i(f_i(\bar{x}) - \alpha_i) = 0, \qquad \forall i = 1, \ldots, p.$

We call $(\bar{\lambda}, \bar{\mu})$ **the Lagrange multipliers** or the **Kuhn–Tucker multipliers**

2

Dynamic Models in Discrete Time

To make it easier to approach to continuous-time models, we present here a dynamic model in discrete time and with a finite horizon. This allows us specify such concepts as self-financing, arbitrage and complete markets, and to show how martingales, though not part of the initial data of the model, can be used to give the problem a pleasing form.

We consider a model in discrete time, with a finite horizon. There are $d+1$ assets, whose prices at each time n are represented by a random variable. We define the concept of a self-financing strategy, and of an arbitrage opportunity.

Under the assumption that the underlying probability space is finite, we provide two proofs that the assumption of NAO is equivalent to the existence of a probability measure under which discounted prices are martingales. The first uses a separation theorem, and the second, the results obtained in Chap. 1. Indeed, we establish that to any informational structure, we can associate a tree, and that the condition of "no arbitrage opportunities" defined previously is equivalent to a concept of no arbitrage opportunities at each node of the tree. This enables us to use the results obtained for a one period model, and to give an alternative proof of the existence of a probability measure under which discounted prices are martingales.

Once we have defined complete markets, we show that under the assumption of NAO, this property is equivalent to the uniqueness of the martingale measure, which is equivalent to P, and under which discounted prices are martingales.

Next we tackle the problem of valuation. First we focus on replicable variables, which are such that they can be attained with a self-financing strategy, whose value we will determine. In the case of a complete market, we show that the value at time 0 of any given strategy, is equal to the expectation of its discounted payoff under the unique martingale measure.

As an example, we describe the binomial model of Cox, Ross and Rubinstein, and by taking its limit, we then obtain the Black–Scholes model.

Next, under a finite horizon, we present two models of optimal portfolio choice. In the first, the criterion is the maximization of the utility of final wealth, and in the second, it is the maximization of the utility of the consumption stream.

Finally, in the last section, we describe the problems encountered under an infinite horizon.

The annex recalls the definitions and main properties of conditional expectation and martingales.

2.1 A Model with a Finite Horizon

Let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with a filtration $(\mathcal{F}_n)_{n=0}^{N}$, that is to say, with an increasing family of sub-σ-fields $\mathcal{F} : \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_N = \mathcal{F}$. The paths $\omega \in \Omega$ are a generalization of the concept of different states of the world.

The σ-field $\mathcal{F}_0$ is the trivial σ-field ($\mathcal{F}_0 = \{\phi, \Omega\}$). The σ-field $\mathcal{F}_n$ represents the information that is known at time n. The increasing nature of the family of σ-fields translates the fact that there is no "loss" of information.

A sequence of random variables $(S_n, n \leq N)$ is $\mathcal{F}_n$-adapted if S_n is $\mathcal{F}_n$-measurable for all $n \leq N$. In other words, the information known at time n includes knowledge of the value taken by the random variable S_n. The increasing nature of the filtration implies that any $\mathcal{F}_n$-measurable variable is also $\mathcal{F}_k$-measurable for all $k > n$. A $\mathcal{F}_0$-measurable variable is (a.s.) equal to a constant.

When the space $(\Omega, \mathcal{F})$ is endowed with two probability measures P and Q, we specify that a property is true for $(\Omega, \mathcal{F})$ equipped with P (respectively with Q), by saying that the property is true "under P" (respectively under Q).

The financial market is made up of $d+1$ assets, whose prices at time n are given by a random vector $S_n = (S_n^0, S_n^1, \ldots, S_n^d)^T$, with values in $\mathbb{R}_+^{d+1}$. The asset 0 is assumed to be riskless: $S_n^0 = (1+r)^n S_0^0$ where r is the interest rate, which is assumed to be constant for the sake of simplicity. We take $S_0^0 = 1$.

We suppose that the vector S_n is $\mathcal{F}_n$-measurable.

Definition 2.1.1. *A portfolio strategy is a family $\theta = (\theta_n)_{n=1}^{N}$ of random vectors $\theta_n = (\theta_n^0, \theta_n^1, \ldots, \theta_n^d)$ such that*

$$\forall\, 1 \leq n \leq N, \quad \forall i \geq 0, \quad \theta_n^i \quad \text{is } \mathcal{F}_{n-1}\text{-measurable} .$$

The vector θ_n is the portfolio at time n: θ_n^i represents the number of shares of the asset i that are held at time n. We do not place any restrictions on the sign of θ_n^i, which means that short sales and borrowing are allowed. The condition of measurability reflects the fact that θ is "predictable[1]". This means that the investor chooses his portfolio of assets θ_n^i "just before n"; i.e., with only knowledge of the information described by $\mathcal{F}_{n-1}$. In particular, he does not yet know the price S_n.

We denote by $V_n(\theta) = \theta_n \cdot S_n$ the scalar product $\sum_{i=0}^{d} \theta_n^i S_n^i$, which represents the value of the portfolio at time $n > 0$. The random variable $V_n(\theta)$ is $\mathcal{F}_n$-measurable. For any sequence of r.v. (X_n), we use the notation $\Delta X_n = X_n - X_{n-1}$.

Definition 2.1.2. *A portfolio strategy θ is self-financing if*

$$\theta_n \cdot S_n \;=\; \theta_{n+1} \cdot S_n, \quad n \in \{1, \ldots, N-1\} \,.$$

This means that no additional funds are introduced, no money is withdrawn, and that transactions do not entail costs. The variations of the value of the portfolio are due only to the variations in the assets' prices. The amount $\theta_n \cdot S_n$ is the value of the portfolio after time n and (strictly) before time $n+1$, and $\theta_{n+1} \cdot S_n$ is the value of the portfolio after it has been readjusted, strictly before time $n+1$, and before the prices change (the predictability of θ_n). We write $V_0(\theta) = \theta_1 \cdot S_0$, which represents the initial value of the portfolio. We note that V_0 is $\mathcal{F}_0$-measurable.

We can write the self-financing condition as follows:

$$\begin{aligned}
V_n(\theta) &= \theta_{n-1} \cdot S_{n-1} + \theta_n \cdot S_n - \theta_n \cdot S_{n-1} \\
&= V_{n-1}(\theta) + \theta_n \cdot \Delta S_n, \; n \geq 1 \,.
\end{aligned} \tag{2.1}$$

2.2 Arbitrage with a Finite Horizon

2.2.1 Arbitrage Opportunities

Definition 2.2.1. *An arbitrage opportunity is a self-financing strategy θ such that*

$$\begin{aligned}
&\text{(i)} \quad P(V_0(\theta) = 0) &= 1 \\
&\text{(ii)} \quad P(V_N(\theta) \geq 0) &= 1\,; \quad P(V_N(\theta) > 0) > 0 \,.
\end{aligned}$$

The investor has no capital initially, and at time N he ends up with non-negative wealth, which is positive on a set of positive measure (and so $E(V_N(\theta)) > 0$).

---

[1] We will come back to this concept in Chap. 3.

Definition 2.2.2. *Two probability measures P and Q defined on the same probability space $(\Omega, \mathcal{F})$ are equivalent if, for all $A \in \mathcal{F}$,*

$$P(A) = 0 \Longleftrightarrow Q(A) = 0 .$$

It is easy to see that the set of arbitrage opportunities is identical for two equivalent probability measures.

2.2.2 Arbitrage and Martingales

Let us now show that the assumption of NAO is equivalent to the existence of a probability measure, which is equivalent to P, and under which discounted prices are martingales.

Notation 2.2.3. We denote by $\hat{S}$ the vector of discounted prices, i.e.,

$$\hat{S}_n^i = S_n^i / S_n^0 .$$

We have $\hat{V}_n(\theta) = V_n(\theta)/S_n^0 = \sum_{i=0}^d \theta_n^i \cdot \hat{S}_n^i = V_0(\theta) + \sum_{k=1}^n \theta_k \cdot \Delta \hat{S}_k$ for $n > 0$ and $\hat{V}_0(\theta) = V_0(\theta)$. The same arbitrage opportunities arise for S and for $\hat{S}$. A self-financing strategy satisfies

$$\theta_n \cdot \hat{S}_n \;=\; \theta_{n+1} \cdot \hat{S}_n \; ,$$

hence

$$\hat{V}_{n+1}(\theta) \;=\; \hat{V}_n(\theta) + \theta_{n+1} \cdot \Delta \hat{S}_{n+1}, \qquad n \in \{1, \ldots, N\} . \tag{2.2}$$

Proposition 2.2.4. *We suppose that there exists a martingale measure, in other words, a probability measure Q that is equivalent to P, and such that the vector $\hat{S} = (1, \hat{S}^1, \hat{S}^2, \ldots, \hat{S}^d)$ is a martingale under Q. Then $\left(\hat{V}_n(\theta), \, n \geq 0 \right)$ is a martingale under Q.*

Proof. Let Q be a probability measure that is equivalent to P and such that $\hat{S}$ is a martingale under Q. Let θ be a self-financing strategy. Let us show that $\hat{V}(\theta)$ is a martingale under Q. It is obvious that $\hat{V}_n(\theta)$ is $\mathcal{F}_n$-adapted. We denote by $E_Q(\cdot|\mathcal{F})$ the conditional expectation under Q. Using (2.2), we obtain

$$E_Q(\hat{V}_{n+1}(\theta) - \hat{V}_n(\theta) \,|\, \mathcal{F}_n) \;=\; E_Q(\theta_{n+1} \cdot \Delta \hat{S}_{n+1} \,|\, \mathcal{F}_n)$$

$$=\; E_Q(\theta_{n+1} \cdot (\hat{S}_{n+1} - \hat{S}_n) \,|\, \mathcal{F}_n) \;=\; \theta_{n+1} \cdot E_Q(\hat{S}_{n+1} - \hat{S}_n \,|\, \mathcal{F}_n) \; ,$$

since θ_{n+1} is $\mathcal{F}_n$-measurable (we use property g of the conditional expectation[2]). Hence, as $\hat{S}$ is a $\mathcal{F}_n$-martingale under Q,

$$E_Q(\hat{V}_{n+1}(\theta) - \hat{V}_n(\theta)|\mathcal{F}_n) \;=\; 0 .$$

$\square$

---

[2] See annex.

Proposition 2.2.5. *We suppose that there exists a probability measure Q that is equivalent to P and such that the vector of discounted prices $\hat{S} = (1, \hat{S}^1, \hat{S}^2, \ldots, \hat{S}^d)$ is a martingale under Q. Then, there are no arbitrage opportunities.*

Proof. Since the process $\hat{V}_n(\theta)$ is a martingale for any self-financing strategy θ, we have (using properties a and c of the conditional expectation, and the fact that $\hat{V}_0(\theta) = V_0(\theta)$):

$$E_Q(\hat{V}_N(\theta)) = E_Q(\hat{V}_0(\theta)) = E_Q(V_0(\theta)) .$$

If the strategy has zero initial value, we have $E_Q(\hat{V}_N(\theta)) = 0$, hence θ cannot be an arbitrage opportunity. □

We can establish the converse of Proposition 2.2.5.

Theorem 2.2.6. *We suppose that there are no arbitrage opportunities. Then there exists a probability measure Q that is equivalent to P and such that, under Q, the vector of discounted prices is a martingale.*

Proof. Let us assume that Ω is finite, with $P(\{\omega\}) > 0$, for all $\omega \in \Omega$; we reproduce the proof of Harrison–Pliska [178]. The proof is analogous to the one in Chap. 1, as we are working in a space of finite dimension. Another proof is provided in the next section.

Let C be the set of non-negative $\mathcal{F}$-measurable random variables such that $E(X) = 1$. This set is convex and compact.

Let Γ be the set of $\mathcal{F}$-measurable random variables X, such that there exists a self-financing strategy θ satisfying $V_0(\theta) = 0$ and $X = V_N(\theta)$. The set is a closed vector space. If there are no opportunities for arbitrage, then Γ and C are disjoint. According to Minkowski's theorem, there exists a linear functional L such that

$$L(X) = 0 , \quad X \in \Gamma ,$$
$$L(X) > 0 , \quad X \in C .$$

Hence, there exists a random variable l such that

$$\forall X \in C \qquad \sum_{\omega \in \Omega} l(\omega) X(\omega) > 0$$

$$\forall X = V_N(\theta) \in \Gamma \qquad \sum_{\omega \in \Omega} l(\omega) V_N(\theta)(\omega) = 0 .$$

It follows from the first property that, $\forall \omega \in \Omega$, $l(\omega) > 0$. Let $\Lambda = \sum_{\omega \in \Omega} l(\omega)$ and

$$Q(\{\omega\}) = \frac{l(\omega)}{\Lambda} .$$

We check that Q is equivalent to P: the only null set is the empty set.

It remains to check the martingale property. If θ^* is a predictable process taking values in $\mathbb{R}^d$, $\theta^* = (\theta_n^1, \theta_n^2, \ldots, \theta_n^d)_{n \leq N}$, we can construct θ^0, a process taking values in $\mathbb{R}$ and such that the strategy $(\theta^0, \theta^1, \ldots, \theta^d)$ is self-financing and has zero initial value. By induction, it suffices to choose θ_1^0 such that $\theta_1^0 S_0^0 = -\sum_{i=1}^d \theta_1^i S_0^i$ (zero initial value), and θ_{n+1}^0 such that

$$\theta_{n+1}^0 S_n^0 = \theta_n^0 S_n^0 + \sum_{i=1}^d S_n^i (\theta_n^i - \theta_{n+1}^i) \qquad \text{(self-financing condition)} .$$

By construction (θ^0, θ^*) is predictable.

Thus, writing $\theta = (\theta^0, \theta^*)$, for all $n \leq N$:

$$\hat{V}_n(\theta) = \theta_n^0 + \theta_n^1 \hat{S}_n^1 + \cdots + \theta_n^d \hat{S}_n^d .$$

Using (2.2) and the equalities $\Delta \hat{S}_n^0 = 0$ and $\hat{V}_0(\theta) = 0$, we obtain

$$\hat{V}_n(\theta) = \sum_{j=1}^n (\theta_j^1 \Delta \hat{S}_j^1 + \cdots + \theta_j^d \Delta \hat{S}_j^d) .$$

Thus, using the definition of Q, we obtain, for all θ^*

$$E_Q\left(\sum_{j=1}^N \theta_j^1 \Delta \hat{S}_j^1 + \cdots + \theta_j^d \Delta \hat{S}_j^d\right) = E_Q(\hat{V}_N(\theta))$$

$$= \frac{1}{\Lambda} \sum_\omega l(\omega) \, \hat{V}_N(\theta)(\omega) = 0$$

and thus that $E_Q(\theta_n^i \hat{S}_n^i) = E_Q(\theta_n^i \hat{S}_{n-1}^i)$ for any $\mathcal{F}_{n-1}$-measurable random variable θ_n^i. This is equivalent to the martingale property of $\hat{S}^i$ (under Q). $\qquad \square$

The theorem remains true in discrete time when Ω is no longer finite. See for example Dalang, Morton and Willinger [80], Schachermayer [325], and Kabanov and Kramkov [225].

Remark 2.2.7. In Harrison and Kreps [177] for example, one comes across the following definition of an arbitrage opportunity: it is a self-financing strategy such that

$$P(V_N(\theta) \geq 0) = 1 ; \qquad P(V_N(\theta) > 0) > 0$$

and

$$P(V_0(\theta) \leq 0) = 1 .$$

We remark that the assumption on V_N can be written $E_P[V_N(\theta)] > 0$; $V_N(\theta) \geq 0$ P-a.s.. It is obvious that the arbitrage opportunities defined in Sect. 2.2.1 are arbitrage opportunities in the sense of Harrison and Kreps.

Theorem 2.2.6 enables us to prove the converse. We suppose that there are no arbitrage opportunities, in the sense of Sect. 2.2.1. Then there exists a probability measure Q under which $\hat{V}_n(\theta)$ is a martingale for any self-financing strategy. We have $\hat{V}_0(\theta) = E_Q(\hat{V}_N(\theta) \mid \mathcal{F}_0)$. Hence, since $\hat{V}_N(\theta)$ is positive and has strictly positive expectation (under P and under Q, as P and Q are equivalent), we obtain that $\hat{V}_0(\theta) = V_0(\theta)$ is positive with strictly positive expectation; there are no arbitrage opportunities in the sense of Harrison and Kreps.

Definition 2.2.8. *The model is arbitrage-free if there exist no arbitrage opportunities.*

Notation 2.2.9. We denote by $\mathcal{P}$ the set of probability measures that are equivalent to P and that make discounted prices into martingales.

2.3 Trees

Let Ω be finite with $P(\{\omega\}) > 0$ for all $\omega \in \Omega$. We want to show that to any filtration we can associate a tree, and that the NAO property defined previously is equivalent to a concept of NAO at each node of the tree.

Example 2.3.1. Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_i, i = 1, \ldots, 3$ be the σ-algebras generated by the following partitions of Ω :

$$\begin{aligned}
F_1 &= \{\{\omega_1, \omega_2, \omega_3\}, \{\omega_4\}, \{\omega_5\}\} \\
F_2 &= \{\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4\}, \{\omega_5\}\} \\
F_3 &= \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}, \{\omega_5\}\} .
\end{aligned}$$

To this filtration, we associate the following tree: at time 0, there is a single node; at time 1, there are three nodes, the first corresponds to the atom of F_1, $\{\omega_1, \omega_2, \omega_3\}$, the second to $\{\omega_4\}$, the third to $\{\omega_5\}$. At time 2, there are four nodes corresponding to the four atoms of F_2. The two first correspond respectively to $\{\omega_1\}$ and to $\{\omega_2, \omega_3\}$ and follow on from the time-1 node $\{\omega_1, \omega_2, \omega_3\}$. Finally, at time 3, there are five nodes, one for each state

of the world. The first corresponds to $\{\omega_1\}$ and follows on from the time-1 node $\{\omega_1, \omega_2, \omega_3\}$ and from the time-2 node $\{\omega_1\}$.

Thus we find that to each node of the tree, we can associate, on the one hand its past, and on the other, the set of states of the world that follow it.

Fig. 2.1. Example 2.3.1
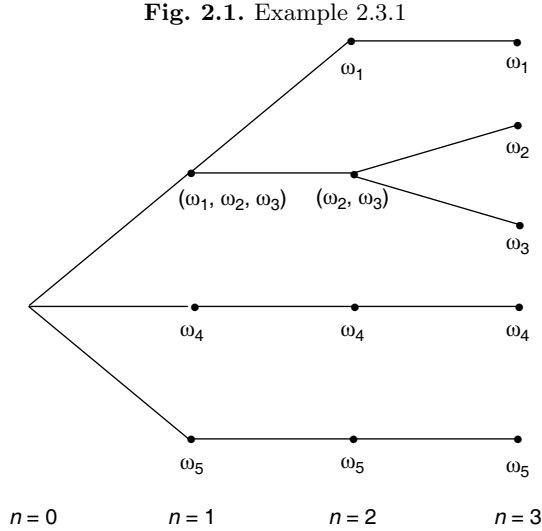
More generally, let us suppose that we are given a filtration $(\mathcal{F}_n)_{n=0}^N$, and let F_n be the partition of Ω induced by the non-trivial atoms of $\mathcal{F}_n$. We suppose that F_n has I_n elements that we denote by A_n^i, $i \leq I_n$. To this filtration we associate the following tree. At time 0, there is a single node. At time 1, there are I_1 nodes, each corresponding to an atom A_1^i of F_1. Each atom of F_1 is the union of atoms of F_2: $A_1^i = \cup_{j \in J_1^i} A_2^j$. The node i therefore has J_1^i successors. At time 2, there are I_2 nodes, each corresponding to an atom of F_2, and succeeding a unique node, the atom from which it emerges. The node j of the tree at time n is the atom A_n^j of F_n.

Let j be a node of the tree at time n. We denote by $\Delta^n(j)$ the set of nodes succeeding j at time $n+1$:

$$\Delta^n(j) = \{A_{n+1}^l \in F_{n+1} \mid A_{n+1}^l \subseteq A_n^j\}\,.$$

We can identify any $\mathcal{F}_n$-measurable real-valued mapping with a mapping from F_n into $\mathbb{R}$. A portfolio strategy $\theta = (\theta_1, \ldots, \theta_N)$ is a family of mappings $\theta = (\theta_n)_{n=1}^N$ such that $\theta_n : F_{n-1} \to \mathbb{R}^{d+1}$.

As the price process is assumed to be adapted to the filtration, S_n is constant on A_n^j. The prices are therefore well-defined at each node of the tree. We denote their value at the node j and at the time n by $S_n(j)$.

Fig. 2.2. A General Tree

$n = 0$ $n = 1$ $n = 2$
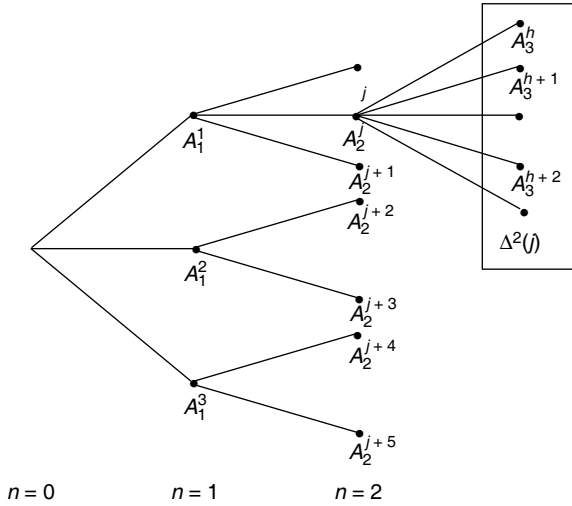
Definition 2.3.2. *We say that there is arbitrage on the tree if there exists n, $0 \leq n < N$, $A_n^j \in F_n$ and $\theta_{n+1} \in \mathbb{R}^{d+1}$ such that one of the two following equivalent conditions is satisfied:*

1) $\theta_{n+1} \cdot (\hat{S}_{n+1} - \hat{S}_n) \geq 0$, $\forall \; \omega \in A_n^j$ *with a strict inequality for* $\omega \in A_{n+1}^l \in \Delta^n(j)$,

2) $\theta_{n+1} \cdot S_n \leq 0$, $\forall \; \omega \in A_n^j$ *and* $\theta_{n+1} \cdot S_{n+1} \geq 0$, $\forall \; \omega \in A_n^j$ *with a strict inequality for* $\omega \in A_{n+1}^l \in \Delta^n(j)$.

Lemma 2.3.3. *Conditions 1) and 2) of the previous definition are equivalent.*

Proof. Let us first show that 1) implies 2). Let θ_{n+1} satisfy 1), and let $\tilde{\theta}_{n+1}$ be defined by

$$\tilde{\theta}_{n+1}^0 = -\sum_{j=1}^d \theta_{n+1}^j \hat{S}_n^j \quad \text{and} \quad \tilde{\theta}_{n+1}^j = \theta_{n+1}^j, \quad \forall j \geq 1 \; .$$

By construction, $\tilde{\theta}_{n+1} \cdot S_n = 0$ and

$$\tilde{\theta}_{n+1} \cdot S_{n+1} = -S_{n+1}^0 \sum_{j=1}^{d} \theta_n^j \hat{S}_n^j + \sum_{j=1}^{d} \theta_n^j S_{n+1}^j$$

$$= S_{n+1}^0 (\theta_{n+1} \cdot (\hat{S}_{n+1} - \hat{S}_n)) \geq 0, \quad \forall \omega \in A_n^j$$

with the inequality being strict for $\omega \in A_{n+1}^l$. Hence 2) holds.

Conversely, if θ_{n+1} satisfies 2)

$$\theta_{n+1}^0 + \sum_{j=1}^{d} \theta_{n+1}^j \hat{S}_n^j \leq 0 \quad \text{and} \quad \theta_{n+1}^0 + \sum_{j=1}^{d} \theta_{n+1}^j \hat{S}_{n+1}^j \geq 0 \text{ on } A_n^j$$

with the inequality being strict for all $\omega \in A_{n+1}^l$. Subtracting one inequality from the other,

$$\theta_{n+1} \cdot (\hat{S}_{n+1} - \hat{S}_n) \geq 0, \ \forall \, \omega \in A_n^j$$

with the inequality being strict for $\omega \in A_{n+1}^l$. Hence 1) holds.   □

Proposition 2.3.4. *Under the assumption of NAO on the tree, there exists a probability measure Q on Ω, which is equivalent to P, and under which discounted prices are martingales.*

Proof. Let us consider the set of nodes at time $n + 1$ and springing from A_n^j:

$$\Delta^n(j) = \{A_{n+1}^l \in F_{n+1} \mid A_{n+1}^l \subseteq A_n^j\} .$$

As there is NAO between the node j and its successor nodes, according to Proposition 1.2.8 there exists a family of real numbers $\pi_n(A_n^j, A_n^\ell)$, $\ell \in \Delta^n(j)$ denoted by $\pi_n(j, \ell)$ such that

(i) $\pi_n(j, \ell) > 0$,

(ii) $\sum_{\ell \in \Delta^n(j)} \pi_n(j, \ell) = 1$,

(iii) $\hat{S}_n^i(j) = \sum_{\ell \in \Delta^n(j)} \pi_n(j, \ell) \hat{S}_{n+1}^i(\ell)$.

The $\pi_n(j, \ell)$ can be interpreted as transition probabilities between times n and $n + 1$ for going between node j and its successors. Let Q be the unique probability measure on Ω such that $Q(A_{n+1}^l) = \pi_n(j, \ell)Q(A_n^j)$ and $Q(A_0) = 1$. Since $Q(\omega) > 0$, $\forall \, \omega$, Q is equivalent to P. Relation (iii) then becomes:

$$\hat{S}_n^i = E_Q(\hat{S}_{n+1}^i \mid \mathcal{F}_n) .$$

The prices, discounted by the riskless rate of return, are martingales.   □

Proposition 2.3.5. *There is NAO in the model if and only if there is NAO on the tree.*

Proof. If there is NAO on the tree, then there exists a probability measure Q on Ω under which the asset prices, discounted by the riskless rate of return, are martingales. According to Proposition 2.2.5, there is then NAO in the model.

Conversely, if there are no arbitrage opportunities in the model, then there exists Q satisfying $Q(A_n) > 0$, $\forall A_n \in F_n$, $\forall\, n$ and under which prices, discounted by the riskless rate of return, are martingales. Let

$$
\pi_n(j, \ell) = \begin{cases} \dfrac{Q(A^l_{n+1})}{Q(A^j_n)} & \text{if} \quad A^l_{n+1} \subseteq A^j_n \\[2ex] 0 & \text{otherwise .} \end{cases}
$$

As $\hat{S}^i_n = E_Q(\hat{S}^i_{n+1} \mid \mathcal{F}_n)$,

$$
\hat{S}^i_n(j) \;=\; \sum_{\ell \in \Delta^n(j)} \pi_n(j, \ell)\, \hat{S}^i_{n+1}(\ell) \;.
$$

Therefore, $\theta_n \cdot S_{n+1}(l) \geq 0$ with a strict inequality for at least one l, implies $\theta_n \cdot S_n(j) > 0$. There is NAO on the tree. $\qquad\square$

2.4 Complete Markets with a Finite Horizon

Definition 2.4.1. *A $\mathcal{F}_N$-measurable random variable X is said to be* replicable[3] *if there exists a self-financing strategy θ such that $V_N(\theta) = X$.*

Definition 2.4.2. *A market is complete if any $\mathcal{F}_N$-measurable random variable X is replicable*

In other words, a market is complete if any $\mathcal{F}_N$-measurable random variable X satisfies $X = \sum_{i=0}^{d} \theta^i_N S^i_N$ where θ is a self-financing strategy.

Using (2.2), it is easy to show that such a r.v. X can be written

$$
(S^0_N)^{-1} X \;=\; \hat{V}_0(\theta) + \sum_{n=1}^{N} \theta_n \cdot \Delta \hat{S}_n \;,
$$

which we can also write in the form

$$
(S^0_N)^{-1} X \;=\; \hat{V}_0(\theta) + \int_0^N \theta \, d\hat{S} \;,
$$

---

[3] The words *attainable* and *hedgeable* are also used.

where $\int_0^N \theta \, d\hat{S}$ is defined as the integral of the step function that equals θ_n on $]n, n+1]$, with respect to $\hat{S}$. After a change of measure, this is a martingale if the model is arbitrage-free.

In the next chapter, we will extend this concept to continuous-time processes: in a complete arbitrage-free market, any $\mathcal{F}_N$-measurable random variable can be written, after discounting, as the integral of a predictable[4] process, with respect to the discounted price process, which is itself a martingale.

2.4.1 Characterization

Proposition 2.4.3. *An arbitrage-free market is complete if and only if there exists a unique probability measure Q that is equivalent to P and under which discounted prices martingales.*

Proof. Let us first prove the direct implication:
We suppose the market to be arbitrage-free and complete, and we suppose that P_1 and P_2 belong to $\mathbb{P}$. Let X be a $\mathcal{F}_N$-measurable random variable. There exists a strategy θ such that $X = V_N(\theta)$, that is, after discounting, such that

$$\frac{X}{S_N^0} = \hat{V}_N(\theta) .$$

We have already remarked that $\hat{V}_n(\theta)$ is a P_1-martingale (and a P_2-martingale). Hence

$$E_{P_i}(\hat{V}_N(\theta)) = E_{P_i}(V_0(\theta)) = V_0(\theta) .$$

It follows that

$$E_{P_1}\left(\frac{X}{S_N^0}\right) = E_{P_2}\left(\frac{X}{S_N^0}\right) .$$

As this equality holds for all $\mathcal{F}_N$-measurable X, , we have $P_1 = P_2$ on $\mathcal{F}_N$.

Proof of the converse:

Assume that there exists a unique equivalent martingale probability measure P^*. Let us prove that the market is complete.

We reproduce the proof of Lamberton-Lapeyre [250].

We suppose that the market is arbitrage-free, but not complete. Hence there exists a non-replicable random variable. Then the space $\mathcal{E}$ of random variables of the form:

$$U_0 + \sum_{n=1}^N \theta_n \cdot \Delta\hat{S}_n \tag{2.3}$$

---

[4] A property of measurability, which we will come back to.

(with U_0 constant, and with $(\theta_n^1, \ldots, \theta_n^d)_{1 \leq n \leq N}$ a predictable process), is a strict subspace of the space of all the random variables defined on $(\Omega, \mathcal{F})$. Endow the space of random variables with the scalar product $(X, Y) \rightarrow E^*(XY)$. Then there exists a non-zero X, orthogonal to $\mathcal{E}$. Let us now set:

$$P^{**}(\{\omega\}) \;=\; \left(1 + \frac{X(\omega)}{2\|X\|_\infty}\right) P^*(\{\omega\})$$

where $\|X\|_\infty \;=\; \sup_{\omega \in \Omega} |X(\omega)|$. Note that P^{**} is a probability measure ($P^{**}(\Omega) = 1$ as the constant random variable equal to 1 belongs to $\mathcal{E}$, and hence, by orthogonality, $E^*(X) = 0$), equivalent to P, and not equal to P^* (as X is non-zero). As discounted prices are martingales under P^*, we have in addition:

$$E_{P^{**}} \left(\sum_{n=1}^N \theta_n \cdot \Delta \hat{S}_n\right)$$

$$= \; E_{P^*} \left(\sum_{n=1}^N \theta_n \cdot \Delta \hat{S}_n\right) + \frac{1}{2\|X\|_\infty} E_{P^*} \left(X \sum_{n=1}^N \theta_n \cdot \Delta \hat{S}_n\right) \;=\; 0$$

since $\hat{S}_n$ is a P^*-martingale and X is orthogonal to $\mathcal{E}$. This entails, by a reasoning used previously, that $(\hat{S}_n)_{0 \leq n \leq N}$ is a P^{**}-martingale, thus contradicting the assumption that P^* was a unique equivalent martingale measure. Hence the market is complete.   □

Remark 2.4.4. Proposition 2.4.3 is true in general, even in the case of continuous time. The proof then makes use of a predictable representation theorem, and of some very subtle theorems of Jacod. We refer the interested reader to Harrison and Pliska [179], Jacod [205], and to the articles by Delbaen and Schachermayer, [96, 97].

2.5 Valuation

Under the hypothesis of NAO, replicable variables can be valuated.

Proposition 2.5.1. *Let X be a replicable random variable, and let θ be a self-financing strategy such that $V_N(\theta) = X$. The value of $V_0(\theta)$ does not depend on the choice of θ, and is called the value of X at time 0. We have*

$$V_0(\theta) = E_Q(X(S_N^0)^{-1})$$

for any self-financing strategy θ and for any probability measure Q under which discounted prices are martingales.

Proof. Let $\theta_1 = (\theta_{n,1})_{n \leq N}$ and $\theta_2 = (\theta_{n,2})_{n \leq N}$ be two self-financing strategies such that $V_N(\theta_1) = V_N(\theta_2) = X$. We suppose that $V_0(\theta_1) > V_0(\theta_2)$. Let θ^* be the d-dimensional process that corresponds to the risky part of $\theta_2 - \theta_1$:

$$\theta^* := (\theta_{n,2}^1 - \theta_{n,1}^1, \ldots, \theta_{n,2}^d - \theta_{n,1}^d)_{n \leq N}$$

where $\theta_{n,1} = (\theta_{n,1}^i)_{i \leq d}$ (respectively $\theta_{n,2} = (\theta_{n,2}^i)_{i \leq d}$).

As in Theorem 2.2.6, we can construct a process θ^0 such that the strategy (θ^0, θ^*) is self-financing and has zero initial value. Writing $(\psi^0, \mathbf{0})$ for a strategy $(\psi^0, 0, \ldots, 0)$, which only involves investing in the riskless asset, and noticing that V_N is linear with respect to the strategy, we have

$$\begin{aligned} V_N(\theta^0, \theta^*) &= V_N(\theta_2 - \theta_1) + V_N(\theta^0 - (\theta_2^0 - \theta_1^0), \mathbf{0}) \\ &= V_N(\theta^0 - (\theta_2^0 - \theta_1^0), \mathbf{0}) . \end{aligned}$$

As the strategies θ_1 and θ_2 are self-financing, the strategy $(\theta^0 - (\theta_2^0 - \theta_1^0), \mathbf{0})$ also is, as it is the difference of self-financing strategies. Using the definition of θ^0,

$$\begin{aligned} V_N(\theta^0, \theta^*) &= V_N(\theta^0 - (\theta_2^0 - \theta_1^0), \mathbf{0}) = (1 + r)^N V_0(\theta^0 - (\theta_2^0 - \theta_1^0), \mathbf{0}) \\ &= -(1 + r)^N V_0(\theta_2 - \theta_1) > 0 . \end{aligned}$$

Thus we can obtain a strategy with zero initial value, and such that $V_N(\theta^0, \theta^*) > 0$, which is impossible according to the assumption of NAO. Hence $V_0(\theta_1) = V_0(\theta_2)$. If Q_1 and Q_2 are two probability measures under which discounted prices are martingales, then

$$E_{Q_1}(\hat{V}_N(\theta)) = E_{Q_2}(\hat{V}_N(\theta)) \qquad (= V_0(\theta))$$

for any self-financing strategy θ. □

Remark 2.5.2. If we know that $\mathbb{P}$ contains a unique measure Q, the result of Proposition 2.5.1 is obvious, since $V_0(\theta) = E_Q(\hat{V}_N(\theta))$.

The same proof shows that if there exists a strategy θ such that $V_N(\theta) = 0$ and $V_0(\theta) < 0$, then there exists an arbitrage opportunity. This proof will be generalized in Chap. 3.

2.5.1 The Complete Market Case

When the market is arbitrage-free and complete, we can value all $\mathcal{F}_N$-measurable random variables. We need to calculate:

$$E_Q \left(\frac{X}{S_N^0} \right) = V_0(\theta) .$$

More generally, noting that $\hat{V}_n(\theta)$ is a Q-martingale, we have

$$V_n(\theta) \;=\; S_n^0 \, E_Q \left(\frac{X}{S_N^0} \;\middle|\; \mathcal{F}_n \right) .$$

We call $V_n(\theta)$ the price of X at time n. It is interesting to note that the calculation depends on Q (and not on P).

In particular, if we want to price a call with maturity N on the asset S^1, and if asset 0 is riskless and has rate of return r, then the value of the call at time n is

$$V_n(\theta) = (1+r)^n E_Q \left(\frac{(S_N^1 - K)^+}{(1+r)^N} \;\middle|\; \mathcal{F}_n \right) .$$

The value of a put at time n is

$$V_n(\theta) = (1+r)^n E_Q \left(\frac{(K - S_N^1)^+}{(1+r)^N} \;\middle|\; \mathcal{F}_n \right) .$$

We can check that the put–call parity holds, by noticing that

$$(S_n - K)^+ = S_n - K + (K - S_n)^+ .$$

The strategy θ that enabled us to valuate X can be interpreted as a hedging strategy for the seller of the contract X.

More generally, we can valuate a cash flow, that is a process $(X_n)_{n=1}^N$ adapted to the filtration. Its value at time n is

$$V_n(\theta) = (1+r)^n E_Q \left(\sum_{t=n}^N \frac{X_t}{(1+r)^t} \;\middle|\; \mathcal{F}_n \right) .$$

2.6 An Example

2.6.1 The Binomial Model

This model was developed by Cox, Ross and Rubinstein [71]. There are two assets:

- a riskless asset whose rate of return r is independent of the time period and of the state of the world,
- a stock whose rate of return between the time n and the time $n+1$ can be either u, or d with $d < 1 + r < u$.

At time 0, the stock is worth S. At time 1, the stock can be worth either Su or Sd, and at time 2, Su^2 Sud or Sd^2. The price of the stock at time n depends only on the number of up-moves between time 0 and time n, and can be worth Su^n, $Su^{n-1}d$, $\ldots, Sd^n$.

In this model, a state of the world is a sequence of values that the stock takes between time 0 and time N. The partition F_1 has two elements, the set of the states of the world such that the stock is worth either Su or Sd at time 1. Similarly, the partition F_2 has four elements, the set of the states of the world such that the stock is worth Su at time 1, then Su^2 at time 2, Su then Sud, Sd then Sdu and finally Sd then Sd^2. At time n, F_n has 2^n elements.

We can associate a tree with this model. At time n, there are 2^n nodes, and for all j and for all n, $\Delta^n(j)$ has only two elements.

Fig. 2.3. Tree for the Binomial Model

$n = 0$ $n = 1$ $n = 2$ $n = N$
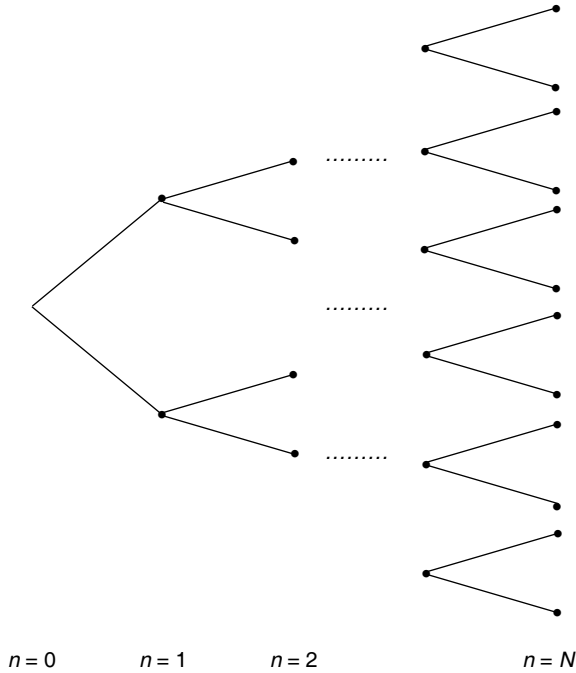
As we have shown in the previous section, if there is NAO in the model, then there exists probability π for moving between a time-n node and its "up" successor at time $n+1$, which is independent of the node and of the instant in time, and where, using (iii) from the proof of Proposition 2.3.4:

$$\pi = \frac{1}{u-d} \{1 + r - d\} .$$

In this binomial model, π depends neither on the instant in time, nor on the position of the node in the tree.

As the market is complete, there exists a unique martingale measure, which we can obtain explicitly as was done in Proposition 2.3.4. As the price of the

stock at time n depends only of the number of up-moves between time 0 and time n, under this new measure, S_n is a random variable which has the binomial distribution with parameters n and π:

$$Q(S_n = u^j d^{n-j} S) = \binom{n}{j} \pi^j (1 - \pi)^{n-j}$$

where

$$\binom{n}{j} = \frac{n!}{j!(n-j)!} .$$

2.6.2 Option Valuation

We can now compute the value of a European option. If we work with n periods, we have

$$C = \frac{1}{(1+r)^n} E_Q(S_n - K)^+$$

and hence

$$C = \frac{1}{(1+r)^n} \sum_{j=0}^n \binom{n}{j} \pi^j (1 - \pi)^{n-j} (u^j d^{n-j} S - K)^+ . \qquad (2.4)$$

Let us rewrite (2.4) in a slightly different form. Let

$$\eta = \inf\{j \in N \mid u^j d^{n-j} S - K > 0\} .$$

If $[\![\alpha]\!]$ denotes the integer part of α, that is to say the integer such that $[\![\alpha]\!] \le \alpha < [\![\alpha]\!] + 1$, we can see that

$$\eta = \left[\!\left[\frac{\ln K/Sd^n}{\ln u/d} \right]\!\right] + 1 \qquad \text{where ln is the natural logarithm .}$$

(in financial terms, η is the minimum number of up-moves that are required over n periods in order to be "in the money", i.e., to make a strictly positive profit). We then have (using the definition of $(\cdot)^+$)

$$C = \frac{1}{(1+r)^n} \sum_{j=\eta}^n \binom{n}{j} \pi^j (1 - \pi)^{n-j} (u^j d^{n-j} S - K) .$$

The coefficient $d(n, j; \pi) = \binom{n}{j} \pi^j (1 - \pi)^{n-j}$ represents the probability that a random variable with the binomial distribution of parameters n and π, takes the value j. Let us introduce the notation:

$$D(n, \eta; \pi) = \sum_{j=\eta}^n d(n, j; \pi) .$$

Using the definition of π and the equality $1 - \frac{\pi u}{1+r} = \frac{d - d\pi}{1+r}$, we obtain the following result.

Proposition 2.6.1. *The price of a call option in the binomial model is given by*

$$C = S \sum_{j=\eta}^{n} nj \left(\frac{\pi u}{1+r}\right)^j \left(\frac{d - d\pi}{1+r}\right)^{n-j} - \frac{K}{(1+r)^n} D(n, \eta; \pi)$$

$$= S D\left(n, \eta; \frac{\pi h}{1+r}\right) - \frac{K}{(1+r)^n} D(n, \eta; \pi) \,.$$

(2.5)

The same reasoning establishes the price of a put:

$$P = \frac{1}{(1+r)^n} \sum_{j=0}^{n} \binom{n}{j} \pi^j (1-\pi)^{n-j} (K - u^j d^{n-j} S)^+ \,.$$

We can show that the put–call parity is still satisfied.

2.6.3 Approaching the Black–Scholes Model

We would like to take the limit of (2.5) in order to approach a continuous-time model. We study the financial market on an interval of time T. Firstly, we suppose that we have n periods of length $\dfrac{T}{n}$, and we study a binomial model over these n periods. Next we let n tend to infinity, whilst T stays fixed.

The Choice of Parameters

We have supposed that the coefficients u, d, r and π, which appear in (2.5), do not depend on the period we are looking at, but they do of course depend on the length of the period. Here, the length of each period is $\dfrac{T}{n}$. The coefficients will thus depend on n. Let us write them u_n, d_n, r_n and π_n.

If our aim is to approach the continuous-time model, we must ensure the equality between the returns in continuous-time, and those in discrete time when we let $n \to \infty$. Thus we must have:

$$\lim_{n \to \infty} (1 + r_n)^n = e^{\rho T}$$

where ρ is the instantaneous rate of return. To explain our choice of u_n and d_n, we need to carry out a few calculations.

We divide the interval $[0, T]$ up into n periods of length $\dfrac{T}{n}$. Let S_n be the asset price evaluated in the binomial model after n periods.

If j denotes the number of up-moves, this price is $S_n(j) := S u_n^j d_n^{n-j}$, which can be written

$$\ln \frac{S_n}{S} = J \ln \frac{u_n}{d_n} + n \ln d_n$$

where J is a random variable which has the binomial distribution with parameters n and p_n, where p_n denotes the probability of an up-move (we suppose that this probability does not depend on the period). In particular, we have $E(J) = n p_n$ and $\text{Var}(J) = n p_n (1 - p_n)$. We denote by $n \nu_n$ the expectation of $\ln \left(\frac{S_n}{S} \right)$, and by $n \sigma_n^2$ its variance.

Let S_T be the asset price at time T. We require the discrete model to produce a price S_n that tends to S_T in the sense that the expectation $E \left(\ln \frac{S_T}{S} \right)$ of the logarithm of the return on the asset is approximated by the expectation of $\ln \left(\frac{S_n}{S} \right)$. Similarly, we impose an analogous condition on the variances: $\text{Var} \left(\ln \frac{S_T}{S} \right)$ must be the limit of $\text{Var} \left(\ln \frac{S_n}{S} \right)$.

Using the distribution of S_n, we obtain:

$$\begin{cases} n \nu_n = n \left(p_n \ln \dfrac{u_n}{d_n} + \ln d_n \right) \\[2mm] n \sigma_n^2 = n p_n (1 - p_n) \left(\ln \dfrac{u_n}{d_n} \right)^2 . \end{cases}$$

We take $p_n = 1/2$ for the sake of simplicity. We noticed earlier that the price of the option did not depend on the value that this probability took.

We then impose on the coefficients u_n and d_n to be such that $n \nu_n$ converges to νT and $n \sigma_n^2$ converges to $\sigma^2 T$, where ν and σ^2 represent the expectation and variance of the "instantaneous" logarithm of the return on the asset, i.e., $\sigma^2 T = \text{Var} \left(\ln \frac{S_T}{S} \right)$ and $V T = E \left[\ln \frac{S_T}{S} \right]$.

To achieve this, we can take for example,

$$u_n = e^{\nu \frac{T}{n} + \sigma \sqrt{\frac{T}{n}}}, \qquad d_n = e^{\nu \frac{T}{n} - \sigma \sqrt{\frac{T}{n}}} . \tag{2.6}$$

Proposition 2.6.2. *Under the previous conditions, S_n/S converges in distribution to $e^{\nu T + \sigma \sqrt{T} G}$ where G is a standard Gaussian variable.*

Proof. Indeed,

$$\ln \left(\frac{S_n}{S} \right) = J \ln \frac{u_n}{d_n} + n \ln d_n$$

$$= \nu T + \frac{\sigma \sqrt{T}}{\sqrt{n}} (2J - n)$$

The central limit theorem applied to J, which is the sum of n independent identically distributed random variables with the Bernoulli distribution of mean $1/2$ and variance $1/4$, implies that

$$\frac{J - n/2}{\sqrt{n/4}} = \frac{2J - n}{\sqrt{n}}$$

converges in distribution to a standard normal variable. It follows that

$$\nu T + \frac{\sigma\sqrt{T}}{\sqrt{n}}(2J - n)$$

converges in distribution to a Gaussian variable with mean νT and variance $\sigma^2 T$. $\qquad\qquad\square$

Limit of the Option Price

We study the behaviour of

$$S\,D\left(n\,,\,\eta_n;\,\frac{\pi_n u_n}{1 + r_n}\right) - \frac{K}{(1 + r_n)^n}\,D(n, \eta_n; \pi_n)$$

with

$$\eta_n \;=\; \left[\!\!\left[\frac{\ln K/Sd_n^n}{\ln u_n/d_n}\right]\!\!\right] + 1 \quad \text{and} \quad \pi_n \;=\; \frac{1 + r_n - d_n}{u_n - d_n}\,.$$

We produce detailed calculations for $D(n, \eta_n; \pi_n)$.

$D(n, \eta; \pi) = 1 - P(Y_n < \eta)$ where Y_n is the sum of n independent random variables with the Bernoulli distribution with parameters π (here π depends on n as well as on η, we omit the subscript n to lighten the notation). Then

$$P(Y_n < \eta) \;=\; P\left(\frac{Y_n - n\pi}{\sqrt{n\pi(1 - \pi)}} < \frac{\eta - n\pi}{\sqrt{n\pi(1 - \pi)}}\right)\,.$$

Using the definition of η and (2.6), we see that

$$\eta_n = \frac{1}{2}n + \frac{\ln\frac{K}{S} - \nu T}{2\sigma\sqrt{T}}\sqrt{n} + o(\sqrt{n})\,.$$

Replacing π_n by its value as a function of u_n and d_n, we obtain

$$\pi_n - \frac{1}{2} \simeq \frac{\sqrt{T}}{2\sigma\sqrt{n}}\left(\rho - \nu - \frac{1}{2}\sigma^2\right)\,.$$

As a result, $\dfrac{\eta_n - n\pi_n}{\sqrt{n\pi_n(1 - \pi_n)}}$ converges to $\dfrac{\ln K/S - \left(\rho - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}\,.$

It remains to apply the central limit theorem. The classic version of the theorem does not apply here, as the probability laws involved depend on n. We can use Lindeberg's theorem[5]. We can also check the result ourselves:

Proposition 2.6.3. *For all n, let $(X_{1,n}; X_{2,n}; \ldots ; X_{n,n})$ be n independent identically distributed random variables distributed according to the probability law $P(X_{i,n} = 1) = 1 - P(X_{i,n} = 0) = \pi_n$. Let $Y_n = \sum_{i=1}^{n} X_{i,n}$.*

Then $\frac{Y_n - n\pi_n}{\sqrt{n\pi_n(1-\pi_n)}}$ converges in distribution (when n tends to infinity) to the standard normal distribution.

Proof. We use characteristic functions. Let $\phi(t) = E(\exp itX)$ be the characteristic function associated with a random variable X. If X has zero expectation and variance 1, then we have $\phi(t) = 1 - \frac{t^2}{2} + o(t^2)$. As the random variables $X_{i,n}$ are independent,

$$E\left(\exp it \, \frac{Y_n - n\pi_n}{\sqrt{n\pi_n(1-\pi_n)}}\right) = \left\{E\left(\exp it \, \frac{X_{1,n} - \pi_n}{\sqrt{n\pi_n(1-\pi_n)}}\right)\right\}^n .$$

The variables $\dfrac{X_{1,n} - \pi_n}{\sqrt{n\pi_n(1-\pi_n)}}$ have zero expectation and variance $1/n$, hence we can check that

$$E\left(\exp it \, \frac{X_{1,n} - \pi_n}{\sqrt{n\pi_n(1-\pi_n)}}\right) = 1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right) .$$

Hence the characteristic function of $\dfrac{Y_n - n\pi_n}{\sqrt{n\pi_n(1-\pi_n)}}$ converges to $\exp\left(-\dfrac{t^2}{2}\right)$. $\qquad\square$

The Black–Scholes Formula

Bringing together the results above, we obtain

$$D(n, \eta_n; \pi_n) \rightarrow 1 - \phi\left(\frac{\ln K/S - (\rho - 1/2\,\sigma^2)T}{\sigma\sqrt{T}}\right)$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} \, du .$$

It remains to note that $1 - \phi(x) = \phi(-x)$, to obtain

---

[5] See Chung [58].

$$D(n, \eta_n; \pi_n) \;\to\; \phi\!\left(\frac{\ln S/K + (\rho - 1/2\,\sigma^2)T}{\sigma\sqrt{T}} \right).$$

The same methods can be applied to study $D(n, \eta_n;\, \pi_n u_n/1 + r_n)$.

Theorem 2.6.4. *When $n \to \infty$, under conditions (2.6), C converges to*

$$S\,\phi(d) - K\,\mathrm{e}^{-\rho T}\,\phi(d - \sigma\sqrt{T})$$

with

$$d \;=\; \frac{\ln\,(S/K) + \rho T}{\sigma\sqrt{T}} + \frac{1}{2}\,\sigma\,\sqrt{T}\,.$$

This formula is known as the *Black–Scholes* formula. We will give a direct proof of the formula in a more general setting in Chap. 3. We notice that the limit does not depend on μ.

The Case of a Put

We obtain that the price of a put converges to

$$K\,\mathrm{e}^{-\rho T}\phi(\delta + \sigma\sqrt{T}) - S\,\phi(\delta)$$

with

$$\delta \;=\; \frac{\ln\,(K/S) - \rho T}{\sigma\sqrt{T}} - \frac{1}{2}\,\sigma\,\sqrt{T}\,.$$

Hence we can check the put–call parity formula holds:

$$C = P + S - \mathrm{e}^{-\rho T}K\,.$$

2.7 Maximization of the Final Wealth

In this section, we suppose Ω to be finite. An investor, with initial wealth x at time 0, maximizes the expectation of a utility function of his final wealth. More specifically, let $U : \mathbb{R}_+ \to \mathbb{R}$ satisfy:

U1 U is strictly concave, strictly increasing, and of class C^1,

U2 $\lim_{x \to +\infty} U'(x) = 0$, $\lim_{x \to 0} U'(x) = \infty$.

Let $I :\,]0, \infty[\,\to\,]0, \infty[$ be the inverse of U'. If follows from **U1** that I is strictly decreasing and continuous.

We suppose that the investor solves the following problem $\mathcal{P}$:

Maximize $E_P\big[U\left(V_N(\theta)\right)\big]$ under the contraints

$$\begin{cases} V_N(\theta) \geq 0 \\ V_0(\theta) = x \end{cases}$$

where $V_N(\theta) = \theta_N \cdot S_N$ and where θ is a self-financing strategy.

Proposition 2.7.1. *There is an optimal strategy if and only if there is NAO.*

Proof. If there is an arbitrage opportunity $\tilde{\theta}$, then for any self-financing strategy θ with initial value x, the strategy $\tilde{\theta} + \theta$ is a self-financing strategy with initial value x and $V_N(\tilde{\theta} + \theta) \geq V_N(\theta)$ a.e. with a strict inequality on a set of strictly positive measure. Therefore, there cannot be an optimal solution.

Conversely, let us consider the set C of positive terminal values of wealth that the investor can achieve with a self-financing strategy:

$$C = \{V_N(\theta) \mid V_N(\theta) \geq 0,\ V_0(\theta) = x,\ \theta \text{ self-financing}\}.$$

Let us show that the set C is the positive cone of the set

$$\mathcal{K} = \{V_N(\theta) \mid V_0(\theta) = x,\ \theta \text{ self-financing}\}.$$

We show that

$$\mathcal{K} = \left\{ \left(x + \sum_{k=1}^{N} \theta_k \Delta \hat{S}_k \right) S_N^0,\ \theta \text{ a predictable process taking values in } \mathbb{R}^d \right\}.$$

Indeed, it follows from (2.2) that if θ is self-financing, then

$$\hat{V}_N(\theta) = V_0(\theta) + \sum_{k=1}^{N} \theta_k \Delta \hat{S}_k = x + \sum_{k=1}^{N} \theta_k \Delta \hat{S}_k.$$

Conversely, for a given predictable process θ taking values in $\mathbb{R}^d$, we define the process θ^0 by the self-financing condition

$$\theta_1^0 = x - \sum_{j=1}^{d} \theta_1^j S_1^j,\ \theta_{n+1}^0 = \theta_n^0 + \sum_{j=1}^{d} (\theta_n^j - \theta_{n+1}^j) \hat{S}_n^j.$$

The final wealth associated with this strategy depends only on θ, and equals

$$\left(x + \sum_{k=1}^{N} \theta_k \Delta \hat{S}_k \right) S_N^0.$$

The set $\mathcal{K}$ is the translation of a vector space of finite dimension, and is therefore a closed convex set. We show that the set C is compact. As we have NAO, there exists a martingale measure Q. According to Proposition 2.2.4,

$$E_Q\left(\frac{V_N}{S_N^0}\right) = x .$$

Since $Q(\omega) > 0$, $\forall\omega$, then $0 \le V_N(\omega) \le \frac{xS_N^0}{Q(\omega)}$, $\forall\omega$, and so C is bounded. The function U being continuous, the optimization problem has a solution. The optimal final wealth V_N^* is unique, as U is strictly concave.   □

Proposition 2.7.2. *The optimal final wealth V_N^* is strictly positive, and*

$$Q(\omega) = \frac{P(\omega)S_N^0 U'(V_N^*(\omega))}{E_P(S_N^0 U'(V_N^*))}$$

is a martingale measure.

Proof. Let us first show that V_N^* is strictly positive. We suppose that there exists $A_N \in F_N$ such that $V_N^* = 0$ on A_N. We denote by θ the strategy that involves investing everything in the riskless asset up until time N: $\theta_n^0 = x$, $\forall n \ge 1$, and $\theta_n^j = 0$, $\forall n \ge 1$, $\forall j \ge 1$. We then have $V_N(\theta) = S_N^0 x$. Let us consider the strategy $\varepsilon\theta + (1-\varepsilon)\theta^*$ where $\theta_N^* \cdot S_N = V_N^*$. We have

$$E_P\left[U((1-\varepsilon)V_N^* + \varepsilon V_N) - U(V_N^*)\right] \ge \varepsilon\Big[E_P(1_{A_N}U'[(1-\varepsilon)V_N^* + \varepsilon V_N]V_N)$$
$$+ E_P(1_{A_N^c}U'[(1-\varepsilon)V_N^* + \varepsilon V_N](V_N - V_N^*))\Big]$$

For a sufficiently small ε, the expression on the left-hand side is strictly positive since $U'((1-\varepsilon)V_N^* + \varepsilon V_N) \to \infty$ on A_N when $\varepsilon \to 0$. This contradicts the optimality of V_N^*.

We remark that for any self-financing strategy θ with initial value x, we have

$$V_N(\theta) = S_N^0\left[x + \sum_{k=1}^N \theta_k \cdot \left(\hat{S}_k - \hat{S}_{k-1}\right)\right] .$$

As in the proof of Theorem 2.2.6, if $\tilde{\theta}$ is a predictable process taking values in $\mathbb{R}^d$, $\tilde{\theta} = (\theta_n^1, \theta_n^2, \ldots, \theta_n^d)_{n \le N}$, we can construct θ^0, a process taking values in $\mathbb{R}$ and such that the strategy $(\theta^0, \theta^1, \ldots, \theta^d)$ is self-financing and with initial value x. Let

$$W(\tilde{\theta}) = E_P\left[U\left(S_N^0\left(x + \sum_{k=1}^N \theta_k \cdot (\hat{S}_k - \hat{S}_{k-1})\right)\right)\right] .$$

The investor's problem can be written as an optimization problem without constraints:

$$\max W(\tilde{\theta})$$

$$\tilde{\theta} \text{ predictable,}$$

where θ_n^j is $\mathcal{F}_{n-1}$-measurable, and can be identified as a vector of $\mathbb{R}^{I_{n-1}}$ and where $\theta_n^j 1_{A_{n-1}}$ can be identified as one of its components. As the first order conditions are necessary and sufficient, the partial derivatives of $W(\tilde{\theta})$ with respect to the variables $\theta_n^j 1_{A_{n-1}}$ are zero at θ^*, $\forall n$, $\forall A_{n-1} \in F_{n-1}$, $\forall j$. Hence, in particular, by differentiating with respect to $\theta_n^j 1_{A_{n-1}}$, we obtain

$$\sum_{\omega \in A_{n-1}} P(\omega) U'(V_N^*(\omega)) S_N^0(\omega)(\hat{S}_n^j(\omega) - \hat{S}_{n-1}^j(\omega)) = 0 \ ,$$

and hence

$$1_{A_{n-1}}(\omega) \hat{S}_{n-1}^j(\omega) = \frac{\displaystyle\sum_{\omega \in A_{n-1}} P(\omega) U'(V_N^*(\omega)) \ S_N^0(\omega) \hat{S}_n^j(\omega)}{\displaystyle\sum_{\omega \in A_{n-1}} P(\omega) U'(V_N^*(\omega)) S_N^0(\omega)} \ .$$

This expression holds for all $A_{n-1} \in F_{n-1}$ and for all n, so we have

$$\hat{S}_{n-1}^j = \frac{E_P(U'(V_N^*) \hat{S}_n^j \mid \mathcal{F}_{n-1})}{E_P(U'(V_N^*) \mid \mathcal{F}_{n-1})} \ .$$

Hence, using Bayes' rule,

$$\hat{S}_{n-1}^j = E_Q(\hat{S}_n^j \mid \mathcal{F}_{n-1})$$

where

$$Q(\omega) = \frac{P(\omega) S_N^0(\omega) U'(V_N^*(\omega))}{E_P(S_N^0(\omega) U'(V_N^*(\omega)))} \ .$$

$\square$

Proposition 2.7.3. *If the market is complete and without arbitrage, then* $V_N^* = I\left(\frac{\lambda L}{S_N^0}\right)$ *where Q is the unique martingale measure, where $L = \frac{dQ}{dP}$ and where λ is determined by the equation* $E_Q\left(I\left(\frac{\lambda L}{S_N^0}\right) \frac{1}{S_N^0}\right) = x$.

Proof. As the market is complete, there exists a unique martingale measure Q, and the set of positive final wealth attainable from an initial wealth of x is therefore:

$$\left\{ V_N \mid V_N \geq 0, \ E_Q\left(\frac{V_N}{S_N^0}\right) = x \right\} \ .$$

Let $L = \frac{dQ}{dP}$ be the density of Q with respect to P. The investor's problem comes down to an optimization problem with a single constraint.

Maximize $E_P\left[U(V_N)\right]$ under the contraints

$$\begin{cases} V_N \geq 0 \\ E_P\left(\dfrac{LV_N}{S_N^0}\right) = x \end{cases}$$

Let λ be the Lagrange multiplier associated with the problem. V_N^* is optimal if and only if it is the optimal solution to the problem:

$$\text{Maximize } E_P\left[U(V_N) - \lambda\frac{LV_N}{S_N^0}\right] \text{ under the constraint}$$

$$V_N \geq 0$$

and $E_P\left(\frac{LV_N}{S_N^0}\right) = x$. Since $U'(0) = \infty$, we recover the fact that $V_N^* > 0$ and $U'(V_N^*) = \frac{L\lambda}{S_N^0}$. Hence $V_N^* = I(\frac{\lambda L}{S_N^0})$. Let us show that λ is determined by the equation $E_Q\left(I\left(\frac{\lambda L}{S_N^0}\right)\frac{1}{S_N^0}\right) = x$. The mapping $\lambda \to E_Q\left(I\left(\frac{\lambda L}{S_N^0}\right)\frac{1}{S_N^0}\right)$ is strictly decreasing from $\mathbb{R}_+$ into $\mathbb{R}_+$. Therefore there exists a unique λ such that $E_Q\left(\frac{1}{S_N^0}I(\frac{\lambda L}{S_N^0})\right) = x$. □

2.8 Optimal Choice of Consumption and Portfolio

In this section, we suppose once again that Ω is finite. We now suppose that there is a single good for consumption at each date and in each state of the world, and that this good is taken as numéraire. We assume that the investor has initial wealth x at time 0, and maximizes the expectation of the utility of his consumption stream. Hence he solves the problem:

$$\max E\left[\sum_{n=0}^{N} \alpha^n U(c_n)\right]$$

(where U satisfies the assumptions **U1** and **U2**) under self-financing constraints that we write:

$$\begin{aligned} \theta_n \cdot S_n &= c_n + \theta_{n+1} \cdot S_n, \quad 0 < n < N \\ x &= c_0 + \theta_1 \cdot S_0 \end{aligned} \tag{2.7}$$

$$\theta_N \cdot S_N = c_N \tag{2.8}$$

and

$$c_n \geq 0, \quad \forall\, n \tag{2.9}$$

where θ_n is $\mathcal{F}_{n-1}$-measurable and where c_n is $\mathcal{F}_n$-measurable for all n. We say that the strategy θ finances $(c_n)_{n=0}^N$.

Equations (2.7) and (2.8) are equivalent to

$$V_0 = x - c_0$$

$$\hat{V}_n = x + \sum_{k=1}^{n} \theta_k \cdot (\hat{S}_k - \hat{S}_{k-1}) - \sum_{k=0}^{n-1} \frac{c_k}{S_k^0} \quad 0 < n \leq N \qquad (2.10)$$

$$V_N = c_N .$$

Proposition 2.8.1. *There is an optimal strategy if and only if there is NAO. The optimal consumption is unique.*

Proof. We suppose that there is an optimal stream $(c_n^*)_{n=0}^{N}$ financed by θ^*, and an arbitrage opportunity $\tilde{\theta}$. Then the strategy $\tilde{\theta} + \theta^*$, which finances the stream c_n, $0 \leq n < N$, $c_N = (\tilde{\theta}_N + \theta_N) \cdot S_N$, has initial value x and $U(c_N) \geq U(c_N^*)$ with the inequality being strict on a set of strictly positive measure. Therefore, there can be no optimal solution.

Conversely, let us consider the set

$$C = \left\{ (c_n)_{n=0}^{N}, \ c_n \geq 0 \ \forall n \mid \text{ there exists } \theta \text{ satisfying } (2.7) \text{ and } (2.8) \right\} .$$

It is a convex set. As in Proposition 2.7.1, we show that

$$C = \Big\{ (c_n)_{n=0}^{N}, \ c_n \geq 0 \ \forall n \mid \text{ there exists a predictable } \theta ,$$

$$\text{taking values in } \mathbb{R}^d \text{and such that } \sum_{k=0}^{N} \hat{c}_k = x + \sum_{k=1}^{N} \theta_k (\hat{S}_k - \hat{S}_{k-1}) \Big\} .$$

The set $\left\{ x + \sum_{k=1}^{N} \theta_k (\hat{S}_k - \hat{S}_{k-1}), \ \theta \text{ predictable} \right\}$ is closed, as a translated vector space; C is thus closed.

Let us show that C is also compact. As there is NAO, there exists a martingale measure Q. Since

$$\frac{c_N}{S_N^0} = x + \sum_{k=1}^{N} \theta_k \cdot (\hat{S}_k - \hat{S}_{k-1}) - \sum_{k=0}^{N-1} \frac{c_k}{S_k^0} ,$$

and

$$E_Q \left(\sum_{k=0}^{N} \frac{c_k}{S_k^0} \right) = x ,$$

C is bounded. As the function U is continuous, the optimization problem has a solution. Finally, as the function U is strictly concave, the optimal consumption stream is unique. $\qquad \square$

Proposition 2.8.2. *The optimal consumption* $(c_n^*)_{n=1}^N$ *is strictly positive, and*

$$Q(\omega) = \frac{P(\omega)U'(c_N^*(\omega))}{E_P(U'(c_N^*))}$$

is a martingale measure.

Proof. The proof of the strict positivity of the optimal consumption stream $(c_n^*)_{n=0}^N$ is identical to that of Proposition 2.7.2, and is therefore omitted. We remark that the utility associated with a self-financing strategy θ with initial value x, is

$$E\left[U(x - \theta_1 \cdot S_0) + \sum_{k=1}^{N-1} \alpha^k U((\theta_k - \theta_{k+1}) \cdot S_k) + \alpha^N U(\theta_N \cdot S_N)\right]. \quad (2.11)$$

As the optimal consumption stream $(c_n^*)_{n=0}^N$ is strictly positive, we obtain by differentiating (2.11) with respect to $\theta_n^j 1_{A_{n-1}}$,

$$\sum_{\omega \in A_{n-1}} P(\omega)\left[-U'(c_{n-1}^*)S_{n-1}^j + \alpha U'(c_n^*)S_n^j\right] = 0, \quad \forall\, j = 0, \ldots, d$$

and hence

$$\alpha^{n-1}S_{n-1}^j U'(c_{n-1}^*) = \alpha^n E_P(U'(c_n^*)S_n^j \mid \mathcal{F}_{n-1}), \quad j = 0, \ldots, d. \quad (2.12)$$

It follows from (2.12) that $\alpha^n S_n^j U'(c_n^*)$ is a martingale for all $j = 0, \ldots, d$. In particular, for $j = 0$, we obtain that

$$\alpha^n(1+r)^n U'(c_n^*) \quad (2.13)$$

is a martingale, and hence that

$$\frac{S_{n-1}^j U'(c_{n-1}^*)}{(1+r)^{n-1}U'(c_{n-1}^*)} = \frac{E_P(S_N^j U'(c_N^*) \mid \mathcal{F}_{n-1})}{(1+r)^N E_P(U'(c_N^*) \mid \mathcal{F}_{n-1})}.$$

In other words, applying Bayes' rule (see annex),

$$\hat{S}_{n-1}^j = E_Q(\hat{S}_N^j \mid \mathcal{F}_{n-1}) \quad (2.14)$$

where

$$\frac{\mathrm{d}Q}{\mathrm{d}P} = \frac{U'(c_N^*)}{E(U'(c_N^*))}.$$

□

Proposition 2.8.3. *If the market is complete,* $c_n^* = I\left(\frac{\lambda L_n}{\alpha^n S_n^0}\right)$ *where Q is the unique martingale measure, where $L = \frac{\mathrm{d}Q}{\mathrm{d}P}$, where $L_n = E_P(L \mid \mathcal{F}_n)$ and where λ is determined by the equation*

$$\sum_{n=0}^N E_P\left(\frac{L_n}{S_n^0} I\left(\frac{\lambda L_n}{\alpha^n S_n^0}\right)\right) = x.$$

Proof. As this proof is very similar to that of Proposition 2.7.3, we content ourselves with giving its outline.

Since the market is complete, there exists a unique martingale measure Q, and the set of non-negative consumption streams that are attainable with an initial wealth of x is therefore:

$$\left\{ (c_n)_{n=0}^N, \ c_n \geq 0, \ \forall n \ \middle| \ E_Q \left(\sum_{n=0}^N \frac{c_n}{S_n^0} \right) = x \right\} .$$

Let $L_N = \frac{dQ}{dP}$ be the density of Q with respect to P on $\mathcal{F}_N$, and let $L_n = E_P[L_N | \mathcal{F}_n]$. Using $E_Q(X_n) = E_P(L_n X_n)$ for any $\mathcal{F}_n$-measurable random variable X_n,

$$E_Q \left(\sum_{n=0}^N \frac{c_n}{S_n^0} \right) = E_P \left(\sum_{n=0}^N \frac{c_n L_n}{S_n^0} \right) .$$

The investor's problem then comes down to an optimization problem with a single constraint:

$$\text{Maximize} \sum_{n=0}^N \alpha^n E_P \big[U(c_n) \big]$$

$$\begin{cases} c_n \geq 0, \quad 0 \leq n \leq N \\ \displaystyle\sum_{n=0}^N E_P \left(\frac{c_n L_n}{S_n^0} \right) = x . \end{cases}$$

Let λ be the Lagrange multiplier associated with this relation. Since $c_n^* > 0$ we have $\alpha^n U'(c_n^*) = \frac{L_n \lambda}{S_n^0}$, $\quad \forall n$, and hence $c_n^* = I(\frac{\lambda L_n}{S_n^0 \alpha^n})$ where λ is determined by the equation

$$\sum_{n=0}^N E_P \left(\frac{L_n}{S_n^0} I \left(\frac{\lambda L_n}{\alpha^n S_n^0} \right) \right) = x ,$$

which admits a unique solution. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

When the market is incomplete, the optimal solution can be calculated by generalizing to a dynamic framework, the method presented in Chap. 1 Sect. 1.3. It is also possible to use the method of dynamic programming, which we outline here.

For our exposition, we take the tree associated with the filtration $(\mathcal{F}_n)_{0 \leq n \leq N}$. Using the notation introduced in Sec. 2.3, we define a sequence of value functions, as follows. Let $x \in \mathbb{R}_+$ represent wealth.

$$V_N(x) = U(x)$$

$$\forall A_{N-1} \in F_{N-1}, \ V_{N-1}(x, A_{N-1}) = \max U(c_{N-1}) + \alpha E \big(V_N(\theta_N \cdot S_N) \mid A_{N-1} \big)$$

under the contraints

$$x = c_{N-1} + \theta_N \cdot S_{N-1}$$

$$c_{N-1} \geq 0 .$$

We write $c_{N-1}(x, A_{N-1})$ and $\theta_N(x, A_{N-1})$ for the optimal solution of this problem. $V_{N-1}(x, A_{N-1})$ represents the maximal utility that can be attained at the node A_{N-1} of the tree, for an investor who only lives for the last period between times $N-1$ and N, has at his disposal a amount of wealth x, and adjusts his consumption between times $N-1$ and N.

$$\forall A_{N-2} \in F_{N-2}, \; V_{N-2}(x, A_{N-2})$$

$$= \max U(c_{N-2}) + \alpha E\big(V_{N-1}(\theta_{N-1} \cdot S_{N-1}, .\,) \mid A_{N-2}\big)$$

$$x = c_{N-2} + \theta_{N-1} \cdot S_{N-2}$$

$$c_{N-2} \geq 0 ,$$

where $V_{N-1}(y, .\,)$ is the r.v. that equals $V_{N-1}(y, A_{N-1})$ at the node A_{N-1}. By backward induction, we define

$$\forall A_t \in F_t, \; V_t(x, A_t) = \max U(c_t) + \alpha E\big(V_{t+1}(\theta_{t+1} \cdot S_{t+1}, .\,) \mid A_t\big)$$

$$x = c_t + \theta_{t+1} \cdot S_t$$

$$c_t \geq 0 .$$

We denote by $c_t(x, A_t)$ and $\theta_{t+1}(x, A_t)$ the optimal solution to this problem.

The sequences $\big(V_t(x, A_t), \quad c_t(x, A_t), \quad \theta_t(x, A_t)\big)_{0 \leq t \leq N}$ are known, so we use them to deduce the optimal consumption and portfolio:

$$c_0^* = c_0(x), \qquad\qquad \theta_1^* = \theta_1(x)$$

$$c_1^* = \big[c_1(\theta_1^* \cdot S_1, A_1)\big]_{A_1 \in F_1} \qquad \theta_2^* = \big[\theta_2(\theta_1^* \cdot S_1, A_1)\big]_{A_1 \in F_1}$$

and at time t,

$$c_t^* = \big[c_t(\theta_t^* \cdot S_t, A_t)\big]_{A_t \in F_t} \qquad \theta_{t+1}^* = \big[\theta_{t+1}(\theta_t^* \cdot S_t, A_t)\big]_{A_t \in F_t}$$

where a $\mathcal{F}_t$-measurable variable is identified with the vector of the values it takes on the atoms of time t.

Remark 2.8.4. As in Propositions 2.7.2 and 2.8.2, it is possible to show that under the assumptions **U1** and **U2**, $c_t(x, A_t) > 0$, $\forall \; (t, \; A_t, \; x)$. Having eliminated the consumption, we deduce, using the envelope theorem, that $V_{N_1}(., A_{N-1})$ is differentiable $\forall \; A_{N-1}$, and by backward induction, that $V_t(x, A_t)$ is differentiable $\forall \; A_t$ and that $V_t'(x, A_t) = U'(c_t(x, A_t))$. In particular,

$$V_t'(\theta_t^* \cdot S_t, A_t) = U'(c_t^*) . \tag{2.15}$$

However, at time t, once the consumption has been eliminated, the investor solves the optimization problem without contraints

$$\max_{\theta_{t+1}} U(x - \theta_{t+1} \cdot S_t) + \alpha E\big(V_t(\theta_{t+1} \cdot S_{t+1}, \cdot) \mid A_t\big) .$$

By differentiating this expression with respect to θ_{t+1} and using (2.15), we recover equation (2.12):

$$S_t^j = \alpha E_P \left(\frac{U'(c_{t+1}^*)}{U'(c_t^*)} S_{t+1}^j \mid \mathcal{F}_t \right), \quad j = 0, \ldots, d .$$

2.9 Infinite Horizon

The theory that we have developed in the previous sections comes up against a number of difficulties when the dynamic model has an infinite and countable number of periods, even if we assume that at each date, there are only a finite number of states of nature.

The first difficulty is linked to the fact that an investor can "roll over" his debt indefinitely. As an illustration, let us give an example, which is due to Pliska [301].

We consider the binomial model of Sect. 2.6.1 with an infinite horizon, and where we suppose that $u = 1$, $d = 0.9$ and $r = 0$. We are going to show that an investor can ensure that he gains 1 euro in the future, without owning anything at time 0.

At time 0, he borrows 10 euros, which he invests in stock. At time 1, if the price of the stock has risen, he sells his portfolio and receives 11 euros, he repays his debt, and thus has a gain of 1 euro. If on the contrary, the price has fallen, and the value of his portfolio is now only 9 euros, he borrows 11 euros, which he invests in the stock, so that with what he held previously, he has constructed a portfolio of 20 euros in stock. His debt is now $10 + 11 = 21$ euros. At time 2, if the price of the stock goes up, he sells his portfolio, receives 22 euros, repays his debt and so has a gain of 1 euro. If the price goes down, the value of the portfolio is now only 18 euros, and he borrows 22 euros (so doubling his stake) and he now owes a total of $22 + 11 + 10$ euros. At time $n \geq 2$, if the price has fallen n times, the value of his portfolio is $9 \times 2^{n-1}$ euros, the investor has borrowed $10 + 11(1 + \cdots + 2^{n-2}) = 11 \times 2^{n-1} - 1$ euros, and his net wealth is $-(11 \times 2^{n-1} - 1) + 9 \times 2^{n-1} = -2^n + 1$ euros.

When there are a finite number of periods, the event "the price falls n times running" has a strictly positive probability, and the strategy put forward above is not an arbitrage opportunity. However, when there are an infinite and countable number of periods, the probability that the price falls

every time is zero, and the strategy is an arbitrage opportunity.

To avoid this type of strategy, it suffices for example to impose that the wealth be uniformly bounded below, or that strategies be uniformly bounded.

The definition of NAO runs into still more difficulties. We specify our model, in order to explain them.

Let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with a filtration $(\mathcal{F}_n)_{n=0}^{\infty}$ where $\mathcal{F}_0 = \{\phi, \Omega\}$.

As in the previous sections, there are $d+1$ assets. Their prices at time n, $S_n = (S_n^0, S_n^1, \ldots, S_n^d)^T$, are taken to be $\mathcal{F}_n$-adapted. Asset 0 is assumed to be riskless: $S_0^0 = 1$, $S_n^0 = (1+r)^n$ where r is the interest rate, which we assume to be constant.

A portfolio strategy is a family $\theta = (\theta_n)_{n=1}^{\infty}$ of random vectors $\theta_n = (\theta_n^0, \theta_n^1, \ldots, \theta_n^d)$ such that

$$\forall\, n,\ \forall i \geq 0, \quad \theta_n^i \quad \text{is } \mathcal{F}_{n-1}\text{-measurable.}$$

$$\text{and} \quad \exists k > 0 \text{ such that } \|\theta_n\| \leq k,\ \forall\, n\,.$$

A portfolio strategy θ is self-financing if

$$\theta_n \cdot S_n = \theta_{n+1} \cdot S_n, \quad \forall\, n \geq 1\,.$$

At first, it seems natural to generalize Definition 2.2.1 as follows:
A "finite" arbitrage opportunity is a self-financing strategy θ such that

(i) $P(V_0(\theta) = 0) = 1,$

(ii) $\exists N$ such that $P(V_N(\theta) \geq 0) = 1; \quad P(V_N(\theta) > 0) > 0.$

It is then easy to generalize Proposition 2.2.5, and to show that if there exists a martingale measure equivalent to P, then there are no finite arbitrage opportunities. The following example, which we owe to Schachermayer [326], shows that the converse is not true. In the following example, there are no finite arbitrage opportunities, and there exists a martingale measure Q, but it is not equivalent to P.

We consider a binomial model with an infinite horizon, and where $r = 0$ and $S_n = S_{n-1} + \varepsilon_n + \alpha_n,\ \ \forall n \geq 1$ with the α_n being constants in $]0, 1[$. We suppose that under P, the r.v. $(\varepsilon_n)_{n=1}^{\infty}$ are independent and that $P(\varepsilon_n = 1) = P(\varepsilon_n = -1) = \frac{1}{2},\ \forall\, n \geq 1.$

In such a model, there exists a unique martingale measure Q defined by

$$Q(\varepsilon_n = 1) = \frac{1 - \alpha_n}{2}, \qquad Q(\varepsilon_n = -1) = \frac{1 + \alpha_n}{2}, \quad \forall\, n \geq 1\,,$$

where the random variables $(\varepsilon_n)_{n=1}^\infty$ are independent. We then have with $\eta_i = \pm 1$, $1 \leq i \leq n$,

$$\frac{\mathrm{d}Q}{\mathrm{d}P}(\varepsilon_1 = \eta_1, \varepsilon_2 = \eta_2, \dots, \varepsilon_n = \eta_n) = \prod_{i=1}^{n}(1 - \eta_i \alpha_i)\,,$$

and $\frac{\mathrm{d}Q}{\mathrm{d}P}\mid_{\mathcal{F}_n} = \Pi_{i=1}^n(1 - \varepsilon_i \alpha_i)$. It follows from Kakutani's theorem (see Williams [367] p. 150) that Q is equivalent to P on $\mathcal{F}_\infty$ if and only if $\sum_{n=1}^{\infty} \sqrt{1 + \alpha_n} + \sqrt{1 - \alpha_n} < \infty$, which is equivalent to $\sum_{n=1}^{\infty} \alpha_n^2 < \infty$. It follows from this same theorem that unless this condition is satisfied, Q is singular with respect to P.

Thus we are led to define a more restrictive notion of NAO. Developing the ideas of Kreps [242] and Schachermayer [326] has shown that by giving a good definition of NAO, we can obtain an equivalence between NAO and the existence of an equivalent martingale measure.

Moreover, we have shown in Sects. 2.7 and 2.8 that when there are a finite number of states of nature and periods, there is an equivalence between NAO, the existence of an equivalent martingale measure, and the existence of an optimal solution to the problem of an optimal consumption–portfolio choice. We now show that under certain assumptions, we can extend Proposition 2.8.2.

We denote by C, the set of adapted real-valued processes $(c_n)_{n \geq 0}$ such that $\sup_n \|c_n\|_\infty < \infty$.

As in Sect. 2.8, we suppose that there is a single consumption good at each time and in each state of the world, and that it is taken as numéraire. We assume that an investor, with initial wealth x at time 0, maximizes the expectation of the utility of his consumption stream, i.e., he solves the problem:

$$\max E\left[\sum_{n=0}^{\infty} \alpha^n U(c_n)\right]$$

(with U satisfying the assumptions **U1** and **U2**) under the self-financing constraints, which we write

$$\begin{aligned}\theta_n \cdot S_n &= c_n + \theta_{n+1} \cdot S_n, \quad \forall\, n > 0 \\ x &= c_0 + \theta_1 \cdot S_0\,,\end{aligned}$$

where we suppose moreover that $(c_n, \theta_{n+1})_{n \geq 0} \in C^{d+2}$ and $c_n \geq 0$, $\forall\, n$.

Proposition 2.9.1. *If there is an optimal consumption stream* $(c_n^*)_{n=1}^\infty$, *and if there exists* $m > 0$ *such that* $c_n^* \geq m$, $\forall n$, *then there exists a martingale measure* Q, *which is equivalent to* P, *and* $\frac{\mathrm{d}Q}{\mathrm{d}P} = \lim_n \frac{U'(c_n^*(\omega))}{E_P(U'(c_n^*))}$.

Proof. The proof is analogous to that of Proposition 2.8.2. We note that the utility associated with a self-financing strategy θ with initial value x, is

$$E\left[U(x - \theta_1 \cdot S_0) + \sum_{k=1}^\infty \alpha^k U((\theta_k - \theta_{k-1}) \cdot S_k)\right], \qquad 0 < \alpha < 1 .$$

By differentiating with respect to θ_n and proceeding as in the proof of Proposition 2.8.2, we obtain that

$$\alpha^n (1 + r)^n U'(c_n^*)$$

is a martingale, that $\frac{U'(c_n^*(\omega))}{E_P(U'(c_n^*))}$ is a martingale and that

$$\hat{S}_{n-1}^j = E_{\hat{Q}}\left(\hat{S}_n^j \mid \mathcal{F}_{n-1}\right)$$

where $\frac{\mathrm{d}Q}{\mathrm{d}P}\big|_{\mathcal{F}_n} = \frac{U'(c_n^*)}{E_P(U'(c_n^*))}$. The martingale $\frac{U'(c_n^*(\omega))}{E_P(U'(c_n^*))}$ is positive and uniformly integrable since $m \leq c_n^* \leq M$, $\forall n \geq 0$. It thus converges in $L^1(P)$ to $\lim_n \frac{U'(c_n^*(\omega))}{E_P(U'(c_n^*))}$. Hence we have

$$\hat{S}_{n-1}^j = E_Q(\hat{S}_n^j \mid \mathcal{F}_n), \quad \text{with} \quad \frac{\mathrm{d}Q}{\mathrm{d}P} = \lim_n \frac{U'(c_n^*(\omega))}{E_P(U'(c_n^*))} .$$

$\square$

Remark 2.9.2.

1. The assumption that $c_n^* \geq m$, $\forall n$ means that (c_n^*) is in the interior of l_+^∞.

2. If we introduce an exogenous dividend process $(d_n)_{n \in N} \in C^{d+1}$ with $d_0 = 0$, $\forall\, n$, we write the self-financing constraints in the form

$$\begin{aligned}
\theta_n \cdot (S_n + d_n) &= c_n + \theta_{n+1} \cdot S_n, \quad \forall\, n > 0 \\
x &= c_0 + \theta_1 \cdot S_0 .
\end{aligned}$$

We define the discounted gains process by

$$\hat{G}_n^j = \hat{S}_n^j + \sum_{k=1}^n \hat{d}_k^j, \quad \text{with } \hat{d}_k = \frac{d_k}{(1+r)^k} .$$

Then, under the same assumptions and using the same notation as in Proposition 2.9.1, we find that the discounted gains process is a martingale

$$\hat{G}_n^j = E_Q(\hat{G}_{n+1}^j \mid \mathcal{F}_n), \quad j = 1, \ldots, d \quad \text{with } \frac{\mathrm{d}Q}{\mathrm{d}P} = \lim_n \frac{U'(c_n^*(\omega))}{E_P(U'(c_n^*))} \ ,$$

which can also be written in the following form, called "Lucas' formula"

$$S_n^j = \frac{1}{U'(c_n^*)} E_P \left(\sum_{k=n+1}^{\infty} \alpha^{k-n} U'(c_k^*) d_k^j \mid \mathcal{F}_n \right), \quad j = 1, \ldots, d \ .$$

Notes

Sect. 2.2 is heavily based on the papers Harrison and Kreps, [177], (1979) and Harrison and Pliska, [178], (1981). Theorem 2.2.6 has been extended to the case of an infinite number of states of nature by Dalang, Morton and Willinger, [80], (1989). The problems of arbitrage with portfolio constraints or with transaction costs have been studied by Jouini and Kallal, [221, 222], (1995), Schuger, [328], (1996), Jouini and Napp [223] (2001).

The procedure for taking limits in Sect. 2.6 is due to Cox, Ross and Rubinstein, [71], (1979). It may be tempting to work in discrete time and obtain results in continuous time by taking the limit of the discrete-time model. This approach is often very difficult to implement. We refer to the book of Prigent [303] (2002) for recent results. American options, superhedging under constraints are studied in Föllmer and Schied [162] (2004).

For Sect. 2.8, we could have considered more general forms of utility function. For example, Epstein and Zin, [153], (1989) use recursive utilities, Dunn and Singleton [130], (1986) use utilities that at time t depend both on the consumption at time t and on past consumption, and Epstein and Wang, [152], (1994) assume that agents have multiple priors over states of the world. We can also introduce portfolio constraints or obtain an explicit solution when the market is incomplete (see Pliska, [301], 1997).

We have not dealt with the dynamic version of the CCAPM. The reader is referred to Huang and Litzenberger, [197], (1988). For a model with heterogeneous beliefs, see Jouini and Napp [220].

For the case where the state process is a Markov chain, we can consult Duffie, [114], (1991), and for the dynamic programming and economic modeling aspects, Stokey and Lucas, [346], (1989).

When the market is incomplete, one can consult Föllmer and Schied [162] (2004) for superhedging and minimization of the hedging error.

ANNEX 2

Conditional Expectation and Martingales

The book Chung [58] is a good reference for the definitions and properties relating to martingales in discrete time.

Definition *Let X be an integrable random variable, and let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$. The conditional expectation of X with respect to $\mathcal{G}$, denoted by $E(X|\mathcal{G})$, is the unique (up to an equality that holds almost surely) random variable such that*

(i) $E(X|\mathcal{G})$ is a $\mathcal{G}$-measurable random variable;

(ii) $\int_G E(X|\mathcal{G})\mathrm{d}P = \int_G X\mathrm{d}P, \forall G \in \mathcal{G}$.

The conditional expectation satisfies the following properties:

Property

a. *The conditional expectation is linear: if Y is integrable and if a and b are two real numbers, $E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$.*

b. *It is increasing: if $Y \leq X$, $E(Y|\mathcal{G}) \leq E(X|\mathcal{G})$.*

c. *If X belongs to $L^2(P)$, $E(X|\mathcal{G})$ is the projection of X onto $L^2(\mathcal{G})$.*

d. *If X is $\mathcal{G}$-measurable, $E(X|\mathcal{G}) = X$.*

e. *$E(E(X|\mathcal{G})) = E(X)$.*

f. *If $\mathcal{G}_i$ are two sub-σ-fields of $\mathcal{F}$ such that $\mathcal{G}_1 \subseteq \mathcal{G}_2$, we have*

$$E(E(X|\mathcal{G}_1)|\mathcal{G}_2) = E(E(X|\mathcal{G}_2)|\mathcal{G}_1) = E(X|\mathcal{G}_1).$$

g. *If X and Y belong to $L^2(P)$ and if Y is $\mathcal{G}$-measurable:*

$$E(XY|\mathcal{G}) = YE(X|\mathcal{G}).$$

h. *Fatou's lemma: if $|X_n| \leq Y$ where Y is integrable, and if X_n converges a.s. to X, then $E(X_n|\mathcal{G})$ converges a.s. to $E(X|\mathcal{G})$.*

i. *Jensen's inequality: for g a convex function, we have*

$$E(g(X))|\mathcal{F}) \geq g[E(X|\mathcal{F})].$$

All the inequalities and equalities above hold almost surely.

Definition *A sequence of r.v.'s $(M_n)_{0 \le n \le N}$ adapted to $(\mathcal{F}_n; 0 \le n \le N)$ is a martingale under P if*

(i) M_n is integrable with respect to P, $0 \le n \le N$,

(ii) $E_P(M_{n+1} | \mathcal{F}_n) = M_n$, $0 \le n \le N - 1$,

where $E_P(\cdot | \mathcal{F}_n)$ denotes the conditional expectation with respect to $\mathcal{F}_n$, and where the space $(\Omega, \mathcal{F})$ is equipped with the probability measure P.

A sequence of random vectors $S_n = (S_n^i, i \le d)$ is a martingale if the sequences $(S_n^i, 0 \le n \le N)$ are martingales. We denote by $E_P(S_{n+1} | \mathcal{F}_n)$ the vector with components $E_P(S_{n+1}^i | \mathcal{F}_n)$.

Uniform Integrability

Definition *A family of random variables $(X_i, i \in I)$ is said to be uniformly integrable if*

$$\lim_{a \to \infty} \sup_{i \in I} \int_{|x_i| \le a} |X_i| \mathrm{d}P = 0 \ .$$

In particular, if there exists an integrable Y such that $|X_i| \le Y, \forall i \in I$, then the family $(X_i, i \in I)$ is uniformly integrable.

Change of Probability Measure: the Radon-Nicodym Density

Definition *Let P and Q be two equivalent probability measures on $(\Omega, \mathcal{F})$. There exists a $\mathcal{F}$-measurable positive random variable f such that for every $A \in \mathcal{F}$, $Q(A) = E_P(f1_A)$. We write: $\dfrac{\mathrm{d}Q}{\mathrm{d}P} = f$.*

Conditional Expectation

If is important to be able to express the conditional expectation of a random variable X under Q with respect to its conditional expectation under P.

Property [Bayes' Rule] *We have*

$$E_Q(X | \mathcal{G}) = \frac{E_P(Xf | \mathcal{G})}{E_P(f | \mathcal{G})} \ .$$

3

The Black–Scholes Formula

In Chap. 2, we obtained the Black–Scholes formula by taking the limit of the binomial model. In this chapter, we present two further methods for obtaining the formula, and then show how analogous results can be obtained in a more general framework. The financial market comprises d risky assets, and one bond or riskless asset. Asset prices are modeled by means of a Brownian motion, using the notion of stochastic integral.

The first section is for the benefit of readers who are not familiar with stochastic calculus. We recall the definitions and basic results: Brownian motion, the stochastic integral with respect to Brownian motion, Itô processes, Itô's lemma, and Girsanov's theorem.

We approach the issue of the NAO assumption in the second section. An immediate warning: the equivalences obtained in Chap. 2 cannot entirely be generalized to continuous time.

The third section is devoted to the classic Black–Scholes formula. We present two methods for deriving it: one is based on solving partial differential equations, and the other on martingale theory. We then study how the value of an option varies as a function of the model's parameters.

In the fourth section, we price a financial product that pays dividends at each date, as well as a final dividend, in a financial market comprising one security modeled by a Markov diffusion process. We show that the arbitrage price is a solution to a partial differential equation, and can be expressed in terms of a conditional expectation.

Complementary probability results are to be found in the annex.

3.1 Stochastic Calculus

In this section, we recall some useful concepts and theorems from probability, and briefly present the model that we will be using.

We work on a finite interval of time $[0, T]$ and on a probability space $(\Omega, \mathcal{F}, P)$. When working with several different probability measures on the same space $(\Omega, \mathcal{F})$, we specify that we are working with probability measure P (respectively Q), by saying that we are under P (respectively under Q). We use the notation E_P for the expectation under P.

3.1.1 Brownian Motion and the Stochastic Integral

We have chosen to model the random phenomena that occur, by means of a Brownian motion. This is a process with *continuous* paths and stationary independent increments. More precisely:

Definition 3.1.1. $B = (B_t, t \geq 0)$ *is a real-valued Brownian motion starting from 0 on* $(\Omega, \mathcal{F}, P)$ *if*

a) $P(B_0 = 0) = 1$,

b) $\forall \, 0 \leq s \leq t$, *the real-valued random variable* $B_t - B_s$ *follows the normal distribution with mean 0 and variance* $t - s$,

c) $\forall \, 0 = t_0 < t_1 < \cdots < t_p$, *the variables* $(B_{t_k} - B_{t_{k-1}}, 1 \leq k \leq p)$ *are independent.*

We also use the notation $B(t)$ for B_t.

It can be shown that the mapping $t \to B_t(\omega)$ is continuous for almost all ω, i.e., the Brownian motion's paths are (almost surely) continuous.

An "intuitive" approach to Brownian motion is given in Appendix A. Merton [276] provides a justification for the use of Brownian motion.

We generalize the definition above, to obtain a Brownian motion with values in $\mathbb{R}^k$:

Definition 3.1.2. $B = (B^i, i \leq k)$ *is a k-dimensional Brownian motion if* B^i *are independent real-valued Brownian motions.*

A filtration $(\mathcal{F}_t, t \geq 0)$ on the probability space $(\Omega, \mathcal{F}, P)$ describes the flow of information available to an investor: if $A \in \mathcal{F}_t$, the investor knows at time t, whether or not A has occurred. We assume that a Brownian motion $(B_t, t \geq 0)$ is constructed on $(\Omega, \mathcal{F}, P)$ and we set $\mathcal{F}_t^B = \sigma(B_s, s \leq t)$. We will use the augmented filtration $\mathcal{F}_t$ generated by $\mathcal{F}_t^B$ and the P-null sets (for technical reasons). This filtration is increasing, i.e., $\mathcal{F}_s \subset \mathcal{F}_t$ for $s \leq t$. We say that a process $(X_t, t \geq 0)$ is $\mathcal{F}_t$-adapted if X_t is $\mathcal{F}_t$-measurable for any t. We write $E_t(\cdot) = E(\cdot | \mathcal{F}_t)$ for the conditional expectation of a variable with respect to $\mathcal{F}_t$.

Consider for a moment a model in which the price S_t of a stock is given by a Brownian motion, so that $S_t = B_t$. If an agent owns $\theta(t)$ shares at time t,

and trades at times t_k, the value at time t_K of a self-financing portfolio with initial value x, is $x + \sum_{k=1}^{K} \theta(t_k)\{B(t_k) - B(t_{k-1})\}$ (formula (2.1) in Chap. 2).

If we want trading to occur at any time t, we need to define a mathematical tool allowing us to replace the sum above – which resembles a Riemann sum – by something on which time acts continuously. We will replace it by an integral, written $\int_0^T \theta(s)\,dB(s)$, which we call the stochastic integral of θ with respect to B, and define as the limit (in L^2) of the sum above[1] (see annex).

It is easy to show that, for a fairly large class of *processes* θ, we can define the stochastic integral of θ with respect to B. It soon becomes apparent that we need to impose measurability conditions on θ: in Chap. 2 we assumed $\theta(t_k)$ to be $\mathcal{F}_{t_{k-1}}$-measurable, and described the resulting process as being predictable. The concept of a predictable process[2] is defined for processes indexed by t, $t \in \mathbb{R}_+$. In particular, a left-continuous process (such that the mapping $t \mapsto X_t(\omega)$ is continuous on the left for almost all ω) is predictable.

To define the stochastic integral $\int_0^T \theta(s)\,dB(s)$, we also need integrability conditions for θ.

Definition 3.1.3. *We denote by Θ, the set of predictable processes such that*

$$\int_0^T \theta^2(t)\,dt \;<\; \infty \qquad a.s. \,.$$

It can be shown that the stochastic integral $\int_0^T \theta(s)\,dB(s)$ is well-defined for $\theta \in \Theta$. We have of course $\int_s^t dB(u) = B(t) - B(s)$.

The stochastic integral with respect to Brownian motion has an extremely interesting property: it is a martingale[3], as long as additional integrability conditions hold.

The following result is an important tool:

Proposition 3.1.4. *If $\theta \in \Theta$ and if $E\left[\int_0^T \theta^2(s)\,ds\right] < \infty$, then the stochastic integral $\left(M_t = \int_0^t \theta(s)\,dB(s)\,; t \leq T\right)$ defines a process M, which is a martingale with zero expectation and with variance $E\left(M_t^2\right) = E\left[\int_0^t \theta^2(s)\,ds\right]$.*

Exercise 3.1.5. Show that $\left(B^2(t) - t\,,\; t \geq 0\right)$ is a martingale. One could proceed by showing that $E\left(B^2(t) - B^2(s)\,|\mathcal{F}_s\right) = t - s$, and noticing that this equals $E\left([B(t) - B(s)]^2|\mathcal{F}_s\right)$.

---

[1] For a detailed study of stochastic integrals with respect to Brownian motion, see Chung and Williams [59], or Karatzas and Shreve [233].

[2] An exact definition is to be found in the annex.

[3] See annex.

This exercise emphasizes one of the difficulties of stochastic calculus. The formula $\int_0^t B(s)\,\mathrm{d}B_s = \frac{1}{2}B_t^2$ is obviously wrong, as it would imply that the process equal to t is a martingale. Itô's lemma provides explicit formulae to use for integration. Levy's theorem (see Revuz–Yor [307]) states that, if X is a continuous process such that X and $(X_t^2 - t, t \geq 0)$ are martingales, then X is a Brownian motion.

3.1.2 Itô Processes. Girsanov's Theorem

Definition 3.1.6. *The process $X = (X_t,\ t \in [0,T])$ is a real-valued Itô process, if there exists an adapted process $\mu(t)$ and a predictable process $\sigma(t)$ satisfying*

$$\int_0^T |\mu(s)|\,\mathrm{d}s \ < \ \infty \quad P\text{-}a.s.\,;\qquad \int_0^T \sigma^2(s)\,\mathrm{d}s \ < \ \infty \quad P\text{-}a.s.\,,$$

and such that $X_t = X_0 + \int_0^t \mu(s)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}B(s)$, $t \in [0,T]$, where $(B_t,\ t \geq 0)$ is a real-valued Brownian motion.

The equality above is written more concisely as follows,

$$\begin{aligned}
\mathrm{d}X_t &= \mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B_t \\
X(0) &= X_0\,,
\end{aligned} \tag{3.1}$$

in order to develop a formal calculus that is analogous to differential calculus. The coefficient $\mu(t)$ is called the drift; σ is the diffusion coefficient (see Appendix A for an intuitive approach).

The drift term often complicates calculations; moreover, it causes the process to loose its martingale property: an Itô process is only a martingale when $\mu \equiv 0$. To revert to this case, we use Girsanov's theorem. Under reasonable assumptions, a change of probability measure transforms an Itô process into a stochastic integral. Let us state the theorem – without proof – in a special case (refer to the annex for a more general statement of the theorem, and additional comments).

Theorem 3.1.7 (Girsanov's Theorem). *Let $(L_t,\ t \geq 0)$ be the process defined by*

$$L_t \ = \ \exp\left\{ \int_0^t h(s)\,\mathrm{d}B_s - \frac{1}{2}\int_0^t h^2(s)\,\mathrm{d}s \right\},$$

where $(h(s),\ 0 \leq s \leq T)$ is an adapted bounded process.
The process $(L_t,\ t \geq 0)$ is the unique solution[4] to

$$\mathrm{d}L_t = L_t h_t\,\mathrm{d}B_t, \qquad L_0 = 1\,,$$

---

[4] Cf. the annex to Chap. 4.

and satisfies $E(L_t) = 1$, $\forall t \in [0, T]$.

The process $(L_t \ , \ t \geq 0)$ is a martingale.

Let Q be the probability measure defined on $(\Omega, \mathcal{F}_T)$ by $Q(A) = E_P(1_A L_T)$. Under Q, the process B^* defined by $B_t^* = B_t - \int_0^t h(s)\mathrm{d}s$ is a Brownian motion.

In particular, if $h(t) = -\mu(t)\sigma^{-1}(t)$ is bounded, the process $B_t^* = B_t + \int_0^t \mu(s)\,\sigma^{-1}(s)\,\mathrm{d}s$ is a Q-Brownian motion, and the Itô process $(X_t \ , \ t \geq 0)$ defined by (3.1) can be written

$$\begin{aligned} \mathrm{d}X_t &= \mu(t)\,\mathrm{d}t + \sigma(t)\left\{\mathrm{d}B_t^* - \mu(t)\sigma^{-1}(t)\,\mathrm{d}t\right\} \\ &= \sigma(t)\,\mathrm{d}B_t^* \ . \end{aligned}$$

Thus X_t is a stochastic integral with respect to a Brownian motion under Q. It is a Q-martingale if

$$E_Q\left\{\int_0^T \sigma^2(s)\,\mathrm{d}s\right\} = E_P\left\{L_T\int_0^T \sigma^2(s)\,\mathrm{d}s\right\}$$

$$= E_P\left\{\int_0^T L_s\sigma^2(s)\,\mathrm{d}s\right\} < \infty \ .$$

It is useful to note that the Girsanov transformation alters the drift but leaves the diffusion coefficient unchanged.

3.1.3 Itô's Lemma

If we assume that all asset prices follow Itô processes, then we will need to evaluate expressions of the form $f(t, X_t)$, and to specify their dynamics. The method for doing this is given by Itô's lemma.

Denote by $C_b^{1,2}([0, T] \times \mathbb{R} \,, \ \mathbb{R})$ the set of functions $f(t, x)$ that are continuous, of class C^1 with respect to t and C^2 with respect to x, and with bounded derivatives.

Lemma 3.1.8 (Itô's lemma[5]).
Let $f \in C_b^{1,2}([0, T] \times \mathbb{R} \,, \ \mathbb{R})$ and let X be an Itô process:

$$\mathrm{d}X_t = \mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B(t) \ .$$

Let $Y_t = f(t, X_t)$. Then Y is an Itô process that satisfies

$$\begin{aligned} \mathrm{d}Y_t &= \frac{\partial f}{\partial t}(t, X_t)\mathrm{d}t + \frac{\partial f}{\partial x}(t, X_t)\,\mu(t)\mathrm{d}t \\ &\quad + \frac{\partial f}{\partial x}(t, X_t)\,\sigma(t)\mathrm{d}B(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t)\,\sigma^2(t)\mathrm{d}t \end{aligned} \tag{3.2}$$

---

[5] See Appendix A for an intuitive approach.

or, in the more concise form,

$$dY_t = \frac{\partial f}{\partial t}\, dt + \frac{\partial f}{\partial x}\, dX(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\, \sigma^2(t)dt\,.$$

The boundedness condition on the derivatives is required for the integrals to make sense. It can be dropped when the integrals are known to exist. The stochastic integral will only be a martingale if the derivative of f with respect to the space variables, is sufficiently integrable. Whilst taking a vigilant approach to the validity of the results, we do not give an explicit set of conditions, so as not to weigh down our exposition.

Remark 3.1.9. The formula above differs from that for the differentiation of a composition of functions, which would lead to $\frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t$, by the addition of the term $\frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2(t)dt$. However, the formula is still very straightforward to use.

Here are two examples:

Example 3.1.10. Let

$$dX_t = aX_t\, dt + bX_t\, dB(t)\,,$$
$$X_0 > 0\,,$$

where a and b are *constants*.

It is possible to show that X_t takes strictly positive values. This enables us to define $Y_t = \ln X_t$ where $\ln$ denotes the natural logarithm. The function $\ln x$ does not belong to $C_b^{1,2}$, but Itô's formula remains valid nevertheless; the stochastic integral involved is well-defined, as are the ordinary integrals. Thus we have:

$$dY_t = 0 + \frac{1}{X_t}aX_t\, dt + \frac{1}{X_t}bX_t\, dB(t) - \frac{1}{2}\frac{1}{X_t^2}(bX_t)^2\, dt$$
$$= \left(a - \frac{1}{2}b^2\right)dt + b\, dB(t)\,,$$

which means that (see (3.1))

$$Y(t) = Y_0 + \int_0^t \left(a - \frac{1}{2}b^2\right)ds + \int_0^t b\, dB(s)$$
$$= \ln X_0 + \left(a - \frac{1}{2}b^2\right)t + b\, B(t)\,.$$

The variable Y_t is normally distributed with mean $\ln X_0 + \left(a - \frac{1}{2}b^2\right)t$ and variance $b^2 t$. Hence the name "lognormal" (whose logarithm is normally distributed) for the process X, which is also called the geometric Brownian motion.

Example 3.1.11. Let $X_t = x + \int_0^t b(s)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}B_s$ and $Y_t = \exp X_t$. To calculate $\mathrm{d}Y_t$, set $f(x) = \mathrm{e}^x$. We have $f'(x) = f''(x) = \mathrm{e}^x$. In addition, $\mathrm{d}X_t = b(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B_t$. Hence

$$\mathrm{d}Y_t = Y_t b(t)\,\mathrm{d}t + Y_t \sigma(t)\,\mathrm{d}B_t + 1/2 Y_t \sigma^2(t)\,\mathrm{d}t$$
$$= Y_t \left(\mathrm{d}X_t + 1/2\sigma^2(t)\mathrm{d}t\right)\,,$$

which can be expressed as $\mathrm{d}\left(\exp X_t\right) = \left(\exp X_t\right)\left(\mathrm{d}X_t + \dfrac{1}{2}\sigma^2(t)\mathrm{d}t\right).$

Exercise 3.1.12. Let B be a real-valued Brownian motion.

Using Itô's lemma, and assuming the stochastic integral to be well-defined, show that

$$B^2(t) - t = 2 \int_0^t B(s)\,\mathrm{d}B(s)\,.$$

Exercise 3.1.13. Let X be a process satisfying

$$\mathrm{d}X_t = X_t\left(\mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B_t\right)$$
$$X_0 > 0\,,$$

where $\sigma(t)$ is assumed to be bounded. Let $Y_t = X_t \exp - \int_0^t \mu(s)\mathrm{d}s$. Show that $\mathrm{d}Y_t = Y_t \sigma(t)\,\mathrm{d}B_t$.

Hence deduce that $X_t = E\left(X_T \exp - \int_t^T \mu(s)\mathrm{d}s \,|\mathcal{F}_t\right).$

3.1.4 Multidimensional Processes

We take a d-dimensional Itô process, that is a vector $S^* = (S^1, S^2, \ldots, S^d)^T$ driven by a k-dimensional Brownian motion $B = (B^1, B^2, \ldots, B^k)^T$:

$$\mathrm{d}S^i(t) = \mu^i(t)\,\mathrm{d}t + \sigma^i(t)\,\mathrm{d}B(t)\,, \tag{3.3}$$

where σ^i is the row vector $(\sigma^{i,1}, \sigma^{i,2}, \ldots, \sigma^{i,k})$ and where we use the matrix notation:

$$\int_0^t \sigma^i(s)\,\mathrm{d}B(s) = \sum_{j=1}^k \int_0^t \sigma^{i,j}(s)\,\mathrm{d}B^j(s)\,.$$

We can also write (3.3) in the form $\mathrm{d}S^*(t) = \mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B(t)$. We suppose that

> For all (i,j), μ^i is an adapted process
> and $\sigma^{i,j}$ is a predictable process, such that
> $$\int_0^T |\mu^i(t)|\,\mathrm{d}t < \infty \quad \text{a.s.;} \quad \int_0^T |\sigma^{i,j}(t)|^2\,\mathrm{d}t < \infty \quad \text{a.s.}\,. \tag{3.4}$$

3.1.5 Multidimensional Itô's Lemma

Let $X = (X^1, X^2, \ldots, X^d)^T$ be a d-dimensional Itô process following

$$\mathrm{d}X_t \;=\; \mu_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}B_t \,, \tag{3.5}$$

where μ_t is a $\mathcal{F}_t$-adapted process with values in $\mathbb{R}^d$, where σ_t is a random and predictable $(d \times k)$-matrix, and where B is a k-dimensional Brownian motion, with (μ, σ) satisfying (3.4). We write μ_t and σ_t instead of $\mu(t)$ and $\sigma(t)$ to lighten the notation[6].

Notation 3.1.14. If A is a square matrix $(A_{i,j})_{i,j}$, we use the notation $\mathrm{tr}\, A := \sum_{i=1}^n A_{i,i}$ for the sum of its diagonal terms.

There is a generalization of Itô's lemma to multidimensional processes:

Lemma 3.1.15. *Let $f \in C^{1,2}([0,T] \times \mathbb{R}^d \,; \mathbb{R})$. We write $f_x(t,x)$ for the row vector* $\left[\dfrac{\partial f}{\partial x_i}(t,x) \right]_{i=1,\ldots,d}$ *; $f_{xx}(t,x)$ for the matrix* $\left[\dfrac{\partial^2 f}{\partial x_i \, \partial x_j}(t,x) \right]_{i,j}$ *, and write $f_t(t,x) = \dfrac{\partial f}{\partial t}(t,x)$.*

Let $Y_t = f(t, X_t)$, where X_t satisfies (3.5). Then

$$\mathrm{d}Y_t \;=\; \left\{ f_t(t,X_t) + f_x(t,X_t)\mu_t + \frac{1}{2}\,\mathrm{tr}\,\left[\sigma_t \sigma_t^T f_{xx}(t,X_t) \right] \right\} \mathrm{d}t$$
$$+\, f_x(t,X_t)\sigma_t \, \mathrm{d}B_t \,. \tag{3.6}$$

We can give a more concise form to this formula by introducing the notation $\mathcal{L}$ for the operator defined on $C^{1,2}([0,T] \times \mathbb{R}^d, \mathbb{R})$, and depending on the coefficients μ and σ of the process for X:

$$\mathcal{L}f(t,x) \;=\; f_t(t,x) + f_x(t,x)\mu_t + \frac{1}{2}\,\mathrm{tr}\,\left[\sigma_t \sigma_t^T f_{xx}(t,x) \right] \,.$$

The operator $\mathcal{L}$ is called the infinitesimal generator of the diffusion. Using this notation, equality (3.6) becomes

$$\mathrm{d}Y_t \;=\; \mathcal{L}f(t,X_t)\mathrm{d}t + f_x(t,X_t)\sigma_t \, \mathrm{d}B_t \,.$$

We can also write

$$\mathrm{d}Y_t = f_t(t,X_t) + \sum_i f_{x_i}(t,X_t) \, \mathrm{d}X_t^i + \frac{1}{2} \sum_{i,j} f_{x_i x_j}(t,X_t) \, \mathrm{d}X_t^i \mathrm{d}X_t^j$$

where the product $\mathrm{d}X^i \mathrm{d}X^j$ is evaluated using the convention $\mathrm{d}t\,\mathrm{d}t = 0$, $\mathrm{d}t\,\mathrm{d}B_t^i = 0$ and $\mathrm{d}B_t^i \, \mathrm{d}B_t^j = \delta_{i,j}\,\mathrm{d}t$, where $\delta_{i,j}$ is the Kronecker delta, which is worth 1 when $i = j$ and 0 otherwise.

---

[6] Taking care not to confusion them with partial derivatives.

3.1.6 Examples

a) Let $(B_t \ , \ t \geq 0)$ be a real-valued Brownian motion and let X be a 2-dimensional Itô process, defined by $X = (X^1, X^2)^T$ where

$$\begin{cases} dX_t^1 & = \ \mu_t^1 \, dt + \sigma_t^1 \, dB_t \\ dX_t^2 & = \ \mu_t^2 \, dt + \sigma_t^2 \, dB_t \ . \end{cases}$$

We write this process as

$$dX_t \ = \ \mu_t \, dt + \sigma_t \, dB_t \quad \text{with} \quad \mu_t = \begin{pmatrix} \mu_t^1 \\ \mu_t^2 \end{pmatrix} \quad \text{and} \quad \sigma_t = \begin{pmatrix} \sigma_t^1 \\ \sigma_t^2 \end{pmatrix}.$$

Let $Y_t = X_t^1 X_t^2$. We apply Itô's formula with $f(x_1, x_2) = x_1 x_2$. We have

$$\sigma_t \sigma_t^T \ = \ \begin{bmatrix} |\sigma_t^1|^2 & \sigma_t^1 \sigma_t^2 \\ \sigma_t^1 \sigma_t^2 & |\sigma_t^2|^2 \end{bmatrix} \quad \text{and} \quad f_{xx}(t, x) \ = \ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Hence $\operatorname{tr}\left(\sigma_t \sigma_t^T f_{xx}\right) = 2\sigma_t^1 \sigma_t^2$ and

$$dY_t \ = \ \left(X_t^1 \mu_t^2 + X_t^2 \mu_t^1 + \sigma_t^1 \sigma_t^2\right) \, dt + \left(X_t^1 \sigma_t^2 + X_t^2 \sigma_t^1\right) \, dB_t \ ,$$

i.e.,

$$d\left(X_t^1 X_t^2\right) \ = \ X_t^1 \, dX_t^2 + X_t^2 \, dX_t^1 + \sigma_t^1 \sigma_t^2 \, dt \ .$$

This rule is known as integration by parts. It reads:

$$\int_0^t X_s^1 \, dX_s^2 \ = \ X_t^1 X_t^2 - X_0^1 X_0^2 - \int_0^t X_s^2 \, dX_s^1 - \int_0^t \sigma_s^1 \sigma_s^2 \, ds \ .$$

b) Let $B = (B^1, B^2)^T$ be a 2-dimensional Brownian motion and let $dX_t = \mu_t \, dt + \Sigma_t \, dB_t$ be a 2-dimensional Itô process where

$$\mu_t = \begin{pmatrix} \mu_t^1 \\ \mu_t^2 \end{pmatrix} \qquad \Sigma_t = \begin{bmatrix} \sigma_t^{1,1} & \sigma_t^{1,2} \\ \sigma_t^{2,1} & \sigma_t^{2,2} \end{bmatrix} \qquad \text{and} \quad X_t = \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix}.$$

Using the expanded form:

$$dX_t^i = \mu_t^i dt + \sigma_t^{i,1} dB_t^1 + \sigma_t^{i,2} dB_t^2 \ .$$

We write σ_t^1 for the vector $\begin{pmatrix} \sigma_t^{1,1} \\ \sigma_t^{2,1} \end{pmatrix}$ and σ_t^2 for the vector $\begin{pmatrix} \sigma_t^{1,2} \\ \sigma_t^{2,2} \end{pmatrix}$. We get

$$\Sigma_t \Sigma_t^T \ = \ \begin{bmatrix} \|\sigma_t^1\|^2 & \sigma_t^1 \cdot \sigma_t^2 \\ \sigma_t^1 \cdot \sigma_t^2 & \|\sigma_t^2\|^2 \end{bmatrix}$$

where $\sigma^1 \cdot \sigma^2$ denotes the scalar product of the two vectors.

Let $Y_t = X_t^1 X_t^2$. Itô's formula leads to

$$\begin{aligned}
\mathrm{d}Y_t &= X_t^1 \,\mathrm{d}X_t^2 + X_t^2 \,\mathrm{d}X_t^1 + \sigma_t^1 \cdot \sigma_t^2 \,\mathrm{d}t \\
&= X_t^1 \,\mathrm{d}X_t^2 + X_t^2 \,\mathrm{d}X_t^1 + \left(\sigma_t^{1,1}\sigma_t^{1,2} + \sigma_t^{2,1}\sigma_t^{2,2} \right) \mathrm{d}t .
\end{aligned}$$

This rule of calculation is easy to apply: we write formally $\mathrm{d}Y_t = X_t^1 \,\mathrm{d}X_t^2 + X_t^2 \,\mathrm{d}X_t^1 + \mathrm{d}X_t^1\mathrm{d}X_t^2$ and evaluate the product $\mathrm{d}X_t^1\mathrm{d}X_t^2$ using the following formal rules

$$\mathrm{d}t\,\mathrm{d}t = 0 ; \quad \mathrm{d}B_t^1 \,\mathrm{d}B_t^2 = 0 , \quad \mathrm{d}B_t^i \,\mathrm{d}B_t^i = \mathrm{d}t .$$

Exercise 3.1.16. Let $\mathrm{d}S^i(t) = S^i(t)(r\mathrm{d}t + \sigma_i \mathrm{d}B^i(t))$ be two Itô processes, with B^1 and B^2 independent Brownian motions. Set $S(t) = \sqrt{S_t^1 S_t^2}$.

Show that $\mathrm{d}S(t) = S(t)\left(r - \frac{1}{8}(\sigma_1^2 + \sigma_2^2)\right)\mathrm{d}t + \dfrac{1}{2}\left(\sigma_1\,\mathrm{d}B_t^1 + \sigma_2\,\mathrm{d}B_t^2\right)$. Furthermore, show that there exists a Brownian motion B^3 such that

$$\mathrm{d}S(t) = S(t)(r - \frac{1}{8}\left(\sigma_1^2 + \sigma_2^2\right))\mathrm{d}t + \frac{1}{2}\sqrt{\sigma_1^2 + \sigma_2^2}\,\mathrm{d}B_t^3 .$$

Exercise 3.1.17. Suppose

$$\begin{cases} \mathrm{d}X_t = m(\mu - X_t)\mathrm{d}t + \gamma\mathrm{d}B_t^1 \\ \mathrm{d}Y_t = a(\alpha - Y_t)\mathrm{d}t + c\sqrt{Y_t}\,\mathrm{d}B_t^2 , \end{cases}$$

where B^1 and B^2 are independent Brownian motions. Let $f \in C^{1,2,2}(\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R})$. Apply d'Itô's formula to $f(t,x,y)$. Show that the infinitesimal generator of the pair (X,Y) is

$$\mathcal{L}f(t,x,y) = \frac{1}{2}\gamma^2 f_{xx}'' + \frac{1}{2}c^2 y f_{yy}'' + m(\mu - x)f_x' + a(\alpha - y)f_y' + f_t' .$$

Exercise 3.1.18. Suppose

$$\begin{cases} \mathrm{d}X_t = X_t(\mu_X \mathrm{d}t + \sigma_X \mathrm{d}B_t) \\ \mathrm{d}Y_t = Y_t(\mu_Y \mathrm{d}t + \sigma_Y \mathrm{d}B_t) , \end{cases}$$

and let $Z = \frac{X}{Y}$. Show that $\mathrm{d}Z_t = Z_t(\mu_Z \mathrm{d}t + \sigma_Z \mathrm{d}B_t)$ with $\mu_Z = \mu_X - \mu_Y + \sigma_Y(\sigma_Y - \sigma_X)^T$ and $\sigma_Z = \sigma_X - \sigma_Y$.

3.2 Arbitrage and Valuation

3.2.1 Financing Strategies

Let Z be a $\mathcal{F}_T$-measurable random variable. Can we attain Z with a financing strategy? Let us specify our vocabulary.

We assume that there are d risky assets, whose prices $S^i, i \in \{1,\ldots,d\}$ are assumed to follow the Itô processes

$$\mathrm{d}S^i(t) \;=\; \mu^i(t)\,\mathrm{d}t + \sigma^i(t)\,\mathrm{d}B(t)\,,$$

where the coefficients μ and σ satisfy conditions (3.4). Asset 0 is a riskless asset

$$\mathrm{d}S_t^0 \;=\; S_t^0\, r(t)\,\mathrm{d}t\,,\qquad S^0(0)=1\,,$$

where $r(t)$ is a positive adapted process satisfying $\int_0^T r(t)\,\mathrm{d}t < \infty$ a.s.. Asset 0 is said to be riskless, even when $r(t)$ follows a stochastic process, as S_t^0 is known once r is known: $S_t^0 = \exp \int_0^t r(s)\mathrm{d}s$. On the other hand, S_t^j is not known explicitly, even when the coefficients μ and σ are, because the Brownian motion is a source of randomness.

We write S_t for the price vector of the $(d+1)$ assets

$$S_t \;=\; (S_t^0, S_t^1, \ldots, S_t^d)^T\,,$$

and $S_t^* = (S_t^1, \ldots, S_t^d)^T$ for the price vector of risky assets.

If $\theta = (\theta^0, \ldots, \theta^d)$ represents the number of stocks held of each type, that is to say the portfolio, then the wealth at time T is given by $x + \int_0^T \theta(s)\mathrm{d}S(s)$, where x is the initial wealth and where

$$
\int_0^T \theta(s)\,\mathrm{d}S(s) \;=\; \sum_{i=0}^d \int_0^T \theta^i(s)\,\mathrm{d}S^i(s)
$$

$$
= \int_0^T \theta^0(s)S^0(s)r(s)\mathrm{d}s + \sum_{i=1}^d \left[\int_0^T \theta^i(s)\mu^i(s)\,\mathrm{d}s + \int_0^T \theta^i(s)\sigma^i(s)\,\mathrm{d}B(s) \right].
$$

In order for the stochastic integrals involved to make sense, we need to impose integrability conditions on θ, μ and σ, in addition to the measurability conditions:

Definition 3.2.1. *For $i \geq 1$, denote by $\Theta(S^i)$ the set of processes θ^i that are predictable, and satisfy*

$$
\int_0^T |\theta^i(s)\,\mu^i(s)|\,\mathrm{d}s \;<\; \infty \qquad a.s.\,,\qquad \int_0^T (\theta^i(s))^2\,\|\sigma^i(s)\|^2\,\mathrm{d}s \;<\; \infty \qquad a.s.
$$

where the norm of vector v is defined by $\|v\|^2 = \sum_{j=1}^k v_j^2$.

Denote by $\Theta(S^0)$ the set of processes θ^0 that are adapted, and satisfy $\int_0^T |\theta^0(s)|\, r(s)\, S_s^0\mathrm{d}s \;<\; \infty$.
We write $\theta \in \Theta(S)$ to express $\{\theta^j \in \Theta(S^j);\;\; 0 \leq j \leq d\}$.

Definition 3.2.2. $\theta \in \Theta(S)$ *finances Z if*

$$
\theta_t \cdot S_t \;=\; \theta_0 \cdot S_0 + \int_0^t \theta_s\,\mathrm{d}S_s \qquad \forall\, t \leq T\,,\quad \text{P-a.s.} \qquad\qquad (3.7.\mathrm{i})
$$

$$
\text{and}\qquad \theta_T \cdot S_T \;=\; Z \qquad\qquad\qquad\qquad \text{P-a.s.} \qquad\qquad (3.7.\mathrm{ii})
$$

Condition (3.7.i) is referred to as the self-financing condition, and is the continuous time version of condition (2.1) in Chap. 2. Indeed, suppose that the prices S_t are constant on the interval $[n, n+1[$, and that θ is constant on $]n, n+1]$ [7]. Condition (3.7.i) can be written as

$$\theta_{n+1} \cdot S_{n+1} \; = \; \theta_n \cdot S_n + \int_n^{n+1} \theta_s \, \mathrm{d}S_s \; = \; \theta_n \cdot S_n + \theta_{n+1} \cdot (S_{n+1} - S_n) \ .$$

Hence $\theta_{n+1} \cdot S_n = \theta_n \cdot S_n$.

3.2.2 Arbitrage and the Martingale Measure

Definition 3.2.3. *An arbitrage opportunity is a strategy that satisfies* (3.7.i) *with*

$$P(\theta_T \cdot S_T \geq 0) \; = \; 1$$
$$P(\theta_T \cdot S_T > 0) \; > \; 0$$
$$P(\theta_0 \cdot S_0 = 0) \; = \; 1 \ .$$

In other words, $\theta_T \cdot S_T$ is a positive random variable that has strictly positive expectation.

We can replace the first two conditions by $E(\theta_T \cdot S_T) > 0$ and $P(\theta_T \cdot S_T \geq 0) = 1$.

We work under the assumption that there are no opportunities for arbitrage. This assumption is often replaced by the assumption:

(H) There exists a probability measure Q that is equivalent to P and such that the vector of discounted prices S_t/S_t^0 is a Q-martingale.

Definition 3.2.4. *An equivalent martingale measure Q (EMM for short) is a probability measure that is equivalent to P and such that, under Q, discounted prices are martingales. Such a measure is also called a risk-neutral measure.*

Let us study the link between hypothesis (H) and the assumption of NAO.

Assume that (H) holds. Suppose that we are in the case where the riskless asset has a constant price $S_t^0 = 1$. Assume that the risky asset follows, under Q, the stochastic differential equation

$$\mathrm{d}S_t \; = \; S_t \sigma(t) \, \mathrm{d}B_t \ .$$

In   this   case,   there   is   no   arbitrage   opportunity   such   that $E_Q \int_0^T \|\theta(s)\sigma(s)S_s\|^2 \mathrm{d}s < \infty$: under this integrability condition, $\int_0^t \theta_s \, \mathrm{d}S_s$ is a Q-martingale with zero expectation. The self-financing condition is satisfied

---

[7] The asymmetry is due to the predictable nature of the portfolio.

under Q (P and Q are equivalent). Hence taking expectations under Q in (3.7.i), it follows that

$$E_Q(\theta_T \cdot S_T) \;=\; \theta_0 \cdot S_0 + E_Q\left[\int_0^T \theta_s \, dS_s^1\right] \;=\; \theta_0 \cdot S_0 \;.$$

As $\theta_0 \cdot S_0$ is $\mathcal{F}_0$-measurable and hence constant, it follows that $\theta_0 \cdot S_0$ cannot be zero when $E_Q(\theta_T \cdot S_T) > 0$, which would be the case if we had $E_P(\theta_T \cdot S_T) > 0$.

In the most general case, hypothesis (H) does not imply NAO. Suppose now that we are in the case of a single risky asset, whose price satisfies $S_t = B_t$, and of a bond with a constant price equal to 1. Dudley [109] showed that for any positive and $\mathcal{F}_T$-measurable random variable Y, there exists a predictable process θ^1 such that $\int_0^T \theta_s^1 \, dB_s = Y$. We can then construct θ^0 such that the strategy (θ^0, θ^1) is self-financing and has zero initial value: it is enough to take

$$\theta_t^0 \;=\; \int_0^t \theta_s^1 \, \mathrm{d}B_s - \theta_t^1 B_t \;.$$

The strategy (θ^0, θ^1) is then an arbitrage opportunity.

Stricker [348] showed that under (H), there does not exist an elementary strategy that constitutes an arbitrage opportunity. An elementary strategy is such that there exists a sequence of real numbers $t_0 < t_1 < \cdots < t_p$ with

$$\theta_s \;=\; \sum_{i=0}^{p-1} 1_{]t_i, t_{i+1}]}(s) \, \psi_i$$

for ψ_i a $\mathcal{F}_{t_i}$-measurable variable.

Another way to avoid arbitrage opportunities under hypothesis (H) is to limit ourselves to strategies θ such that $M_t = \int_0^t \theta(s) \, dS_s$ is bounded below[8]. In this case, $(M_t \,,\; t \geq 0)$ is a local[9] martingale bounded below, and thus a supermartingale under Q, and satisfies $E_Q(M_t) \leq 0$, $t \in [0, T]$. This implies NAO, for if θ were a strategy with zero initial wealth, we would have

$$E_Q(\theta_t \cdot S_t) \;=\; E_Q\left(\int_0^t \theta_s \, \mathrm{d}S_s\right) \;\leq\; 0 \;.$$

The study of the converse is even more delicate. Thus, we cannot show (H) and NAO to be equivalent in all generality. Specific references are given in the notes.

---

[8] Dybrig and Huang [132].
[9] The definition is to be found in the annex.

3.2.3 Valuation

As in the previous chapters, assuming that there is NAO, we will show that, if θ and φ are two strategies that finance Z, then $\theta_0 \cdot S_0 = \varphi_0 \cdot S_0$.

Definition 3.2.5. *Under NAO, if θ is a strategy that finances Z, then $\theta_0 \cdot S_0$ is the arbitrage price of Z and $\pi(Z)_t := \theta_t \cdot S_t$ is the "implicit price" of Z at time t.*

Proposition 3.2.6. *Under NAO, the arbitrage price is well-defined.*

Proof. Let θ and φ be two self-financing strategies such that $\theta_T \cdot S_T = \varphi_T \cdot S_T$. Let us show that $\theta_0 \cdot S_0 = \varphi_0 \cdot S_0$ under the assumption of NAO. Suppose that $\theta_0 \cdot S_0 > \varphi_0 \cdot S_0$. Intuitively, we would buy a portfolio φ at price $\varphi_0 \cdot S_0$, sell θ at price $\theta_0 \cdot S_0$ and invest the difference between the two, which is positive, in the riskless asset. At time T, we would sell φ and buy θ at the same price, and make a profit from the riskless investment. More formally, let

$$\psi_t := \varphi_t - \theta_t = (\varphi_t^i - \theta_t^i)_{0 \le i \le d} ,$$

and let ψ^*, which corresponds to the risky assets, be such that

$$\psi_t^* = (\varphi_t^i - \theta_t^i)_{1 \le i \le d} .$$

We can construct γ_t^0 such that $\gamma = (\gamma^0, \psi^*)$ is a self-financing strategy with zero initial value: it is enough to take $\gamma^0(0)$ such that

$$\gamma^0(0)S^0(0) + \psi^*(0) \cdot S^*(0) = 0 ,$$

and to construct $\gamma^0(t)$ as a solution to

$$\gamma^0(t)S^0(t) + \psi^*(t) \cdot S^*(t) = \int_0^t \gamma^0(s)\, dS^0(s) + \int_0^t \psi^*(s)\, dS^*(s) .$$

This equation can be solved as a differential equation with a perturbation function, by noticing that $\int_0^t \gamma^0(s)\, dS^0(s) = \int_0^t \gamma^0(s)r(s)S^0(s)\, ds$ is an ordinary integral. From the expressions

$$\gamma^0(0)S^0(0) + \psi^*(0) \cdot S^*(0) = 0$$

and

$$\psi^0(0)S^0(0) + \psi^*(0)S^*(0) < 0 ,$$

we obtain $\gamma^0(0) - \psi^0(0) > 0$.

The strategy $\gamma - \psi$ is self-financing (as it is the difference of self-financing strategies), its risky component is zero, and $(\gamma - \psi)_T \cdot S_T = (\gamma^0 - \psi^0)(0)S_T^0$. Hence $\gamma_T \cdot S_T = (\gamma - \psi)_T \cdot S_T + \psi_T \cdot S_T = (\gamma - \psi)_T \cdot S_T > 0$, and γ would be an arbitrage opportunity. $\square$

We have not studied the market's completeness: can any random variable Z be financed? In broad terms, the market is complete when $k = d$ and if the matrix σ is invertible. We will come back to this topic in the next chapter, and in Chap. 9.

3.3 The Black–Scholes Formula: the One-Dimensional Case

3.3.1 The Model

Consider a first security, to be called a bond, with the price process $(S_t^0, t \geq 0)$ defined by the differential equation

$$\begin{cases} dS_t^0 = rS_t^0 \, dt \, , & r \text{ positive and constant} \\ S_0^0 = 1 \, , \end{cases}$$

and a second security, whose price $(S_t^1, t \geq 0)$ satisfies the stochastic differential equation

$$\begin{cases} dS_t^1 = \mu S_t^1 \, dt + \sigma S_t^1 \, dB_t \\ S_0^1 > 0 \, , \end{cases}$$

where μ and σ are two *constants*, σ is non-zero, and $(B_t, t \geq 0)$ is a real-valued Brownian motion.

In Sect. 3.2.3 we described how the existence of a probability measure, under which the discounted price $S_t^{1,a} := e^{-rt} S_t^1$ is a martingale, is related to the absence of arbitrage opportunities. Let us show that such a probability measure exists in the model we have here.

A direct application of Itô's lemma shows that

$$dS_t^{1,a} = S_t^{1,a} \left[(\mu - r)dt + \sigma \, dB_t \right] \, .$$

Girsanov's theorem will enable us to transform $(S_t^{1,a}, t \geq 0)$ into a martingale.

Let $(L_t, t \geq 0)$ be the process satisfying $dL_t = -(\mu - r)\sigma^{-1} L_t dB_t$, $L_0 = 1$.

Girsanov's theorem shows that under the probability measure Q defined on $\mathcal{F}_T$ by $\dfrac{dQ}{dP} = L_T$, the process $(S_t^{1,a}, t \geq 0)$ satisfies $dS_t^{1,a} = S_t^{1,a} \sigma \, dB_t^*$, where $B_t^* := B_t + (\mu - r)\sigma^{-1}t$ is a Q-Brownian motion. The probability measure Q is equivalent to P.

Using the first part of Girsanov's theorem, which ensures that the solution to $dX_t = X_t h \, dB_t$ is a martingale when h is bounded, we find that $S^{1,a}$ is a Q-martingale. The discounted price of the riskless asset is constant and equal to 1, so it is also a Q-martingale!

Note that $S_t^{1,a} = \exp \left[\sigma B_t^* - \dfrac{1}{2}\sigma^2 t \right]$, which entails that $S_t^{1,a}$ (and hence S_t^1) is positive.

Under Q, the risky asset's price satisfies

$$\mathrm{d}S_t^1 \;=\; S_t^1 (r\,\mathrm{d}t + \sigma\,\mathrm{d}B_t^*) \,.$$

We call Q the risk-neutral measure (the measure that is neutral with respect to risk), as under Q the two assets have the same expected rate of return r.

Theorem 3.3.1. *In a model with two assets*

$$\mathrm{d}S_t^0 \;=\; rS_t^0 \,\mathrm{d}t \,, \qquad \mathrm{d}S_t^1 \;=\; \mu S_t^1 \,\mathrm{d}t + \sigma S_t^1 \,\mathrm{d}B_t \,,$$

there exists a probability measure Q that is equivalent to P, and such that discounted prices are martingales under Q.

We now take a random variable $Z = g\left(S_T^1\right)$ where g is a non-negative function, and we seek to obtain its implicit price. In the case where $g(x) = (x - K)^+ = \max(x - K\,,\,0)$, we are looking at a European option with strike price K.

Definition 3.3.2. *A European call option (a European option to buy) is a contract that gives the right (but not the obligation) to buy at time T (the maturity) a stock at price K (the strike or exercise price), which is fixed when the contract is signed.*

If $S_T^1 \geq K$, the option enables its owner to buy the asset at price K and then sell it immediately at price S_T^1 : the difference $S_T^1 - K$ between the two prices is the realized gain. If $S_T^1 < K$, the gain is zero.

3.3.2 The Black–Scholes Formula

Let $C^{1,2}([0,T] \times \mathbb{R}_+\,,\,\mathbb{R})$ be the set of functions f from $[0,T] \times \mathbb{R}_+$ into $\mathbb{R}$, and of class C^1 with respect to t and C^2 with respect to x.

We suppose that there exists $C \in C^{1,2}([0,T] \times \mathbb{R}_+\,,\,\mathbb{R})$ such that

$$\begin{aligned}
\pi(Z)_t &= C(t, S_t^1)\,, & t &< T \\
g(x) &= C(T, x)\,, & x &\in \mathbb{R}_+ \,.
\end{aligned}$$

Let $Y_t = C(t, S_t^1)$. From Itô's lemma:

$$\mathrm{d}Y_t \;=\; \left(\mu\, S_t^1\, C_x(t, S_t^1) + C_t(t, S_t^1) + \frac{1}{2}\sigma^2 (S_t^1)^2\, C_{xx}(t, S_t^1) \right) \mathrm{d}t$$
$$+ \sigma S_t^1\, C_x(t, S_t^1)\mathrm{d}B_t \quad (3.8)$$

where

$$C_t \;=\; \frac{\partial C}{\partial t}\,, \quad C_x \;=\; \frac{\partial C}{\partial x}\,, \quad C_{xx} \;=\; \frac{\partial^2 C}{\partial x^2}\,.$$

Using the notation $\mathcal{L}$ for the infinitesimal generator of the diffusion $(S_t, t \geq 0)$, defined on $C^{1,2}([0,T] \times \mathbb{R}_+\,,\,\mathbb{R})$ by

$$\mathcal{L}C \;=\; \mu x C_x + C_t + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 C}{\partial x^2} \;,$$

we can rewrite (3.8) as

$$\mathrm{d}Y_t \;=\; \mathcal{L}C(t, S_t^1)\mathrm{d}t + \sigma S_t^1 C_x(t, S_t^1)\,\mathrm{d}B_t \;. \tag{3.9}$$

Suppose that there exists a strategy θ that finances Z. We have not yet shown that such a strategy exists, but we will construct one shortly. The strategy θ is represented by a pair (α, β) where α (respectively β) is the number of bonds (respectively risky assets) held.

Thus, if $S_t = (S_t^0, S_t^1)$:

$$
\begin{aligned}
\theta_t \cdot S_t \;&=\; \alpha_t S_t^0 + \beta_t S_t^1 \\
&=\; \theta_0 \cdot S_0 + \int_0^t \theta_s\,\mathrm{d}S_s \\
&=\; \alpha_0 S_0^0 + \beta_0 S_0^1 + \int_0^t \alpha_s\,\mathrm{d}S_s^0 + \int_0^t \beta_s\,\mathrm{d}S_s^1 \\
&=\; C(t, S_t^1) =\; Y_t \;,
\end{aligned}
$$

hence a new expression for $\mathrm{d}Y_t$ is:

$$
\begin{aligned}
\mathrm{d}Y_t \;&=\; \alpha_t\,\mathrm{d}S_t^0 + \beta_t\,\mathrm{d}S_t^1 \\
&=\; r\alpha_t S_t^0\,\mathrm{d}t + \beta_t\left(\mu S_t^1\,\mathrm{d}t + \sigma S_t^1\,\mathrm{d}B_t\right) \\
&=\; \left(r\alpha_t S_t^0 + \mu\beta_t S_t^1\right)\,\mathrm{d}t + \sigma\beta_t S_t^1\,\mathrm{d}B_t \;. \tag{3.10}
\end{aligned}
$$

Comparing (3.9) and (3.10), and identifying $\mathrm{d}t$ terms, we obtain

$$\mathcal{L}C(t, S_t^1) \;=\; r\alpha_t S_t^0 + \mu\beta_t S_t^1 \;. \tag{3.11}$$

Identifying the coefficients of $\mathrm{d}B_t$, we have

$$\sigma S_t^1 C_x(t, S_t^1) \;=\; \sigma\beta_t S_t^1 \;, \tag{3.12}$$

and we still have

$$\alpha_t S_t^0 + \beta_t S_t^1 \;=\; C(t, S_t^1) \;. \tag{3.13}$$

From (3.12) we draw

$$\beta_t \;=\; C_x(t, S_t^1) \;,$$

which we substitute into (3.13) to get

$$\alpha_t \;=\; \left\{C(t, S_t^1) - S_t^1 C_x(t, S_t^1)\right\}\left(S_t^0\right)^{-1} \;.$$

Thus we have obtained a financing strategy for Z as a function of its implicit price. By substitution into (3.11),

$$\mathcal{L}C(t, S_t^1) \ = \ r\left[C(t, S_t^1) - S_t^1 C_x(t, S_t^1)\right] + \mu C_x(t, S_t^1) S_t^1 \ .$$

After replacing $\mathcal{L}C$ with its full expression, and carrying out simplifications, this last equality can be written

$$r S_t^1 C_x(t, S_t^1) + C_t(t, S_t^1) + \frac{1}{2}\sigma^2 (S_t^1)^2 C_{xx}(t, S_t^1) \ = \ r C(t, S_t^1)$$

$$t \in [0, T] \ ; \ \text{P-a.s.} \quad (3.14)$$

with of course

$$C(T, S_T^1) \ = \ g(S_T^1) \qquad \text{a.s. .}$$

It is important to note that μ does not appear in (3.14).

We can easily show that, whilst (3.14) a priori holds for all t, and for almost every ω, it is also satisfied when we replace S_t^1 by x with $x > 0$, since the support of the law for S_t^1 is $[0, \infty]$.

Hence, we find that C satisfies the parabolic equation[10] :

$$\begin{cases} r x\, C_x(t, x) + C_t(t, x) + \dfrac{1}{2}\sigma^2 x^2 C_{xx}(t, x) \ = \ r\, C(t, x) \,, \\[2mm] \qquad\qquad\qquad x \in]0, \infty[\,, \ t \in]0, T[\,, \\[2mm] C(T, x) \ = \ g(x) \,, \quad x \in]0, \infty[\qquad \text{(boundary condition) .} \end{cases} \quad (3.15)$$

Let us summarize the results:

Theorem 3.3.3. *Let $(S_t^0, t \geq 0)$ be the price of a bond $\mathrm{d}S_t^0 = r S_t^0\,\mathrm{d}t$, and let $(S_t^1, t \geq 0)$ be the price of the risky asset satisfying $\mathrm{d}S_t^1 = \mu S_t^1\,\mathrm{d}t + \sigma S_t^1\,\mathrm{d}B_t$. Let $Z = g\left(S_T^1\right)$ be a positive random variable, with $\pi(Z)_t$ as its implicit price. We assume that there exists $C \in C^{1,2}([0, T] \times \mathbb{R}_+, \mathbb{R})$ such that*

$$\begin{aligned} \pi(Z)_t \ &= \ C(t, S_t^1) \,, & t &< T \\ g(x) \ &= \ C(T, x) \,, & x &\in \mathbb{R}_+ \ . \end{aligned}$$

Then C satisfies the parabolic equation

$$\begin{cases} r x\, C_x(t, x) + C_t(t, x) + \dfrac{1}{2}\sigma^2 x^2 C_{xx}(t, x) \ = \ r\, C(t, x) \,, \\[2mm] \qquad\qquad\qquad x \in]0, \infty[\,, \ t \in]0, T[\,, \\[2mm] C(T, x) \ = \ g(x) \quad , \quad x \in]0, \infty[\,. \end{cases} \quad (3.15)$$

A strategy θ that finances Z is given by $\theta = (\alpha, \beta)$ where

$$\begin{aligned} \alpha_t \ &= \ \left\{C(t, S_t^1) - S_t^1 C_x(t, S_t^1)\right\} \left(S_t^0\right)^{-1} \\ \beta_t \ &= \ C_x(t, S_t^1) \ . \end{aligned}$$

---

[10] The boundary condition is fixed at the final point in time, which is not the case with the classical types of parabolic equations.

In general, it is not straightforward to obtain an explicit solution to the parabolic equation above. Probabilistic calculations provide an expression for its solution.

Suppose x and t to be fixed, and $Z_s^{x,t}$ to be the process indexed by s, $(t \leq s \leq T)$, and defined by

$$Z_s^{x,t} = x + r \int_t^s Z_u^{x,t} \, \mathrm{d}u + \sigma \int_t^s Z_u^{x,t} \, \mathrm{d}B_u \, .$$

$Z_s^{x,t}$ is initialized at point x at time t: $Z_t^{x,t} = x$. After time t, it follows the same dynamics as $\mathrm{d}Z(u) = rZ_u \mathrm{d}u + \sigma Z_u \mathrm{d}B_u$.

We have the following result, known as the Feynman-Kac formula [11]:

Theorem 3.3.4. *For a positive-valued function $g \in C^2(\mathbb{R})$ such that g, g' and g'' are all Lipschitz[12], the function*

$$C(t,x) := E\left[\mathrm{e}^{-r(T-t)} \, g(Z_T^{x,t})\right] \tag{3.16}$$

is the unique Lipschitz solution to (3.15).

Remark 3.3.5. In the case of European option "pricing", $g(x) = (x-K)^+$ does not satisfy the theorem's regularity conditions at the point $x = K$. However the theorem's result can be extended to piecewise regular, Lipschitz functions.

A good exercise involves checking that the theorem implies that our choice of α and β is satisfactory, in that (α, β) is admissible (i.e., the pair belongs to $\Theta(S)$) and finances $g(S_T^1)$. We can show that the model is complete, that is to say, that there exists a financing strategy for any square integrable $\mathcal{F}_T$-measurable random variable Z. This result requires new probabilistic tools (the representation theorem) and will be discussed in the next chapter.

3.3.3 The Risk-Neutral Measure

We are going to give another valuation method, which involves working under the risk-neutral probability measure Q defined by $\mathrm{d}Q = L_T \mathrm{d}P$ (see Sect. 3.3.1).

We have seen that under Q, the risky asset's price satisfies

$$\mathrm{d}S_t^1 = S_t^1(r \, \mathrm{d}t + \sigma \, dB_t^*) \, .$$

If we construct a portfolio as in (3.13),

---

[11] The proof of this result rests on Itô's lemma and on the martingale properties of stochastic integrals. See Sect. 5 of the annex.

[12] A function g is Lipschitz on $\mathbb{R}$ if there exists $k > 0$ such that $|g(x)-g(y)| \leq k|x-y|$ for all x, y.

$$Y_t = \alpha_t S_t^0 + \beta_t S_t^1 ,$$

we can check that under Q

$$\mathrm{d}Y_t = \alpha_t S_t^0 r \, \mathrm{d}t + \beta_t S_t^1 (r \, \mathrm{d}t + \sigma \, \mathrm{d}B_t^*)$$
$$= Y_t r \, \mathrm{d}t + \mathrm{d}M_t ,$$

where $(M_t, \, t \geq 0)$ is defined by $\mathrm{d}M_t = \beta_t S_t^1 \sigma \, \mathrm{d}B_t^*$. Under integrability conditions, $(M_t, \, t \geq 0)$ is a martingale. This implies (using Itô's lemma) that $(\mathrm{e}^{-rt}Y_t, \, t \geq 0)$ is a martingale[13] (equal to M_t up to a constant). Hence we get (from the martingale property) :

$$\mathrm{e}^{-rt}Y_t = E_Q(\mathrm{e}^{-rT}Y_T \,|\, \mathcal{F}_t)$$

and, using the fact that the process $(S_t, \, t \geq 0)$ is Markov,

$$Y_t = C(t, S_t^1) = E_Q(\mathrm{e}^{-r(T-t)}g(S_T^1) \,|\, \mathcal{F}_t) .$$

Theorem 3.3.6. *The implicit price of $g\left(S_T^1\right)$ is given by*

$$E_Q \left[\mathrm{e}^{-r(T-t)} g\left(S_T^1\right) \,|\, \mathcal{F}_t \right]$$

where Q is the risk-neutral measure.

In particular, at time $t = 0$

$$C(0, x) = E_Q \left[\mathrm{e}^{-rT} g(S_T^1) \right] = E_P \left[\mathrm{e}^{-rT} g(Z_T) \right]$$

where $(Z_t, t \geq 0)$ satisfies $\mathrm{d}Z_t = Z_t (r \, \mathrm{d}t + \sigma \, \mathrm{d}B_t), \quad Z_0 = x$, and we recover formula (3.16).

For a general time t, elementary calculations involving conditional expectations lead to the same result as (3.16), using the fact that under Q

$$S_T^1 = S_t^1 \exp\left[r(T - t) + \sigma(B_T^* - B_t^*) - \frac{1}{2}\sigma^2(T - t) \right] .$$

We can easily recover the hedging portfolio, by applying Itô's lemma to $C(t, S_t^1)$. Indeed,

$$\mathrm{d}C(t, S_t^1) = \frac{\partial C}{\partial t}(t, S_t^1) \, \mathrm{d}t + \frac{\partial C}{\partial x}(t, S_t^1) \, \mathrm{d}S_t^1 + \frac{1}{2}(S_t^1 \sigma)^2 \frac{\partial^2 C}{\partial x^2}(t, S_t^1) \mathrm{d}t .$$

Using the fact that $\mathrm{e}^{-rt}C(t, S_t^1)$ is a martingale, and setting the coefficient of the $\mathrm{d}B_t$ term to zero, we recover partial differential equation (3.15). We can decompose $C(t, S_t^1)$ into

$$C(t, S_t^1) = \alpha_t S_t^0 + \beta_t S_t^1$$

where $\alpha_t = \{C(t, S_t^1) - S_t^1 C_x(t, S_t^1)\}(S_t^0)^{-1}$ and $\beta_t = C_x(t, S_t^1)$. It is enough to transfer the values of α and β into the expression for $\mathrm{d}C$ to check that α , β is indeed a self-financing hedging portfolio.

---

[13] This is a generalization of the results of Chap. 2.

3.3.4 Explicit Calculations

In our model, the coefficients r and σ being constants, we can take the calculations further, using equation (3.16).

We saw (Example (3.1.10)) that a process satisfying

$$dZ_s = rZ_s \, ds + \sigma Z_s \, dB_s$$

has a lognormal law. Here, if the initial point in time is t, then the logarithm of $Z_s^{x,t}$ is distributed according to $N\left[\log Z_t^{x,t} + \left(r - \frac{1}{2}\sigma^2\right)(s - t), \, \sigma^2(s - t)\right]$, or alternatively, $Z = \exp U$ where U is normally distributed.

Hence we calculate

$$E\left[e^{-r(T-t)} g(Z_T^{x,t})\right] = e^{-r(T-t)} E\left[g(Z_T^{x,t})\right]$$

$$= e^{-r(T-t)} \int_{-\infty}^{+\infty} g(e^u) f_{T-t}(u) \, du \, ,$$

where $f_{T-t}(u)$ is the probability density function of the normal distribution with mean

$$\log x + \left(r - \frac{1}{2}\sigma^2\right)(T - t)$$

and variance $\sigma^2(T - t)$.

When g has an explicit form, we can develop these calculations further. Let us take the case $g(x) = (x - K)^+$.

Let $\phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-u^2/2} \, du = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-x}^{+\infty} e^{-u^2/2} \, du$. Using the expression

$$\int_{-\infty}^{+\infty} g(e^u) f_{T-t}(u) \, du$$

$$= \int_{\{u > \ln K\}} e^u \, f_{T-t}(u) \, du - K \int_{\{u > \ln K\}} f_{T-t}(u) \, du \, ,$$

we obtain

$$C(t, x) = x \, \phi\left(\frac{1}{\sigma\sqrt{T-t}} \left\{\ln\left(\frac{x}{K}\right) + (T - t)\left(r + \frac{\sigma^2}{2}\right)\right\}\right)$$

$$- Ke^{-r(T-t)} \phi\left(\frac{1}{\sigma\sqrt{T-t}} \left\{\ln\left(\frac{x}{K}\right) + (T - t)\left(r - \frac{\sigma^2}{2}\right)\right\}\right) , \quad (3.17)$$

with $C(T, x) = (x - K)^+$. Hence we obtain the same formula as in Sect. 2.6.

Theorem 3.3.7 (The Black–Scholes Formula). *The price of a call is given by*

$$C(0, x) = x\phi(d_1) - K e^{-rT} \phi(d_2)$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T}} \left\{ \log\left(\frac{x}{K}\right) + T\left(r + \frac{\sigma^2}{2}\right) \right\}, \quad d_2 = d_1 - \sigma\sqrt{T}.$$

We also have

$$C(t, x) = x\phi(d_1(t)) - K e^{-r(T-t)} \phi(d_2(t))$$

where

$$d_1(t) = \frac{1}{\sigma\sqrt{T-t}} \left\{ \ln\left(\frac{x}{K}\right) + (T-t)\left(r + \frac{\sigma^2}{2}\right) \right\}, \quad d_2(t) = d_1(t) - \sigma\sqrt{T-t}.$$

Exercise 3.3.8. Rework the previous calculations as follows.

a. Write $S_T = S_t \exp Y$ where S_t and Y are independent. Determine the distribution of Y.

b. Show that if U follows a normal distribution with mean m and variance σ^2 under Q, then $Q(e^U > u) = \phi((m - \ln u)/\sigma)$ and $E_Q(e^U 1_{U>u}) = e^{m+\sigma^2/2}\phi(m + \sigma^2 - u)$.

c. Thence deduce the value of $C(0, x)$ and of $C(t, x)$.

These results can be generalized to the case where r, μ and σ are deterministic functions of t. Then we obtain:

$$C(t, x) = x\phi(d_1(t)) - K e^{-\int_t^T r(s)\mathrm{d}s} \phi(d_2(t))$$
$$C(T, x) = (x - K)^+$$

where

$$d_1(t) = \left\{ \ln\left(\frac{x}{K}\right) + \int_t^T \left(r(s) + \frac{\sigma^2(s)}{2}\right) \mathrm{d}s \right\} \left\{ \sqrt{\int_t^T \sigma^2(s)\mathrm{d}s} \right\}^{-1}$$

$$d_2(t) = \left\{ \ln\left(\frac{x}{K}\right) + \int_t^T \left(r(s) - \frac{\sigma^2(s)}{2}\right) \mathrm{d}s \right\} \left\{ \sqrt{\int_t^T \sigma^2(s)\mathrm{d}s} \right\}^{-1}.$$

Exercise 3.3.9. Establish the Black–Scholes formula for the price $P(t, x)$ of a put (the option to sell) where $g(x) = (K - x)^+$. Establish the so-called put–call parity formula $C(t, x) = P(t, x) + x - K e^{-r(T-t)}$.

Exercise 3.3.10. Assume that the coefficients r and σ are time-dependent. Write $R_T^t = \exp\left[-\int_t^T r(s)\mathrm{d}s\right]$ and $\Sigma^2(t,T) = \int_t^T \sigma^2(s)\mathrm{d}s$. Let Φ_{m,Σ^2} be the normal probability density function of mean $m = \ln(x/R_T^t) - \Sigma^2(t,T)/2$ and variance $\Sigma^2(t,T)$. We want to evaluate $g(S_T^1)$.

Show that $C(t,x) = R_T^t \int_{\mathbb{R}} g(e^y)\Phi_{m,\Sigma^2}(y)\,\mathrm{d}y$ and, by differentiating under the integral sign, that $\dfrac{\partial C}{\partial x}(t,x) = \displaystyle\int_{\mathbb{R}} g'(e^{y+\Sigma^2(t,T)/2})\Phi_{m,\Sigma^2}(y)\,\mathrm{d}y.$

3.3.5 Comments on the Black–Scholes Formula

We have established that the price of a call with strike K and maturity T is given, in the case of constant coefficients, by:

$$C(0,x) \;=\; x\,\phi(d_1) - K\,\mathrm{e}^{-rT}\,\phi(d_2) \tag{3.18}$$

where d_1 and d_2 depend on the parameters x, K, T, r and σ:

$$d_1 \;=\; \frac{1}{\sigma\sqrt{T}}\left\{\log\left(\frac{x}{K}\right) + T\left(r + \frac{\sigma^2}{2}\right)\right\}; \quad d_2 \;=\; d_1 - \sigma\sqrt{T}\,.$$

We note that $C(0,x)$ is a homogeneous function of degree 1 in (x,K).

Dependence of C on x

It is interesting to see how the call price evolves as a function of the underlying asset's price, i.e., to evaluate $\dfrac{\partial C}{\partial x}$. This is called the *delta*, and is a generalization of the Δ of Sect. 1.1.5 of Chap. 1. It is the amount of risky asset held in the hedging portfolio, which is made up of the underlying asset and the bond (see Theorem 3.3.3).

Intuitively, if the market rate of the underlying asset increases, then so does that of the call option.

Differentiating (3.18), a straightforward but tedious calculation, shows that $\dfrac{\partial C}{\partial x} = \phi(d_1)$. Another approach involves writing $C(0,x)$ as $E(\mathrm{e}^{-rT}(Z_T^{x,0} - K)^+)$ (as in (3.16)). Since

$$Z_T^{x,0} = x\,\exp\left[(r - \frac{1}{2}\sigma^2)T + \sigma B_T\right] \qquad \text{(Example 3.1.10)}\,,$$

we get

$$
\begin{aligned}
C(0,x) \;&=\; E\big(x\exp\big(\sigma B_T - \frac{1}{2}\sigma^2 T\big) - \mathrm{e}^{-rT}K\big)^+ \\
&=\; E\big(x\exp\big(-\sigma B_T - \frac{1}{2}\sigma^2 T\big) - \mathrm{e}^{-rT}K\big)^+,
\end{aligned}
$$

as B_T has the same distribution as $-B_T$. We differentiate the expression under the expectation sign.

The derivative of the integrand equals $\exp\left(-\sigma B_T - \frac{1}{2}\sigma^2 T\right)\mathbf{1}_{B_T \le d_2\sqrt{T}}$, except at points such that $B_T = d_2\sqrt{T}$, which constitute a negligible set. Thus we obtain

$$\frac{\partial C}{\partial x} = E\left[\exp\left(-\sigma B_T - \frac{1}{2}\sigma^2 T\right)\mathbf{1}_{B_T \le d_2\sqrt{T}}\right].$$

Defining the probability measure P^* by $\mathrm{d}P^* = \left[\exp\left(-\sigma B_T - \frac{1}{2}\sigma^2 T\right)\right]\mathrm{d}P$, we get

$$\frac{\partial C}{\partial x} = P^*(B_T \le d_2\sqrt{T}) = P^*\left[B_T + \sigma T \le (d_2 + \sigma\sqrt{T})\sqrt{T}\right]$$

$$= P^*\left[\frac{B_T^*}{\sqrt{T}} \le d_2 + \sigma\sqrt{T}\right] = \phi\left(d_2 + \sigma\sqrt{T}\right) = \phi(d_1)$$

since $B_t^* = B_t + \sigma t$ is a Brownian motion under P^*. The term $\phi(d_1)$ is positive (and smaller than 1) so that when the stock price increases, the call price (premium) also increases. A one euro change in the market rate of the stock corresponds to a delta euro change in the price of the call.

Sensitivity to Volatility

The option buyer speculates, and the greater the fluctuations in the price of the underlying, the more he is prepared to pay for the option. Let us check that this intuition holds.

When a financial product satisfies the stochastic differential equation

$$\mathrm{d}S_t^1 = \mu S_t^1 \,\mathrm{d}t + \sigma S_t^1 \,\mathrm{d}B_t,$$

it is customary to say that σ represents the *volatility* of the product. Intuitively, σ represents the typical deviation of $\frac{\mathrm{d}S^1}{S^1}$, and is linked to the risk that the asset carries (the higher the coefficient, the greater the impact of the random term). Applying Itô's lemma to $C(t, S_t^1)$ yields

$$\mathrm{d}C_t = \frac{\partial C}{\partial t}\mathrm{d}t + \frac{\partial C}{\partial x}\mathrm{d}S_t^1 + \frac{1}{2}(\sigma S_t^1)^2\frac{\partial^2 C}{\partial x^2}\mathrm{d}t$$

$$= \left(\frac{\partial C}{\partial t} + \frac{\partial C}{\partial x}\mu S_t^1 + \frac{1}{2}(\sigma S_t^1)^2\frac{\partial^2 C}{\partial x^2}\right)\mathrm{d}t + \frac{\partial C}{\partial x}\sigma S_t^1 \,\mathrm{d}B_t,$$

so that the volatility of the call is $\frac{1}{C}\frac{\partial C}{\partial x}S^1\sigma$, which we can write as $v_c = \frac{S^1}{C}\frac{\partial C}{\partial x}v_S$. Setting $\eta = \frac{S^1}{C}\Delta$, we find that the volatility of the call is proportional to the volatility of the underlying stock: $v_c = \eta v_S$. This generalizes the discrete-time result.

Moreover, we have seen that

$$\frac{\mathrm{d}C_t}{C_t} = \mu_c \, \mathrm{d}t + \sigma_c \, \mathrm{d}B_t \, ,$$

with

$$\mu_c = \frac{1}{C_t} \left(\frac{\partial C}{\partial t} + \frac{1}{2} (\sigma S_t^1)^2 \frac{\partial^2 C}{\partial x^2} + \frac{\partial C}{\partial x} \mu S_t^1 \right) \, ;$$

that is, taking into account the differential equation for C,

$$\mu_c = \frac{1}{C_t} \left(r C_t - r S_t^1 \frac{\partial C}{\partial x} \right) + \frac{1}{C_t} \frac{\partial C}{\partial x} \mu S_t^1 \, ,$$

hence

$$\mu_c - r = \frac{S_t^1}{C_t} \frac{\partial C}{\partial x} (\mu - r) \, ,$$

which is a generalization of the results of Sect. 1.1.5.
Notice that

$$\eta = \frac{S^1}{C} \frac{\partial C}{\partial x} = \frac{S^1 \phi(d_1)}{S^1 \phi(d_1) - K \mathrm{e}^{-rT} \phi(d_2)} \geq 1 \, .$$

The option carries greater risk than the underlying asset.

The Market Portfolio

We define a market portfolio as a portfolio made up of one bond and one stock. Its value at time t is $M_t = S_t^0 + S_t^1$, hence

$$\mathrm{d}M_t = \left(r S_t^0 + \mu S_t^1 \right) \mathrm{d}t + \sigma S_t^1 \, \mathrm{d}B_t \, ,$$

which we can write

$$\frac{\mathrm{d}M_t}{M_t} = \mu_M \, \mathrm{d}t + \sigma_M \, \mathrm{d}B_t \, .$$

Recall that $\mathrm{Cov}_t(X, Y) = E_t(XY) - E_t(X) E_t(Y)$. Using the Itô process notation (in which we identify $\mathrm{d}X_t$ with ΔX_t, as is explained in the appendix) and assuming that X^1 and X^2 satisfy

$$\mathrm{d}X_t^i = \mu_i \, \mathrm{d}t + \sigma_i \, \mathrm{d}B_t \, ,$$

we get

$$\mathrm{Cov}_t \left(\mathrm{d}X_t^1, \mathrm{d}X_t^2 \right) = \sigma_1 \sigma_2 \, .$$

It follows that

$$\mathrm{Cov}_t \left(\frac{\mathrm{d}M_t}{M_t}, \frac{\mathrm{d}S_t^1}{S_t^1} \right) = \frac{\sigma^2 S_t^1}{S_t^0 + S_t^1}$$

and hence that

$$\mu_S - r = \left(\frac{\mathrm{Cov}_t \left(\frac{\mathrm{d}M}{M}, \frac{\mathrm{d}S^1}{S^1} \right)}{\mathrm{Var}_t \, \mathrm{d}M/M} \right) (\mu_M - r) \, .$$

The Gamma

In practice, the delta's sensitivity in the variations of the underlying market rate is an important parameter for risk management. We introduce the *gamma* of an option, which is the derivative of the delta with respect to the stock price, that is $\dfrac{\partial^2 C}{\partial x^2}$.

As $\dfrac{\partial C}{\partial x} = \phi(d_1)$, and since $\phi'(d_1) = \dfrac{1}{\sqrt{2\pi}} \exp -\dfrac{d_1^2}{2}$,

$$\frac{\partial^2 C}{\partial x^2} = \frac{1}{x\sigma\sqrt{T}} \phi'(d_1) = \frac{1}{\sqrt{2\pi T}\sigma x} \exp -\frac{d_1^2}{2} .$$

Therefore, the price of a call is a convex function of the price of the underlying.

Time to Maturity

Time has a very significant effect on options reaching maturity. Indeed, we have

$$C(t, S_t^1) = S_t^1 \phi[d_1(t)] - Ke^{-r(T-t)} \phi[d_2(t)] ,$$

with

$$d_1(t) = \frac{1}{\sigma\sqrt{T-t}} \left\{ \log \frac{x}{K} + (T-t)\left(r + \frac{\sigma^2}{2}\right) \right\}$$

and $\frac{\partial^2 C}{\partial x^2}(t, S_t^1) = \frac{1}{S_t^1 \sigma \sqrt{T-t}} \phi'[d_1(t)]$; this last expression tends to 0 as t tends to T.

The greater the time to expiry, the higher the price of the call. Therefore, it is useful to evaluate the price's sensitivity to time, that is to calculate $\dfrac{\partial C}{\partial \tau}$, setting $\tau = T - t$.

$$\frac{\partial C}{\partial \tau} = x\phi'(d_1)d_1'(\tau) + rKe^{-r\tau}\phi(d_2) - Ke^{-r\tau}\phi'(d_2)d_2'(\tau) .$$

Notice that

$$\phi'(d_2) = \phi'(d_1)e^{-\sigma^2 \frac{\tau}{2}} e^{\sigma\sqrt{\tau}d_1} = \frac{x}{K}\phi'(d_1)e^{r\tau}$$

and

$$d_2'(\tau) = d_1'(\tau) - \sigma/2\sqrt{\tau} ,$$

so that

$$\frac{\partial C}{\partial \tau} = rKe^{-r\tau}\phi(d_2) + Ke^{-r\tau}\phi'(d_2)\sigma/2\sqrt{\tau} ,$$

i.e.,

$$\frac{\partial C}{\partial \tau} = \frac{x\sigma}{2\sqrt{\tau}}\phi'(d_1) + Ke^{-r\tau}r\phi(d_2) ,$$

which is positive. The price of a call is an increasing function of maturity.

Exercise 3.3.11. Show that the price of a call is a convex and decreasing function of the strike price.

3.4 Extension of the Black–Scholes Formula

We are now going to generalize the formula obtained in the previous section to the multidimensional case with stocks paying dividends.

3.4.1 Financing Strategies

The market comprises $(d + 1)$ assets: one bond (riskless) and d stocks. The price vector of the $(d + 1)$ assets is denoted by S_t, and that of the d risky assets, by S_t^* .

We suppose that asset i, including the bond, pays a *dividend*. We write $(D_t^i, i \geq 0)$ for the dividend paid by one share of asset i up until time t.

We call the process $G : G_t = S_t + D_t$ the *gains process*. We assume G to be an Itô process. In formal terms, the total gain of a portfolio θ (capital plus dividend) is given by

$$\int_0^t \theta(s)\,\mathrm{d}G_s := \int_0^t \theta(s)\,\mathrm{d}S_s + \int_0^t \theta(s)\,\mathrm{d}D_s .$$

However, we only assume the existence of $\int \theta\,\mathrm{d}G$, that is we impose that θ belongs to $\Theta(G)$.

Definition 3.4.1. *Let Z be a positive-valued $\mathcal{F}_T$-measurable variable. We say that $\theta \in \Theta(G)$ finances Z if*

$$\theta_t \cdot S_t = \theta_0 \cdot S_0 + \int_0^t \theta(s)\,\mathrm{d}G(s) \qquad t \in [0,T] \,;\ a.s.$$
$$\theta_T \cdot S_T = Z \qquad a.s. .$$

We can generalize this definition by adding a dividend rate to Z.

Definition 3.4.2. *Let $Z \in \mathcal{F}_T$ and let $(\zeta_t, t \geq 0)$ be a $\mathcal{F}_t$-adapted process such that $\int_0^T |\zeta(s)|\mathrm{d}s < \infty$ a.s.. We say that $\theta \in \Theta(G)$ finances (ζ, Z) if*

$$\begin{cases} \theta_t \cdot S_t = \theta_0 \cdot S_0 + \int_0^t \theta(s)\,\mathrm{d}G(s) - \int_0^t \zeta(s)\,\mathrm{d}s , & t \in [0,T] \,;\ a.s. \\ \theta_T \cdot S_T = Z & a.s. . \end{cases}$$

We can also interpret ζ in terms of consumption (see Chaps. 4 and 8) or in terms of refinancing costs (see 2.4). The amount $\theta_0 \cdot S_0$ is the implicit price at time 0 of (ζ, Z). Under NAO, this price only depends on (ζ, Z). The implicit price at time t is $\pi(\zeta, Z)_t = \theta_t \cdot S_t$.

If there is a riskless asset with return r, the discounted gain is the process G^d defined by $G_t^\mathsf{d} = R_t S_t + \int_0^t R_s \mathrm{d}D_s$. It is easy to show that if θ finances (ζ, Z), and if we use the notation $V_t = \theta_t \cdot S_t$, we get

$$R_t V_t = \theta_0 \cdot S_0 + \int_0^t \theta(s) \, \mathrm{d}G^{\mathsf{d}}(s) - \int_0^t R(s)\zeta(s) \, \mathrm{d}s \; .$$

In particular, if G^{d} is a Q-martingale, we obtain

$$R_t V_t = E_Q \left(R_T V_T + \int_t^T R(s)\zeta(s)\mathrm{d}s \,|\mathcal{F}_t \right) \; .$$

3.4.2 The State Variable

We now assume that the economy is described by a state vector $Y_t \in \mathbb{R}^d$, satisfying the stochastic differential equation

$$\mathrm{d}Y_t \;=\; \nu(t, Y_t) \, \mathrm{d}t + \eta(t, Y_t) \, \mathrm{d}B_t \tag{3.19}$$

where $(B_t, t \geq 0)$ is a given m-dimensional Brownian motion, ν is a function from $[0, T] \times \mathbb{R}^d$ with values in $\mathbb{R}^d$, and η is a function from $[0, T] \times \mathbb{R}^d$ with $d \times m$-matrix values.

We suppose that ν and η are measurable, and Lipschitz in x, uniformly with respect to t. This ensures that there is a unique[14] process Y satisfying (3.19).

We assume that the *prices of risky assets* are functions of Y_t, so that

$$S_t^* \;=\; \mathcal{Y}(t, Y_t) \, ,$$

where $\mathcal{Y}$ is a function of $C^{1,2}([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$ such that the matrix $\mathcal{Y}_y = \left(\dfrac{\partial \mathcal{Y}_i}{\partial y_j} \right)_{i,j}$ is invertible. This assumption corresponds to the fact that the market is complete, and will enable us to obtain the existence of a portfolio that finances a terminal value.

The bond is assumed to have a constant price $S_t^0 = 1$.

The *dividend* processes are given by the rates, and are assumed to be functions of Y_t :

$$\frac{dD_t}{\mathrm{d}t} \;=\; (r(t, Y_t), \, \delta(t, Y_t)) \, ,$$

where $r(t, y)$ represents the short term interest rate, and $\delta(t, y) \in \mathbb{R}^d$ is the dividend rate.

---

[14] For precise information on stochastic differential equations and on the concepts of the existence and uniqueness of their solutions, one can refer to Rogers and Williams [315], Karatzas and Shreve [233], or Øksendal [294].

3.4.3 The Black–Scholes Formula

We consider financing strategies for (ζ, Z) in the case

$$\zeta_t = \Delta(t, Y_t) \qquad \text{where} \quad \Delta : [0, T] \times \mathbb{R}^d \to \mathbb{R}$$
$$Z = g(Y_T) \qquad \text{where} \quad g : \mathbb{R}^d \to \mathbb{R} \,.$$

We assume that there exists a function $C \in C^{1,2}([0, T] \times \mathbb{R}^d, \mathbb{R})$ such that $\pi(\zeta, Z)_t = C(t, Y_t)$, and a portfolio θ that finances (ζ, Z). We want to determine C. Let

$$V_t := \pi(\zeta, Z)_t = \theta_t \cdot S_t \,.$$

Using the fact that θ finances (ζ, Z), we obtain

$$dV_t = \theta_t \, dG_t - \Delta(t, Y_t)dt = \theta_t \, dS_t + \theta_t \, dD_t - \Delta(t, Y_t)dt$$
$$= \alpha_t \, dS_t^0 + \beta_t \, dS_t^* + \{\alpha_t r(t, Y_t) + \beta_t \cdot \delta(t, Y_t)\} \, dt - \Delta(t, Y_t)dt$$

for $\theta_t = (\alpha_t, \beta_t)$ with $\alpha_t \in \mathbb{R}$, $\beta_t \in \mathbb{R}^d$.

We use Itô's lemma to calculate dS_t^*, noticing that $dS_t^0 = 0$ since $S_t^0 = 1$. Thus, taking $\mathcal{L}$ to be the infinitesimal generator associated with Y,

$$dV_t = \{\beta_t \cdot [\mathcal{L}\mathcal{Y}(t, Y_t) + \delta(t, Y_t)] + [\alpha_t r(t, Y_t) - \Delta(t, Y_t)]\} \, dt$$
$$+ \beta_t \cdot \mathcal{Y}_y(t, Y_t) \, \eta(t, Y_t) \, dB_t \,. \quad (3.20)$$

Moreover, from $V_t = C(t, Y_t)$ we obtain

$$dV_t = \mathcal{L}C(t, Y_t)dt + C_y(t, Y_t) \cdot \eta(t, Y_t) \, dB_t \,. \quad (3.21)$$

Identifying the coefficients of dB_t in (3.20) and (3.21) yields:

$$\beta_t \cdot \mathcal{Y}_y(t, Y_t) = C_y(t, Y_t) \,,$$

and hence, as $\mathcal{Y}_y$ is invertible,

$$\beta_t^T = C_y^T \, \mathcal{Y}_y^{-1}$$

and $C(t, Y_t) = \alpha_t + \beta_t \cdot S_t^* = \alpha_t + C_y \cdot \mathcal{Y}_y^{-1} S_t^*$. Hence

$$\alpha_t = C - C_y \cdot \mathcal{Y}_y^{-1} S_t^* \,.$$

We now identify the dt coefficients in (3.20) and (3.21), and substitute in the above expressions for α_t and β_t. It follows that

$$C_y \cdot \mathcal{Y}_y^{-1} \left(\mathcal{L}\mathcal{Y} + \delta - rS_t^* \right) = \Delta + \mathcal{L}C - rC \,.$$

We now look for an explicit expression for the operator $\mathcal{L}$, which depends on the coefficients ν and η of the process for Y. For $d = 1$, writing $\mathcal{Y}_{yy}$ for the matrix with components $\left(\dfrac{\partial^2 \mathcal{Y}}{\partial y_i \, \partial y_j} \right)_{i,j}$, we get

$$C_y \cdot \mathcal{Y}_y^{-1} \left(\mathcal{Y}_y \nu + \mathcal{Y}_t + \frac{1}{2} \operatorname{tr} \left(\eta \eta^T \, \mathcal{Y}_{yy} \right) + \delta - r S_t^* \right)$$

$$= \Delta - rC + C_y \nu + C_t + \frac{1}{2} \operatorname{tr} \left(\eta \eta^T \, C_{yy} \right) .$$

(For $d > 1$, the term $\operatorname{tr} \left(\eta \eta^T \, \mathcal{Y}_{yy} \right)$ is defined as a vector whose coordinates correspond to the components of $\mathcal{Y}$).

Setting

$$\gamma(t, Y_t) = -\mathcal{Y}_y^{-1}(t, Y_t) \left\{ \mathcal{Y}_t(t, Y_t) + \frac{1}{2} \operatorname{tr} \left(\eta \eta^T \, \mathcal{Y}_{yy}(t, Y_t) \right) + \delta - r \mathcal{Y}(t, Y_t) \right\} ,$$

and cancelling out the $C_y \nu$ terms, we find that

$$C_y \gamma + C_t + \frac{1}{2} \operatorname{tr} \left(\eta \eta^T \, C_{yy} \right) = rC - \Delta .$$

Denote by $\hat{\mathcal{L}}$ the operator

$$\hat{\mathcal{L}} C = C_y \gamma + C_t + \frac{1}{2} \operatorname{tr} \left(\eta \eta^T \, C_{yy} \right) .$$

The function C is a solution to the partial differential equation

$$\hat{\mathcal{L}} C - rC = -\Delta , \tag{3.22}$$

and is subject to the boundary condition

$$C(T, y) = g(y) . \tag{3.23}$$

We can obtain representations of solutions to (3.22) and (3.23) by applying the Feynman-Kac formula:

Theorem 3.4.3. *Under regularity conditions[15], the unique solution to (3.22)-(3.23) is given by*

$$C(t, y) = E \left\{ \int_t^T e^{-\phi(s)} \Delta \left(s, W_s^{y,t} \right) \mathrm{d}s + e^{-\phi(T)} g \left(W_T^{y,t} \right) \right\}$$

where the process $W_s^{y,t}$ is the unique solution to

$$\begin{cases} \mathrm{d} W_s^{y,t} = \gamma(s, W_s^{y,t}) \, \mathrm{d}s + \eta(s, W_s^{y,t}) \, \mathrm{d}B_s \\ W_t^{y,t} = y , \end{cases}$$

and where $\phi(s) = \int_t^s r(u, W_u^{y,t}) \, \mathrm{d}u.$

---

[15] We need to impose regularity conditions (differentiability and growth conditions) on C to obtain the uniqueness of the solution. To have the existence and uniqueness of the solution, it is enough for all the functions involved to be C^2, Lipschitz and with Lipschitz first and second order derivatives. We can refer to Krylov [245] or to Varadhan [358].

The function C represents the implicit price of a financing strategy for (ζ, Z) in the case where $\zeta_t = \Delta(t, Y_t)$ and $Z = g(Y_T)$.

Remark 3.4.4. The results of Theorem 3.4.3 still hold when T is a stopping time. This sort of problem occurs in the pricing of American options, where we are dealing with a free boundary problem.

When Δ depends on the whole history of the process, the results are still valid, but we no longer have a PDE.

3.4.4 Special Case

We are going to write the Black–Scholes formula in the special case where $S_t^* = Y_t$ is a one dimensional process that we write S^1. We assume that

$$\mathrm{d}S^1(t) = \mu(t, S_t^1)\,\mathrm{d}t + \sigma(t, S_t^1)\mathrm{d}B_t \ .$$

In this case, the implicit price of a financing strategy for (ζ, Z) when $\zeta_t = \Delta(t, S_t^1)$ and $Z = g(S_T^1)$ is the unique solution to

$$\hat{\mathcal{L}}C - rC \;=\; -\Delta \ ,$$

subject to the boundary condition

$$C(T, x) \;=\; g(x) \ ,$$

where

$$\hat{\mathcal{L}}C \;=\; r(t, x)xC_x' + C_t' + \frac{1}{2}\sigma(t, x)C_{xx}'' \ .$$

The solution can be written in the form

$$C(t, x) \;=\; E\left\{ \int_t^T \mathrm{e}^{-\phi(s)}\,\Delta(s, W_s^{x,t})\mathrm{d}s + \mathrm{e}^{-\phi(T)}\,g(W_T^{x,t}) \right\}$$

where $W_s^{x,t}$ is the unique process solution to

$$\begin{cases} \mathrm{d}W_s^{x,t} \;=\; W_s^{x,t}r(s, W_s^{x,t})\,\mathrm{d}s + \sigma(s, W_s^{x,t})\,\mathrm{d}B_s \\ W_t^{x,t} \;=\; x \ , \end{cases}$$

and where $\phi(s) = \int_t^s r(u, W_u^{x,t})\,\mathrm{d}u$.

3.4.5 The Risk-Neutral Measure

Let us return to the general case, and show that we can also obtain the results of Theorem 3.4.3 by using the concept of the risk-neutral probability measure.

A risk-neutral probability measure Q is such that the discounted gains process

$$G_t^d = R_t S_t + \int_0^t R_s dD_s$$

is a Q-martingale. Let θ be a financing strategy for (ζ, Z) and let $V_t = \theta_t \cdot G_t$. We have

$$R_t V_t + \int_0^t R_s \zeta_s \, ds = V_0 + \int_0^t \theta_s \, dG_s^a \, .$$

If $\left(\int_0^t \theta_s dG_s^d, t \geq 0 \right)$ were a Q-martingale (which would be the case under integrability conditions placed on θ), we would have $R_t V_t = E_Q \left(Z + \int_t^T R_s \zeta_s ds \,|\, \mathcal{F}_t \right)$.

To determine Q, we note that

$$dG_t^d = R_t \left[\left(\frac{\partial \mathcal{Y}}{\partial t} + \nu \frac{\partial \mathcal{Y}}{\partial y} + \frac{1}{2}\eta^2 \frac{\partial^2 \mathcal{Y}}{\partial y^2} + \delta - r\mathcal{Y} \right)(t, Y_t) \right] dt$$
$$+ \left(\eta \frac{\partial \mathcal{Y}}{\partial y} \right)(t, Y_t) \, dB_t \, .$$

Let

$$h(t, Y_t) = \left(\frac{\partial \mathcal{Y}}{\partial t} + \nu \frac{\partial \mathcal{Y}}{\partial y} + \frac{1}{2}\eta^2 \frac{\partial^2 \mathcal{Y}}{\partial y^2} + \delta - r\mathcal{Y} \right)(t, Y_t) \left[\eta \frac{\partial \mathcal{Y}}{\partial y} \right]^{-1}(t, Y_t) \, ,$$

and let Q be the measure that is equivalent to P and is defined by Girsanov's probability density[16] $dL_t = -h_t L_t dB_t$. This probability measure is such that

$$d\tilde{B}_t = dB_t + h_t dt$$

is a Q-Brownian motion. Under Q the state variable has the dynamics

$$dY_t = -\left(\frac{\partial \mathcal{Y}}{\partial t} + \frac{1}{2}\eta^2 \frac{\partial^2 \mathcal{Y}}{\partial y^2} + \delta - r\mathcal{Y} \right)(t, Y_t) \left[\frac{\partial \mathcal{Y}}{\partial y} \right]^{-1}(t, Y_t) \, dt$$
$$+ \eta(t, Y_t) \, d\tilde{B}_t \, .$$

As a result, the price of a financing strategy for $(\Delta(t, Y_t), g(Y_T))$ is

$$C(t, y) = E_Q \left\{ \int_t^T e^{-\phi(s)} \Delta(s, W_s) ds + e^{-\phi(T)} g(W_T) \,|\, W_t = y \right\}$$

where W_s solves

$$dW_s = \gamma(s, W_s) \, ds + \eta(s, W_s) \, d\tilde{B}_s \, , \quad W_t^t = y$$

with $\phi(s) = \int_t^s r(u, W_u) \, du$ and

$$\gamma(s, x) = -\left(\frac{\partial \mathcal{Y}}{\partial t} + \frac{1}{2}\eta^2 \frac{\partial^2 \mathcal{Y}}{\partial y^2} + \delta - r\mathcal{Y} \right)(t, x) \left[\frac{\partial \mathcal{Y}}{\partial y} \right]^{-1}(t, x) \, .$$

---

[16] Under integrability conditions, which do in fact hold if all the parameters, functions and their partial derivatives are bounded.

3.4.6 Example

Take the case of 3 assets, with $r = 0$, $\Delta = 0$, $D = 0$, and with B, a 2-dimensional Brownian motion. We assume that

$$S_t = \mathcal{Y}(Y_t, t) = Y_t \quad \text{and} \quad g(Y_T) = \max\left(Y_T^1 - Y_T^2, 0\right),$$

where Y satisfies an equation of the same type as (3.19). We assume that

$$dY_t^1 = \nu_t^1 \, dt + \sigma^1 Y_t^1 \, dB_t$$
$$dY_t^2 = \nu_t^2 \, dt + \sigma^2 Y_t^2 \, dB_t,$$

where σ^1 and σ^2 are constant vectors. This corresponds to the option to exchange one unit of asset 1 for one unit of asset 2 at time T.

The generalized Black–Scholes equation is then written $C_t + \frac{1}{2} \operatorname{tr}(\eta \eta^T C_{yy}) = 0$, and the process $W_s^{y,t}$ satisfies $dW_s^{y,t} = \eta(W_s^{y,t}) dB_s$ where

$$\eta(y_1, y_2) = \begin{bmatrix} \sigma^{1,1} y_1 & \sigma^{1,2} y_1 \\ \sigma^{2,1} y_2 & \sigma^{2,2} y_2 \end{bmatrix} \quad \text{with} \quad \sigma^i = \begin{bmatrix} \sigma^{i,1} \\ \sigma^{1,2} \end{bmatrix}.$$

Then

$$C(t, y) = C(t, y_1, y_2) = y_1 \phi(d_1) - y_2 \phi(d_2),$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} \, du \, ;$$

$$d_1 = \log\left\{ \frac{y_1}{y_2} + \frac{1}{2} V^2 (T - t) \right\} \frac{1}{V \sqrt{T - t}},$$

$$d_2 = d_1 - V \sqrt{T - t} \quad ; \quad V^2 = \|\sigma^1 - \sigma^2\|^2.$$

3.4.7 Applications of the Black–Scholes Formula

To summarize, the Black–Scholes formula has enabled us to draw out two concepts.

Asset Prices

A financial agent wants to sell (δ, Z). What price is he willing to accept? Meanwhile, a buyer would pay at most $C(t, Y_t)$; for at a higher price than $C(t, Y_t)$, he could find a strategy that would give him a greater gain than Z.

Hedging Strategies

In order to receive a cash flow (δ, Z), we must be prepared to incur risks. We can hedge ourselves against these risks, by following the strategy $\theta_t = (-a_t, -b_t)$ that finances $(-\delta, -Z)$. To implement this strategy, we have to pay out $-a_0 S_0^0 - b_0 \cdot S_0^*$ initially. This is the hedging cost.

In this model, we have not taken into account transaction costs or other market imperfections (e.g., portfolio constraints).

Notes

Chung and Williams [59], (1983), is a good reference book for stochastic integration. Protter [304], (2005), contains numerous results on stochastic integration and stochastic differential equations, in great generality. Readers that are more interested in Brownian motion, can refer to Karatzas and Shreve [233], (1991), Rogers and Williams [315], (1988), and to Revuz and Yor [307], (1999), who study it in detail. The Feymann-Kac formulae that we have used, are covered by Varadhan [358], (1980), Krylov [245] (1980), and Rogers and Williams [315], (1988). A detailed study of the links between partial differential equations and stochastic calculus is carried out in Varadhan [358], (1980), and in Karatzas and Shreve [233], (1991). The reader can also refer to Øksendal [294], (1998), which is an excellent first approach to all of these issues and to Shreve [338], (2004), Mikosch [278],(1999), and Steele [345],(2001), for stochastic calculus "with finance in view"

We have contented ourselves with models whose price processes are continuous. Stochastic processes involving jumps are introduced in Merton [275], (1976). Models with jumps, in which asset prices are modeled by Poisson processes, are to be found in Jeanblanc-Picqué and Pontier [215], (1990), and in Lamberton and Lapeyre [250], (1997). Extensions to general processes with jumps are given by Aase and Øksendal [1], (1988), Shiryaev [336], (1999), Madan [261], (2001), and are presented in detail in Cont and Tankov [64], (2004), the collective Deutsche Bank book [295], (2002), and in the forthcoming book of Jeanblanc et al. [216].

More general set–ups (the price process as a semi-martingale) are studied by Föllmer and Sonderman [163], (1986) and Schweizer [330] (1991). For a general approach, see Shiryaev [336] (1999).

The classic work concerning the options market is the Cox and Rubinstein book [72], (1985), which contains numerous notes and interesting references. Kat [235], (2001), provides varied and up-to-date applications. The book Jarrow and Rudd [213], (1983), is very accessible, and takes an intuitive approach to the stochastic calculus. The series of Wilmott's books [369, 370], (1998, 2001) presents an interesting and accurate practitioner's approach, as do Overhaus et al. [295], (2000), and Brockhaus et al. [46], (2000).

The Black–Scholes formula, which originated in Black and Scholes, [37], (1973), has since been studied by many authors: one of the first articles was Merton's [274], (1973); later Harrison and Pliska [179], (1983), studied the formula using stochastic calculus, drawing their inspiration from the discrete-time methods of Cox et al. [71], (1979).

We have only looked at the case of European options. American options (giving the right to exercise at any time between 0 and maturity) are covered in Karatzas [229], (1988), Lamberton and Lapeyre [250], (1997) as well as in Elliott and Kopp [149], (1998). The options literature is extensive, due to the diversity of options available. We will return to this topic in Chap. 9.

An altogether different problem is that of valuation in the presence of transaction costs or of constraints. It is no longer possible to replicate the option. The problem of valuation with transaction costs has been addressed by Bensaid, Lesne, Pagès and Sheikmann [27], (1992), in discrete time, and by Davis, Panas and Zariphopoulou [89], (1993) and Cvitanić [32, 74], (1996, 2001), in continuous time.

We have only looked at pricing in the case of complete markets. When there is financing strategy for a product, its valuation is straightforward. In the opposite case (the incomplete market case) the problem is much harder. We will come back to it in Chap. 8.

The Black–Scholes formula only produces an explicit result when the coefficients (r and σ) are deterministic. When volatility is random, the market is in general incomplete. A presentation of such models can be found in Fouque et al. [164], (2000). El Karoui and Jeanblanc-Picqué [139], (1991), show that if we can put deterministic bounds on the volatility, then we can deduce a price range and a hedging strategy.

The study of the relationship between arbitrage and the existence of a probability measure that is equivalent to P and turns discounted prices into martingales, was taken on by Harrison and Kreps [177], (1979), and by Harrison and Pliska [178], (1981), in a space Ω that has a finite number of elements, and in discrete time. A good approach to the problem in continuous time is to be found in the book Müller [284], (1987).

We have shown that in discrete time, the assumption of no arbitrage is equivalent to the existence of a probability measure Q, equivalent to the historic probability P, and under which discounted prices are martingales (a martingale measure). The same is not true in continuous time. In the continuous-time case, we need to introduce restrictions on the strategies that we use (see Dalang et al. [80], (1989); Morton [281], (1989), Delbaen and Schachermayer [96, 97, 98], (1993,1994, 2005), Kabanov [224], (2001), Xia and Yan [372], (2001).

Arbitrage with constraints and/or transaction costs are studied in Kabanov and Stricker, [227], (2002).

More generally, various authors been interested in models with asset prices driven by semi-martingales. The reader can consult for example Shiryaev [336], (1999), for a general presentation and references.

The link between market completeness and the uniqueness of the martingale measure has been studied by Harrison and Pliska [179], (1983). Further information can be found in Björk [34], (1998), Bingham and Kiesel [33], (1998), Shiryaev [336], (1999).

ANNEX 3

In this annex, we present definitions and precise results pertaining to martingales, the stochastic integral, stochastic differential equations, and to the link between partial differential equations and stochastic calculus, without carrying out any proofs.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, let $(B_t)_{t \geq 0}$ be a real-valued Brownian motion and let $\mathcal{F}_t = \sigma(B_s, \, s \leq t)$ be its natural filtration completed by the addition of null sets. Recall that a process X such that for all t, the random variable X_t is $\mathcal{F}_t$-measurable, is said to be adapted. A process is said to be continuous if for P-almost all ω, the mapping $t \to X_t(\omega)$ is continuous.

1 Martingales

Definition $(M_t, t \geq 0)$ *is a **martingale** with respect to the filtration $\mathcal{F}_t$ if*

- $(M_t, t \geq 0)$ *is a $\mathcal{F}_t$-adapted process,*
- M_t *is integrable for all t, that is to say that $E(|M_t|) < \infty$,*
- $M_t = E_t(M_s)$, *when $0 \leq t \leq s$, where E_t denotes the conditional expectation with respect to $\mathcal{F}_t$.*

  $(M_t, t \geq 0)$ *is a **supermartingale** with respect to the filtration $\mathcal{F}_t$ if*

- $(M_t, t \geq 0)$ *is a $\mathcal{F}_t$-adapted process,*
- M_t *is integrable for all t,*
- $M_t \geq E_t(M_s)$, *when $0 \leq t \leq s$.*

If $(M_t, t \geq 0)$ is a martingale, then $E(M_t) = E(M_0), \quad \forall t$.

If $(M_t, \, t \leq T)$ is a martingale, the process is fully determined by its final value: $M_t = E(M_T | \mathcal{F}_t)$.

If M is a martingale, then for any adapted bounded process ψ,

$$E\left(M_T \int_0^T \psi(s)\, ds \right) = E\left(\int_0^T M_s \psi(s)\, ds \right).$$

(Check this holds for any step–function ψ, and use limits).

We need a weaker notion of martingale: that of a local martingale. Let us define this new object.

Definition *A stopping time is a random variable $T : \Omega \to \overline{\mathbb{R}}_+$ such that $(T \leq t) \in \mathcal{F}_t$, $t \in \mathbb{R}_+$. A local martingale $(M_t)_{t \geq 0}$ is an adapted process such that there exists a sequence of stopping times T_n satisfying*

$$T_n \leq T_{n+1} \quad ; \quad T_n \underset{n \to +\infty}{\longrightarrow} +\infty ,$$

and such that for any n, $(M_{t \wedge T_n})_{t \geq 0}$ is a martingale.

When working with local martingales, we can revert to the study of martingales, by introducing the sequence T_n. Some of the properties that are true for $t \wedge T_n$, hold true when $n \to \infty$. We then say that we are working by localization.

A martingale is a local martingale, this follows from Doob's optional sampling theorem:

Theorem *Let M be a continuous martingale and let T be a stopping time. The process $M^T = (M_t^T = M_{t \wedge T}, t \geq 0)$ is a martingale.*

If M is a uniformly integrable martingale, if S and T are two stopping times such that $S \leq T$, then $E(M_T | \mathcal{F}_S) = M_S$.

Proposition *If M is a continuous positive local martingale then it is a supermartingale.*

Proof. As $(M_t, t \geq 0)$ is a local martingale , there exists a sequence of stopping times T_n such that, for all n, $E(M_{t \wedge T_n} | \mathcal{F}_s) = M_{s \wedge T_n}$. Let n tend to infinity. We can apply Fatou's lemma for conditional expectations, as M is positive. It follows, using the continuity of M, that:

$$M_s = \lim_{n \to \infty} E(M_{t \wedge T_n} | \mathcal{F}_s) \geq E(\lim_{n \to \infty} M_{t \wedge T_n} | \mathcal{F}_s) = E(M_t | \mathcal{F}_s) .$$

□

A continuous semi-martingale[17] is a continuous process $(X_t, t \geq 0)$ such that $X_t = M_t + A_t$ where M is a continuous local martingale and A is a continuous process of finite variation.

2 The Itô Integral

To define the stochastic integral of a process with respect to Brownian motion, we start by defining $\int_0^t h(s)\,dB_s$ when h is a "elementary" process:

Definition *We say that h is an elementary process if there exist $(t_i)_{i \leq n}$; $0 = t_0 < t_1 < \cdots < t_n$ such that*

$$h(t) \ = \ h_0 \, 1_{[t_0, t_1]}(t) + \sum_{i=1}^{n-1} h_i \, 1_{]t_i, t_{i+1}]}(t) ,$$

where for any i, h_i is a bounded $\mathcal{F}_{t_i}$-measurable random variable .

The elementary processes play the role of step–functions.

Definition *Let h be an elementary process. We define the random variable $I(h)$, denoted $\int_0^\infty h(s)\,dB_s$, by*

---

[17] See Revuz and Yor [307] p. 121.

$$I(h) \; = \; \sum_{i=0}^{n-1} h_i \left[B_{t_{i+1}} - B_{t_i} \right] \; .$$

We define the process $I_t(h)$, also denoted $\int_0^t h(s)\, \mathrm{d}B_s$, by $I(h) := I\left(h\,1_{[0,t]}\right)$.

Property *Let h be an elementary process. The following properties hold:*

a)   *the process $(I_t(h), t \geq 0)$ is a continuous martingale,*

b)   *the process $(I_t^2(h) - \int_0^t h^2(s)\, \mathrm{d}s, t \geq 0)$ is a continuous martingale,*

c) $E(I_t^2(h)) \; = \; E\left[\int_0^t h^2(s)\, \mathrm{d}s \right].$

These properties are easy to prove using the elementary properties of Brownian motion and of conditional expectations. Property c) follows from b), and shows us that the mapping $h \to I(h)$ is an *isometry* from the space of elementary processes equipped with the norm $L^2(\Omega \times [0, \infty[, \, P \times \mathrm{d}t)$ into $L^2(\Omega, \mathcal{F}, P)$. Using the sets' density, we can extend the isometry:

Definition *Let Λ be the set of processes h such that there exists a sequence h_n of elementary processes that converge to h in $L^2(\Omega \times [0, T], \, \mathrm{d}P \times \mathrm{d}t)$.*

For $h \in \Lambda$, define

$$I(h) \; = \; \lim_{n \to \infty} I(h_n)$$
$$I_t(h) \; = \; I\left(h\,1_{[0,t]} \right) \; .$$

The isometry property shows that $I(h)$ is well-defined, and that we have:

Property   *For $h \in \Lambda$, $(I_t(h), \, t \geq 0)$ and $\left(I_t^2(h) - \int_0^t h^2(s)\, \mathrm{d}s\,, \, t \geq 0 \right)$ are continuous martingales.*

It remains to determine a class of processes that is a subset of Λ.

Definition *A process h is said to be predictable if it is measurable with respect to the σ-field on $(\Omega \times \mathbb{R}_+)$ generated by left-continuous adapted processes.*

This σ-field is also generated[18] by sets of the form $]s, t] \times A_s$ where $A_s \in \mathcal{F}_s$, and $\{0\} \times A_0$ where $A_0 \in \mathcal{F}_0$.

Property *Λ contains predictable processes h such that $E\left[\int_0^\infty h^2(s)\, \mathrm{d}s \right] < \infty$.*

We now recall the properties of the stochastic integral that are used most frequently.

Property
We suppose that the processes h and g belong to Λ. We write $I_t(h) =$ $\int_0^t h(u)\mathrm{d}B_u$. The following properties hold:

---

[18] See Chung and Williams [59] Chap. 3.

a) The martingale property. $I(h)$ *is a martingale:*

$$E\left(\int_0^t h(u)\mathrm{d}B_u|\mathcal{F}_s\right) = \int_0^s h(u)\mathrm{d}B_u , \quad \forall s \leq t .$$

In particular $E(I_t(h)) = 0$.

b) The increasing process associated with $I_t(h)$: *The process* $\left(\left\{\int_0^t h(u)\mathrm{d}B_u\right\}^2 - \int_0^t h^2(s)\mathrm{d}s, t \geq 0\right)$ *is a martingale. Equivalently, the increasing process associated with* $I_t(h)$ *is* $\int_0^t h^2(s)\mathrm{d}s$.

Thence we deduce that the variance of $I_t(h)$ *is* $\int_0^t E(h^2(s)))\mathrm{d}s$. *Similarly we have*

$$E\left(\int_0^t h(u)\,\mathrm{d}B_u \int_0^t g(u)\,\mathrm{d}B_u\right) = E\left(\int_0^t h(s)g(s)\,\mathrm{d}s\right) .$$

Note that Λ contains the left-continuous processes belonging to $L^2(\Omega \times [0,\infty[, P \times \mathrm{d}t)$. In fact, when we integrate with respect to Brownian motion, Λ contains the adapted, measurable (i.e., such that $(t,\omega) \to h(t,\omega)$ is $\mathcal{B}_{\mathbb{R}} \times \mathcal{F}$-measurable) processes belonging to $L^2(\Omega \times [0,\infty[, P \times \mathrm{d}t)$. Predictable processes were introduced to enable us to define the concept of stochastic integration with respect to any square-integrable martingale.

The integrability condition placed on h is very strong. We would like to weaken it, to the detriment of some of the properties of our stochastic integral. Fortunately, the loss is not a serious one.

Definition *Let* $\Theta(B)$ *be the set of predictable processes such that*

$$\int_0^t h^2(s)\,\mathrm{d}s < \infty \quad P\text{-}a.s. \quad \text{for any } t \in \mathbb{R}_+ .$$

By considering the stopping times $T_p = \inf\left\{t > 0 , \int_0^t h^2(s)\,\mathrm{d}s > p\right\}$, we can define the martingale $\int_0^{t \wedge T_p} h(s)\,\mathrm{d}B_s$. It is enough to let p tend to infinity to obtain the following result:

Proposition *For* $h \in \Theta(B)$, *the process* $\int_0^t h(s)\,\mathrm{d}B_s$ *is a local martingale.*

3 Girsanov's Theorem

We are going to study the influence of certain changes of probability measure on Gaussian variables, in order to gain a better understanding of the workings of Girsanov's Theorem.

a) Gaussian Variables

Let X be a normally distributed random variable with expectation $E(X)$ and variance $\operatorname{Var} X$. The Laplace transform of X is

$$E(e^{\lambda X}) = \exp(\lambda E(X) + 1/2\lambda^2 \operatorname{Var} X) .$$

A family of random variables is said to be Gaussian, if any linear combination of these random variables is also Gaussian.

The limit of Gaussian variables is Gaussian. In particular, $\int_s^t X_u du$ is a Gaussian variable if the family $(X_u, s \le u \le t)$ is Gaussian.

If (X, Y) is a Gaussian vector, the distribution of X conditional on Y is a Gaussian distribution with expectation $E(X \,|Y)$ and with the conditional variance as its variance. It follows that $E(e^{\lambda X} \,|Y) = \exp(\lambda E(X \,|Y) + 1/2\lambda^2 \operatorname{Var}_Y X)$.

b) Change of Measure for a Gaussian Variable

Proposition *Let X be a Gaussian variable with mean m and variance σ^2. Let $L = \exp\left[h(X - m) - \frac{1}{2}h^2\sigma^2\right]$ where h is a constant. The variable L is positive, has expectation 1, and defines a probability density function: let Q be such that $dQ = LdP$. Under Q, the variable X is a Gaussian variable with variance σ^2 and mean $m + h\sigma^2$.*

Proof. We give an explicit expression for the density of X under Q, by evaluating

$$E_Q(f(X)) = E_P(Lf(X))$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} f(x) \exp\left(h(x - m) - \frac{1}{2}h^2\sigma^2 - \frac{(x - m)^2}{2\sigma^2}\right) dx$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} f(x) \exp{-\frac{1}{2\sigma^2}(x - m - h\sigma^2)^2}dx .$$

The density of X under Q is then

$$\frac{1}{\sqrt{2\pi}\sigma} \exp{-\frac{1}{2\sigma^2}\left(x - m - h\sigma^2\right)^2} ,$$

which proves the proposition. $\qquad\square$

c) Change of Measure

Recall that two probability measures P and Q defined on $(\Omega, \mathcal{F})$ are equivalent if $P(A) = 0 \Leftrightarrow Q(A) = 0$.

- The Radon–Nikodym density:

Let P and Q be two equivalent probability measures. There exists a positive $\mathcal{F}$-measurable random variable f, such that $Q(A) = E_P(f 1_A)$. We use the notation $\dfrac{dQ}{dP} = f$.

- Conditional expectation:

  It is important to be able to express the conditional expectation of a variable X under Q, with respect to its conditional expectation under P. We have

  $$E_Q(X\,|\mathcal{G}) \;=\; \frac{E_P(Xf\,|\mathcal{G})}{E_P(f\,|\mathcal{G})}\;.$$

d) The Brownian Motion Case

It is easy to show that if B is a Brownian motion, then the process $\left(\exp(\lambda B_t - \dfrac{\lambda^2}{2}t), t \geq 0\right)$ is a martingale, for any λ. The converse can easily be obtained, using the fact that the equality $E\left(e^{\lambda X}|\mathcal{G}\right) = E\left(e^{\lambda X}\right)$ implies the independence of X and $\mathcal{G}$. This result can be generalized: if $X_t = \mu t + \sigma B_t$ then $\exp\left(\lambda X_t - (\mu\lambda + \frac{1}{2}\sigma^2\lambda^2)t\right)$ is a martingale for any λ, and conversely.

Using these properties, and the change of measure rule for conditional expectations, we can check that if $X_t = \mu t + \sigma B_t$, and if Q is defined on $\mathcal{F}_T$ by $dQ = L_T dP$ with

$$L_t = \exp\left(\gamma X_t - (\mu\gamma + \frac{1}{2}\sigma^2\gamma^2)t\right)\;,$$

then the process X_t can be written as

$$X_t = \left(\mu + \gamma\sigma^2\right)t + \sigma\tilde{B}_t\;,$$

where $\tilde{B}$ is a Q-Brownian motion.

e) Girsanov's Theorem

It is easy to prove that a similar result to that of paragraph b) holds for a Gaussian vector, and easy to convince oneself that Girsanov's theorem is valid when the process h is elementary.

Remark Let $L(h)$ be the martingale exponential, defined in Girsanov's theorem by $L_t(h) = 1 + \int_0^t L_s(h)h_s\,dB_s$, and let Q be the measure defined on $\mathcal{F}_T$ by $Q(A) = E_P(1_A L_T)$. This measure has total mass 1 as L_T has expectation 1. Moreover, if $A \in \mathcal{F}_t$ we have $Q(A) = E_P(1_A L_t)$. Indeed, using the properties of conditional expectation, and the martingale property of L_t, we get $E_P(1_A L_T) = E_P[1_A E_P(L_T|\mathcal{F}_t)] = E_P(1_A L_t)$.

Let us now state the multidimensional version of Girsanov's Theorem.

Theorem *Let $B = (B^1, B^2, \ldots, B^d)^T$ be a d-dimensional Brownian motion. Let $h_t = (h_t^1, h_t^2, \ldots, h_t^d)$ be a predictable process such that*

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

The process

$$L_t(h) := \exp \int_0^t h(s)\, \mathrm{d}B_s - \frac{1}{2} \int_0^t \|h_s\|^2 \, \mathrm{d}s$$

satisfies $L_t(h) = 1 + \sum_{i=1}^d \int_0^t L_s(h) h_s^i \, \mathrm{d}B_s^i$; *it is a continuous local martingale.*

If $E(L_T(h)) = 1$, *then the process* $L(h)$ *is a martingale, and* $Q(A) = E(1_A L_T)$ *defines a probability measure on* $(\Omega, \mathcal{F}_T)$ *such that* $Q(A) = E(1_A L_t)$ *for* $A \in \mathcal{F}_t$.

The process $(B_t^* = B_t - \int_0^t h_s \, \mathrm{d}s, 0 \le t \le T)$ *is a* $(\Omega, \mathcal{F}_t, Q)$ *Brownian motion.*

Remark It is enough for h to be adapted and to belong in $L^2(\mathrm{d}t)$, for then we know how to define the stochastic integral with respect to Brownian motion.

If h is bounded, we can show that $E(L_T(h)) = 1$. In fact, for this last condition to hold, it is sufficient to have $E\left(\exp \frac{1}{2} \int_0^T \|h_s\|^2 \, \mathrm{d}s\right) < \infty$ (the Novikov condition). Improved sets of conditions are to be found in Karatzas and Shreve [233].

f) The Martingale Exponential

We can show that if M_t is a continuous, strictly positive martingale (with respect to a Brownian filtration), then there exists q_t such that $\int_0^t q^2(s)\mathrm{d}s < \infty$, *a.s.* and $M_t = \exp\left(\int_0^t q(s)\mathrm{d}B_s - \frac{1}{2}\int_0^t q^2(s)\mathrm{d}s\right)$.

4 Stochastic Differential Equations

A stochastic differential equation is an equation of the form

$$X_t = X_0 + \int_0^t b(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}B_s ,$$

or in the condensed form

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t . \tag{1}$$

The data of the problem is given by b, σ, the space Ω, the Brownian motion B, and the initial condition X_0, which we take to be constant. The unknown is the process X. The problem, as in the case of ordinary differential equations, is to show that when certain conditions are imposed on the coefficients, differential equation (1) has a unique solution, in the sense that two solutions are equal a.s.. We say that X is a diffusion.

It is also a Markov process: let $X_t^{s,x}$ be the solution to (1) with the initial condition at time s: $(X_s^{s,x} = x)$. For $t \ge s$, we have

$$E(f(X_t)|\mathcal{F}_s) = \Phi(X_s) \,,$$

where $\Phi(x) = E\left(f(X_t^{s,x})\right).$

Theorem *We assume that the coefficients b and σ are Lipschitz in x, and this uniformly with respect to t. That is, we assume that there exists K such that for any $t \in [0, T]$, $x \in \mathbb{R}$, $y \in \mathbb{R}$*

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \;\leq\; K|x - y| \,,$$

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \;\leq\; K^2(1 + |x|^2) \,.$$

Then there exists a unique solution to differential equation (1).

In particular, if we want to solve the stochastic differential equation $\mathrm{d}X_t = X_t\,\theta_t\,\mathrm{d}B_t$, $\quad X_0 = 1$, the unique solution is

$$X_t = \exp\left[\int_0^t \theta_s\,\mathrm{d}B_s - \frac{1}{2}\int_0^t \|\theta_s\|^2\,\mathrm{d}s\right] \,.$$

This is a martingale whenever θ is bounded. Moreover, when the integrals are defined, X_t is positive.

5 Partial Differential Equations

We take two functions b and σ from $[0, T] \times \mathbb{R}$ into $\mathbb{R}$, and satisfying the conditions of the previous theorem. Let $\mathcal{L}$ the operator defined on $C^{1,2}([0, T] \times \mathbb{R}, \mathbb{R})$ by

$$\mathcal{L}f(t,x) \;=\; f_t(t,x) + f_x(t,x)b(t,x) + \frac{1}{2}\sigma^2(t,x)f_{xx}(t,x) \,.$$

We say that $\mathcal{L}$ is the infinitesimal generator of diffusion X.

We take as given a final value, that is a function g from $\mathbb{R}$ into $\mathbb{R}$. We now look for solutions to the following problem $\mathcal{P}$: find f such that

$$\mathcal{L}f(t,x) \;=\; 0 \,, \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, T]$$

$$f(T,x) \;=\; g(x) \,, \quad \forall x \in \mathbb{R} \,.$$

For any (t, x), we can define the process $Z^{t,x}$ as the solution to

$$Z^{t,x}(u) \;=\; x + \int_t^u b(s, Z_s^{t,x})\mathrm{d}s + \int_t^u \sigma(s, Z_s^{t,x})\mathrm{d}B_s \,. \tag{2}$$

It is the diffusion with infinitesimal generator $\mathcal{L}$, and initial point x at time t.

Let f be a solution to $\mathcal{P}$. Then applying Itô's Lemma to $f(u, Z_u^{t,x})$ yields

$$f(T, Z_T^{t,x}) = f(t, Z_t^{t,x}) + \int_t^T \mathcal{L}f(s, Z_s^{t,x})\mathrm{d}s + \int_t^T f_x(s, Z_s^{t,x})\sigma(s, Z_s^{t,x})\mathrm{d}B_s \,,$$

which, as $\mathcal{L}f = 0$ and $f(T, x) = g(x)$, leads to

$$g(Z_T^{t,x}) = f(t, x) + \int_t^T f_x(s, Z_s^v) \sigma(s, Z_s^v) dB_s .$$

Hence, under the required integrability conditions, we can deduce that $f(t, x) = E(g(Z_T^{t,x}))$.

We can generalize the problem, and look for solutions to

$$\begin{aligned} \mathcal{L}f(t, x) &= rf(t, x) , \quad x \in \mathbb{R} , \quad \forall t \in [0, T] , \\ f(T, x) &= g(x) , \end{aligned} \tag{3}$$

where r is a constant. We can show that:

Theorem *If $g \in C_b^2(\mathbb{R})$ and if g, g' and g'' are Lipschitz, then the function*

$$f(t, x) := E\big[e^{-r(T-t)} g(Z_T^{t,x})\big]$$

is the unique Lipschitz solution to (3).

Let us give a sketch proof of the theorem. Let f be a solution to (3). We apply Itô's formula to $e^{r(T-s)} f(s, Z_s^{t,x})$, ($t$ fixed, $t \le s \le T$). We obtain

$$f(T, Z_T^{t,x}) - e^{r(T-t)} f(t, Z_t^{t,x})$$

$$= \int_t^T e^{r(T-s)} \left[-rf(s, Z_s^{t,x}) + \frac{\partial f}{\partial s}(s, Z_s^{t,x}) + \frac{1}{2}(\sigma(s, Z_s^{t,x}))^2 \frac{\partial^2 f}{\partial x^2}(s, Z_s^{t,x}) \right] ds$$

$$+ \int_t^T e^{r(T-s)} \frac{\partial f}{\partial x}(s, Z_s^{t,x}) dZ_s^{t,x}$$

$$= \int_t^T e^{r(T-s)} \left\{ -rf + \frac{\partial f}{\partial s} + b\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} \right\}(s, Z_s^{t,x}) ds$$

$$+ \int_t^T e^{r(T-s)} \frac{\partial f}{\partial x}(s, Z_s^{t,x}) \sigma(s, Z_s^{t,x}) dB_s .$$

The first integral (with respect to time) is zero, because f is a solution to (3).

We take the expectations of the two remaining terms. Under regularity conditions, the integral with respect to B is the value at time T of a martingale that is zero at time t, and hence its expectation is also zero. Using the boundary condition, this entails

$$E[g(Z_T^{t,x}) - e^{r(T-t)} f(t, x)] = 0 ,$$

and hence we obtain the result.

4

Portfolios Optimizing Wealth and Consumption

In this chapter, we give a generalization of the model studied in Chap. 1 Sect. 1.3.

The financial market is that of Chap. 3: it is given by a continuous-time model, comprising a riskless asset and d risky assets. The market is complete. An investor with an initial wealth of x at time 0 maximizes the sum of the expected utility of consumption over the planning horizon and the expected utility of wealth at the end of the planning horizon, without running into debt.

Firstly, we show how dynamic programming can be used to study the problem in the case of constant or deterministic coefficients. This method provides an explicit characterization of the optimal portfolio in terms of the value function, which, under regularity conditions, solves the Hamilton–Jacobi–Bellman equation. This is a non-linear equation and in general it is difficult to solve.

Using the techniques of stochastic calculus (our fundamental tool here is the predictable representation theorem, which we recall in the annex), we show that the dynamic constraint on wealth remaining positive is equivalent to a single constraint on an expectation. We then obtain the existence of a portfolio and of a consumption plan enabling us to achieve the required optimization. Thus we prove the market's completeness. We then go on to study the properties of the value function.

Finally, we show how in the case of deterministic coefficients, we can express the solution in terms of partial differential equations of the parabolic type, and thus reduce the problem to the study of two Cauchy problems, which are easier to solve than the Hamilton–Jacobi–Bellman equation.

4.1 The Model

We consider a financial market of $(d+1)$ assets, whose dynamics are modeled as in Chap. 3 Sect. 3.2.

The first asset is riskless, and its price S^0 satisfies the following equation

$$dS^0(t) = S^0(t) r(t) dt,$$
$$S^0(0) = 1.$$

The prices of the d risky assets satisfy the stochastic differential equation

$$dS^i(t) = S^i(t) \left\{ b_i(t) dt + \sum_{j=1}^d \sigma_{i,j}(t) dB_t^j \right\},$$

where $B = (B^1, B^2, \ldots, B^d)^T$ is a d-dimensional Brownian motion. Denote by $\mathcal{F}_t$ the σ-field generated by the paths of B up until t and then completed; $\mathcal{F}_0$ is the σ-field generated by the negligible sets. We work with a finite horizon, that is with $t \in [0, T]$.

We lay down the following hypotheses[1].

H(i)   The processes r, b and σ are measurable, $\mathcal{F}_t$-adapted and uniformly bounded on $([0, T] \times \Omega)$; r is positive.

H(ii)   The matrix $\sigma(t)$ is invertible, its inverse is bounded for all $t \in [0, T]$ and the process $\sigma(t)$ is predictable.

We will show (Proposition 4.4.3 and Sect. 4.7) that under these hypotheses, the market is complete and presents no arbitrage opportunities.

An agent has an initial wealth of x at time 0. He chooses a portfolio. Let $\theta_i(t)$ be the number of shares of type i in his possession at time t. His wealth at time t is therefore

$$X(t) = \sum_{i=0}^d \theta_i(t) S^i(t). \tag{4.1}$$

We suppose that the agent uses a self-financing strategy, and that his consumption between times 0 and t is given by $\int_0^t c(s) ds$, where c is a positive $\mathcal{F}_t$-adapted process: the strategy θ finances c, in the sense of the previous chapter.

As a result, the agent's wealth $X(t)$ satisfies

$$X(t) = \theta(t) \cdot S(t) = x + \int_0^t \theta(s) dS(s) - \int_0^t c(s) ds,$$

where $S = (S^0, S^1, \ldots, S^d)^T$ and where $\cdot$ denotes the scalar product. We can also write this in the form

$$\begin{cases} dX_t = \sum_{i=0}^d \theta_i(t) dS^i(t) - c(t) dt, \\ X_0 = x. \end{cases}$$

---

[1] These assumptions can be refined. See Karatzas and Shreve [233].

We will now write the equation in a form that is more convenient for subsequent calculations. Let π_i be the amount of wealth invested in the i-th risky asset, so that $\pi_i(t) = \theta_i(t)S^i(t)$ for $i \geq 1$. Using (4.1), and writing $X^{\pi,c}$ for the wealth process associated with $(\pi_i, 1 \leq i \leq d)$ and with the consumption c, we get:

$$
\begin{cases}
dX_t^{\pi,c} = \{X_t^{\pi,c} r(t) - c(t)\}\, dt + \displaystyle\sum_{i=1}^{d} \pi_i(t)\, \{b_i(t) - r(t)\}\, dt \\[2mm]
\qquad\qquad + \displaystyle\sum_{i,j=1}^{d} \pi_i(t)\, \sigma_{i,j}(t)\, dB_t^j\,, \\[2mm]
X_0^{\pi,c} = x\,.
\end{cases}
\tag{4.2}
$$

We denote by $\pi(t)$ the vector $(\pi_1(t), \ldots, \pi_d(t))^T$, which is also referred to as the agent's portfolio of risky assets at time t. No sign restrictions are made (the agent may borrow or sell assets short).

For equation (4.2) to have a unique solution[2], we impose the following conditions on the parameters c and π.

H(iii) c is a positive adapted process such that $\int_0^T c(t)\, dt\ <\ \infty$ P-a.s.

H(iv) π is predictable, and satisfies $\int_0^T \|\pi(t)\|^2\, dt\ <\ \infty$ P-a.s. where $\|\cdot\|$ denotes the norm of a vector.

As the coefficients b, r and σ are bounded, the process $X^{\pi,c}$ is an Itô process.

Exercise 4.1.1. Using Itô's formula, check that

$$
S^*(t)\ =\ S^*(0)\ \exp\left\{ \int_0^t \{b(s) - \frac{1}{2}\sigma(s)\sigma^T(s)\}\, ds + \int_0^t \sigma(s)\, dB_s \right\},
$$

where $S^*(t)\ =\ [S^1(t), \ldots, S^d(t)]^T$, and that

$$
X_t^{\pi,c} R(t)\ =\ x + \int_0^t R(s)\, \{-c(s) + \pi(s)^T (b(s) - r(s)\mathbf{1})\}\, ds
$$

$$
+ \int_0^t R(s)\, \pi(s)^T\, \sigma(s)\, dB_s\,,
$$

where $\mathbf{1}$ represents the vector $(1, \ldots, 1)^T$ and where $R(t) = \exp\left[-\int_0^t r(s)\, ds \right]$.

---

[2] See Rogers and Williams [315] or Karatzas and Shreve [233].

4.2 Optimization

We assume that the investor's preferences are represented by an additively separable function of his rate of change in consumption and of his final wealth. More precisely:

Notation 4.2.1. Let U_1 and U_2 be two functions from $\mathbb{R}_+$ into $\mathbb{R}$, and satisfying:

U1 U is strictly concave, strictly increasing and of class C^1,

U2 $\lim_{x \to +\infty} U'(x) = 0$.

Assumption **U1** shows us that the function U' is strictly decreasing. Thus it admits a continuous inverse, denoted I and defined on $]U'(\infty), U'(0)[$. Assumption **U2** shows that I is defined on $]0, U'(0)[$. We extend I at 0, on the right of $U'(0)$, when $U'(0)$ is finite.

Notation 4.2.2. For a given pair (π, c), we use the notation

$$J(x; \pi, c) := E\left\{ \int_0^T U_1(c(t))\mathrm{d}t + U_2(X_T^{\pi,c}) \right\}$$

where E denotes expectation under P and where x is the initial value of $X_t^{\pi,c}$, i.e., $x = X_0^{\pi,c}$.

We suppose here that the investor seeks to maximize $J(x; \pi, c)$ under a path-wise constraint: his wealth $X^{\pi,c}$ must remain positive (the investor is not allowed to run into debt).

4.3 Solution in the Case of Constant Coefficients

4.3.1 Dynamic Programming

We describe the principle of dynamic programming in the case where the coefficients r, b and σ are **deterministic**, with a view to using these results in a later section. We only actually solve the problem in the case of constant coefficients.

The reasoning behind dynamic programming is as follows: we suppose that if we had wealth X_α at time α, then we would be able to optimize between times α and T, and this for all X_α. We then look at the set of strategies that produce wealth X_α at time α, and pick the best of these strategies. To formalize this argument, we need to introduce a new parameter: a time-parameter for our initial point in time (until now, the initial point was always taken to be time 0).

Let $\alpha \in]0, T[$. Let us describe the dynamics for the wealth process $X_t^{\alpha,x}$ of an agent with an initial capital of x at time α. It is easy to show that in our model:

$$
\begin{cases}
\mathrm{d}X_t^{\alpha,x} = [X_t^{\alpha,x} r(t) - c(t)] \, \mathrm{d}t + \sum_{i=1}^{d} \pi_i(t) \, [b_i(t) - r(t)] \, \mathrm{d}t \\
\qquad\qquad + \sum_{i,j=1}^{d} \pi_i(t) \sigma_{i,j}(t) \, \mathrm{d}B_t^j \, , \\
X_\alpha^{\alpha,x} = x \, ,
\end{cases}
\tag{4.3}
$$

for $\alpha \le t \le T$. The process $(X_t^{\alpha,x}, t \ge \alpha)$ also depends on (π, c), but to lighten the notation, we have written $X_t^{\alpha,x}$ instead of $X_t^{\alpha,x;c,\pi}$. We use the notation $\mathbf{1}$ for the vector $(1, \ldots, 1)^T$.

Definition 4.3.1. *A pair (π, c) is admissible for an initial wealth equal to x at time α, if it satisfies $\mathbf{H(iii)}$ and $\mathbf{H(iv)}$, and if the associated wealth $X_t^{\alpha,x}$ remains positive between times α and T. We will denote the set of such admissible pairs by $\mathcal{A}(\alpha, x)$.*

Thus, to optimize between times α and T is a matter of maximizing, over $\mathcal{A}(x, \alpha)$, the value of

$$
J(\alpha, x; \pi, c) := E \left\{ \int_\alpha^T U_1(c_t)\mathrm{d}t + U_2(X_T^{\alpha,x}) \right\} .
$$

4.3.2 The Hamilton–Jacobi–Bellman Equation

Let V be the value function, i.e.,

$$
V(\alpha, x) := \sup \left\{ J(\alpha, x; \pi, c) \, ; \, (\pi, c) \in \mathcal{A}(\alpha, x) \right\} .
$$

A pair (π^*, c^*) is optimal if it is admissible, and if $J(\alpha, x; \pi^*, c^*) = V(\alpha, x)$. The principle of dynamic programming[3] can be written:

The Principle of Dynamic Programming   *For all (α, x), for all $t \ge \alpha$*

$$
V(\alpha, x) = \sup_{(\pi,c)\in\mathcal{A}(\alpha,x)} E \left\{ \int_\alpha^t U_1(c_s)\mathrm{d}s + V(t, X_t^{\alpha,x;\pi,c}) \right\} .
\tag{4.4}
$$

Assuming regularity conditions on the value function[4], this function is to be found as one of the solutions to the so-called Hamilton–Jacobi–Bellman equation:

---

[3] Refer to the annex for an intuitive approach, and to the work of Fleming and Rishel [155] for an exhaustive study.

[4] Refer to the annex for an approach to the proof.

The Hamilton–Jacobi–Bellman Equation

$$\frac{\partial V}{\partial t}(t,x) + \sup_{(\pi,c)\in\mathbb{R}^+\times\mathbb{R}^d} \left\{ [xr_t - c + \pi^T(b_t - r_t\mathbf{1})]\frac{\partial V}{\partial x}(t,x) \right.$$

$$\left. + \frac{1}{2}\|\pi^T\sigma_t\|^2\frac{\partial^2 V}{\partial x^2}(t,x) + U_1(c) \right\} = 0 \quad ; \quad t\in[0,T[\,,\ x\in\mathbb{R}_{++} \quad (4.5)$$

with the boundary conditions

$$\begin{aligned} V(T,x) &= U_2(x)\,, & x &\geq 0\,; \\ V(t,0) &= 0\,, & t &\in[0,T]\,. \end{aligned} \quad (4.6)$$

Using the infinitesimal generator $\mathcal{L}$ associated with the diffusion X, the HJB equation can be written

$$\frac{\partial V}{\partial t}(t,x) + \sup_{(\pi,c)}\left(\mathcal{L}V(t,x) + U_1(c)\right) = 0\,.$$

In the annex, we will give an example of conditions under which the value function solves the HJB equation.

If we know that there exists an optimal pair (π^*,c^*), and that the value function satisfies the HJB equation, then we can determine that pair. Indeed, as the value function V satisfies the HJB equation, we have for all t

$$\frac{\partial V}{\partial t}(t,X_t^*) + \left[r_t X_t^* - c_t^* + (b_t - r_t\mathbf{1})^T\pi^*(t)\right]\frac{\partial V}{\partial x}(t,X_t^*)$$

$$+ \frac{1}{2}\|\sigma_t^T\pi^*(t)\|^2\frac{\partial^2 V}{\partial x^2}(t,X_t^*) \leq -U_1(c_t^*)\,.$$

Hence, by applying Itô's formula to $V(t,X_t^*)$ between times 0 and T, taking expectations (as long as the stochastic integral involved is indeed a martingale), and using the boundary conditions, we get

$$0 = V(0,x) - E(U_2(X_T^*))$$

$$+ E\left\{\int_0^T\left[\frac{\partial V}{\partial t}(s,X_s^*) + \left[r_s X_s^* - c_s^* + (b_s - r_s\mathbf{1})^T\pi^*(s)\right]\right.\right.$$

$$\left.\left.\times\frac{\partial V}{\partial x}(s,X_s^*) + \frac{1}{2}\|\sigma_s^T\pi^*(s)\|^2\frac{\partial^2 V}{\partial x^2}(s,X_s^*)\right]\mathrm{d}s\right\}$$

$$\leq V(0,x) - E\left[\int_0^T U_1(c_s^*)\mathrm{d}s + U_2(X_T^*)\right]\,.$$

As the pair (π^*,c^*) is optimal,

$$V(0,x) = E\left[\int_0^T U_1(c_s^*)\mathrm{d}s + U_2(X_T^*)\right] .$$

Hence, the previous inequality becomes an equality, yielding,

$$E\left\{\int_0^T\left[\frac{\partial V}{\partial t}(s,X_s^*) + \left[r_s X_s^* - c_s^* + (b_s - r_s\mathbf{1})^T\pi^*(s)\right]\frac{\partial V}{\partial x}(s,X_s^*)\right.\right.$$
$$\left.\left. + \frac{1}{2}\|\sigma_s^T\,\pi^*(s)\|^2\,\frac{\partial^2 V}{\partial x^2}(s,X_s^*) + U_1(c_s^*)\right]\mathrm{d}s\right\} = 0 .$$

The expression above can only hold if the integrand is identically zero, for it is a sum of terms that are all either negative or zero. As, by assumption, the value function satisfies the HJB equation, this implies that (π^*, c^*) maximizes, for all t,

$$\left[r_t X_t^* - c_t + (b_t - r_t\mathbf{1})^T\pi(t)\right]\frac{\partial V}{\partial x}(t,X_t^*) + \frac{1}{2}\|\sigma_t^T\,\pi(t)\|^2\,\frac{\partial^2 V}{\partial x^2}(t,X_t^*) + U_1(c_t) ;$$

i.e.,

$$c_t^* = I_1\left(\frac{\partial V}{\partial x}(t,X_t^*)\right)$$

and

$$\pi_t^* = -(\sigma_t\sigma_t^T)^{-1}(b_t - r_t\mathbf{1})\frac{\partial V}{\partial x}(t,X_t^*)\left\{\frac{\partial^2 V}{\partial x^2}(t,X_t^*)\right\}^{-1} .$$

If the stochastic integral is not a martingale, we apply the same reasoning only working by localization.

Proposition 4.3.2. *Suppose that there exists an optimal pair, and that the value function satisfies the HJB equation. Under the previous assumptions, the optimal consumption–portfolio pair corresponding to a wealth equal x at time t, is given by*

$$c_t^* = I_1\left[\frac{\partial V}{\partial x}(t,x)\right] \tag{4.7}$$

$$\pi_t^* = -\left[\sigma_t\sigma_t^T\right]^{-1}\left[b_t - r_t\mathbf{1}\right]\frac{\partial V}{\partial x}(t,x)\left[\frac{\partial^2 V}{\partial x^2}(t,x)\right]^{-1} . \tag{4.8}$$

We say of such a pair that it is in feedback form, to express the fact that it is determined at time t as a function of the optimal wealth at time t.

Remark 4.3.3. It is extremely important to underline the following points.

Firstly, in general it is difficult to solve the HJB equation explicitly and obtain the value function.

Secondly, the optimal pair that we have exhibited depends on the optimal wealth, which would still need to be determined.

Thirdly, the particular form of the HJB equation enabled us to exhibit an optimal couple, on condition that such a couple existed. In particular, the assumption of optimality implies that the pair is admissible. These results can easily be extended to the case where the utility function depends on time.

Equation (4.8) has played a very important role in the financial literature, since $\dfrac{\pi_t}{\|\pi_t\|}$ is independent of the utility functions. The equation provides a mutual fund theorem.

There are solutions to the HJB equation that are not value functions for our problem. Let us clarify the link between the two concepts:

Theorem 4.3.4. *Let v be a function of $C^{1,2}(]0,T[\times\mathbb{R}_+,\mathbb{R}_+)$ satisfying (4.5) and the boundary conditions*

$$
\begin{aligned}
v(T,x) &= U_2(x)\,, & x \geq 0\,; \\
v(t,0) &= 0\,, & t \in [0,T]\,.
\end{aligned}
$$

Then $V(t,x) \leq v(t,x)$, $0 \leq t < T$, $0 \leq x < \infty$, where V is the value function defined in (4.4).

Proof. Let v be a solution to the HJB equation, and let $X^{t,x}$ be the process initialized at x at time t and associated with the admissible pair (π,c). We write

$$
\tau_n = T \wedge \inf\left\{s \in [t,T] \mid X_s^{t,x} = n \text{ or } X_s^{t,x} = \frac{1}{n} \text{ or } \int_t^s \|\pi(u)\|^2\,du = n\right\},
$$

which is a stopping time. Due to the boundedness conditions on the processes up until time τ_n, the stochastic integral

$$
\int_t^{\tau_n} \frac{\partial v}{\partial s}(s,X_s^{t,x})\,\pi^T(s)\,\sigma(s)\,\mathrm{d}B_s
$$

is the value at time τ_n of a martingale that is zero at time t, and it therefore has zero expectation.

Using Itô's Lemma, we show that

$$
\begin{aligned}
& E\left[v(\tau_n, X_{\tau_n}^{t,x})\right] \\
&= v(t,x) + E\left[\int_t^{\tau_n}\left[\frac{\partial v}{\partial s}(s,X_s^{t,x}) + \left\{r_s X_s^{t,x} - c_s + \pi^T(s)(b_s - r_s\mathbf{1})\right\}\right.\right. \\
&\qquad\left.\left.\times\frac{\partial v}{\partial x}(s,X_s^t) + \frac{1}{2}\|\pi_s^T\sigma_s\|^2\frac{\partial^2 v}{\partial x^2}(s,X_s^{t,x})\right]\mathrm{d}s\right] \\
&\leq v(t,x) - E\left[\int_t^{\tau_n} U_1(c_s)\,\mathrm{d}s\right]\,.
\end{aligned}
$$

Hence the result follows by letting n tend to infinity, and by taking the supremum over the set of admissible pairs. $\qquad\square$

It is important to note that we have not used the boundary condition $v(t, 0) = 0$. We could have hoped for the HJB equation to have a unique solution satisfying our boundary conditions, but this is not the case.

However, using general theorems[5], we can prove the following result:

Theorem 4.3.5. *Let v be a solution to the HJB equations that satisfies the conditions of Theorem 4.3.4 , and let (c^*, π^*) be the pair defined in Proposition 4.3.2.*

If the pair is admissible, then v is the value function and the pair (c^*, π^*) is optimal.

This result is derived from the fact that the optimal pair at time t only depends on the optimal wealth at time t (and not on wealth before time t). This is called a "feedback control".

This method may seem convenient. However, we must not forget that it assumes knowledge of the value function (whereas it is not possible in general to solve equation (4.5)) and of the optimal wealth, which must remain positive. Nonetheless, on a simple example, the method allows us to bring the calculations to their conclusion.

Exercise 4.3.6. The same approach can be used to study the case where

$$J(x; \pi, c) := E\left\{ \int_0^T \Gamma(t) U_1(c(t)) \mathrm{d}t + \Gamma(T) U_2(X_T^{\pi,c}) \right\}$$

with $\Gamma(t) = \exp\left[-\int_0^t \gamma(s)\,\mathrm{d}s \right]$, where the process γ is positive and adapted.

The function we are to maximize is:

$$J(\alpha, x; \pi, c) := E\left\{ \int_\alpha^T \Gamma_t^\alpha U_1[c(t)]\mathrm{d}t + \Gamma_T^\alpha U_2(X_T^\alpha) \right\},$$

where $\Gamma_t^\alpha = \exp\left[-\int_\alpha^t \gamma(s)\,\mathrm{d}s \right]$, and the value function is

$$V(\alpha, x) = \sup_{(\pi,c)\in\mathcal{A}(\alpha,x)} E\left\{ \int_\alpha^t \Gamma_s^\alpha U_1(c_s)\mathrm{d}s + \Gamma_t^\alpha V(t, X_t^{\pi,c}) \right\}.$$

Show that the HJB equation is then

$$\frac{\partial v}{\partial t} - \gamma_t v + \sup_{(c,\pi)\in\mathbb{R}^+\times\mathbb{R}^d} \left\{ \frac{\partial v}{\partial x}[xr_t - c + \pi^T(b_t - r_t\mathbf{1})] \right.$$

$$\left. + \frac{1}{2}\frac{\partial^2 v}{\partial x^2}\|\pi^T \cdot \sigma_t\|^2 + U_1(x) \right\} = 0.$$

Find the optimal pair.

---

[5] See Fleming and Rishel [155] p. 158.

4.3.3 A Special Case

Let us find a complete solution to the optimization problem in the case where $U_1(x) = x^\alpha$, $0 < \alpha < 1$, $U_2(x) = 0$, and where the coefficients r, b and σ are constants.

The previous theorem, and a little intuition, will enable us to determine the value function and the optimal pair. We place ourselves in the case where $d = 1$ (one risky asset) so as to avoid excessively heavy calculations.

The HJB equation is then

$$\frac{\partial v}{\partial t} + \sup_{(\pi,c)\in\mathbb{R}\times\mathbb{R}_+} \left[(xr - c + \pi(b - r))\frac{\partial v}{\partial x} + \frac{1}{2}\pi^2\sigma^2\frac{\partial^2 v}{\partial x^2} + c^\alpha \right] = 0 .$$

We shall look for a value function of the form $V(t, x) = [p(t)]^{1-\alpha} x^\alpha$. We will then choose $p(t)$ in such a way that V solves the HJB equation.

Using the previous workings (or solving the HJB equation in this particular case), along with the equality $I(y) = \left(\frac{y}{\alpha}\right)^{\frac{1}{\alpha-1}}$, we find that the optimal consumption at time t for a wealth equal to x at time t, is worth

$$c(t, x) = \frac{x}{p(t)} ,$$

and that the optimal portfolio is

$$\pi(t, x) = \frac{b - r}{\sigma^2}\frac{x}{1 - \alpha} .$$

If V satisfies the HJB equation, with the supremum being attained by the pair (c, π) defined above, then p must satisfy

$$\begin{cases} p'(t) - \nu p(t) + 1 = 0 \\ p(T) = 0 , \end{cases}$$

where

$$\nu = -\frac{\alpha}{1 - \alpha}\left(r + \frac{1}{2}\frac{(b - r)^2}{\sigma^2(1 - \alpha)}\right) .$$

This equation admits as a solution

$$p(t) = \frac{1}{\nu}(1 - \exp -\nu(T - t)) ,$$

which is non-zero on $]0, T[$.

The wealth associated with the pair (c^*, π^*) satisfies

$$\mathrm{d}X_t^* = X_t^* \left[(r - \frac{1}{p(t)} + \eta^2\frac{1}{1 - \alpha})\,\mathrm{d}t - \frac{\eta}{1 - \alpha}\,\mathrm{d}B_t \right] ,$$

with $\eta = -\dfrac{b-r}{\sigma}$, and is therefore positive. The pair (c^*, π^*) is then admissible, and the wealth equals

$$x \exp\left[\int_0^t \left[r - \frac{1}{p(s)} + \eta^2 \frac{1}{1-\alpha} - \frac{1}{2}\left(\frac{\eta}{1-\alpha}\right)^2\right] ds - \int_0^t \frac{\eta}{1-\alpha}\, dB_s\right].$$

Using the explicit form of p, it is straightforward to evaluate the various integrals, to obtain

$$X_t^* = x\, \frac{p(t)}{p(0)} \exp \frac{1}{1-\alpha}\left[\left(r + \frac{1}{2}\eta^2\right)t - \eta B_t\right].$$

The hypotheses **H(iii)** and **H(iv)** of integrability for c and π hold as a result of the integrability of $\exp(aB_t)$ for a constant, and hence as a result of the parameters of the problem being constant.

Exercise 4.3.7. Calculate (c^*, X_T^*) in the case where $U_1(x) = U_2(x) = x^\alpha$, $0 < \alpha < 1$, and then in the case where $U_1(x) = U_2(x) = \ln x$.

In the case of more general utility functions, it is difficult to have an intuition for the form that the value function takes. We will now draw on another approach, that will lead to solving Cauchy problems instead of the HJB equation (which we will come back to in a special case: that of deterministic coefficients).

4.4 Admissible Strategies

We return to the general case: an agent seeks a strategy that will maximize a function of his consumption and final wealth, amongst all those strategies under which his wealth remains positive at all times.

When it is under the form $X(t) \geq 0$, this constraint is difficult to check (it is an infinite-dimensional constraint). We will give it a form that does not depend on the paths of X, and is one-dimensional. To do this, as we have often done in previous chapters, we work under the risk-neutral measure. Let

$$\eta(t) = -\left[\sigma(t)\right]^{-1}\left(b(t) - r(t)\mathbf{1}\right) \qquad (4.9)$$

and let

$$L_t = \exp\left\{\int_0^t \eta(s)\, dB_s - \frac{1}{2}\int_0^t \|\eta(s)\|^2\, ds\right\}.$$

Under the hypothesis **H**, the process η is bounded. We can then use Girsanov's theorem (in its multi-dimensional form).

Let Q be the probability measure equivalent to P and defined on $\mathcal{F}_T$ by $dQ = L_T\, dP$. Let $\tilde{B}_t = B_t - \int_0^t \eta(s)ds$. This is a Q-$\mathcal{F}_t$-Brownian motion.

Under Q, all prices have the same expected return and are governed by the stochastic differential equation

$$\mathrm{d}S^i(t) = S^i(t) \left\{ r(t)\,\mathrm{d}t + \sum_j \sigma_{i,j}(t)\,\mathrm{d}\tilde{B}_t^j \right\} ,$$

and the wealth X satisfies

$$\begin{cases} \mathrm{d}X_t^{\pi,c} = [X^{\pi,c}(t)r(t) - c(t)]\,\mathrm{d}t + \sum_{i,j} \pi_i(t)\,\sigma_{i,j}(t)\,\mathrm{d}\tilde{B}_t^j \\ X^{\pi,c}(0) = x , \end{cases}$$

or, in a closed form,

$$X^{\pi,c}(t)R(t) = x - \int_0^t c(s)R(s)\,\mathrm{d}s + \int_0^t R(s)\pi^T(s)\sigma(s)\,\mathrm{d}\tilde{B}_s , \qquad (4.10)$$

for $R(t)$ the discount factor defined by

$$R(t) = \exp\left[-\int_0^t r(s)\,\mathrm{d}s \right] .$$

Definition 4.4.1. *The pair (π, c) is admissible for an initial wealth x, if the process $X^{\pi,c}$ defined by (4.10) has non-negative values[6]. Denote by $\mathcal{A}(x)$ the set of admissible pairs.*

We give a more practical form to the path-wise constraint $X(t) \geq 0$, $t \in [0, T]$.

Let us return to equation (4.10).

If the pair (π, c) is admissible, then the process

$$M_t := x + \int_0^t R(s)\pi^T(s)\sigma(s)\,\mathrm{d}\tilde{B}_s$$

is a local Q-martingale (σ is bounded) equal to $X^{\pi,c}(t)R(t) + \int_0^t c(s)R(s)\,\mathrm{d}s$, which is a positive local martingale. It is therefore a Q-supermartingale (see annex to Chap. 3), which satisfies $E_Q(M_T) \leq M_0$. Hence we deduce the following result.

Proposition 4.4.2. *Let (π, c) be an admissible pair and let $X^{\pi,c}(T)$ be the associated final wealth. Then*

$$E_Q\left(X^{\pi,c}(T)R(T) + \int_0^T c(s)R(s)\,\mathrm{d}s \right) \leq x . \qquad (4.11)$$

---

[6] Under P or under Q, the result is the same.

Formula (4.11) is the continuous-time equivalent of the formula obtained in Sect. 1.3 of Chap. 1.

Let us prove the converse, which is given by:

Proposition 4.4.3. *Let c be a process that satisfies **H(iii)**, and let Z be a $\mathcal{F}_T$-measurable non-negative random variable, such that*

$$E_Q\left(ZR_T + \int_0^T R(s)c(s)\,\mathrm{d}s \right) \;=\; x\,. \tag{4.12}$$

Then there exists a predictable portfolio π such that the pair (π, c) is admissible and such that the associated final wealth $X_T^{\pi,c}$ is equal to Z. In particular, the market is complete.

Remark 4.4.4. If

$$E_Q\left(ZR_T + \int_0^T R(s)c(s)\,\mathrm{d}s \right) \;\leq\; x\,,$$

then there exists a predictable portfolio π such that the pair (π, c) is admissible and such that the associated final wealth $X_T^{\pi,c}$ is then lesser than or equal to Z.

Indeed, we can construct a positive and $\mathcal{F}_T$-measurable Z_1 such that $Z_1 \geq Z$ and

$$E_Q\left(Z_1 R_T + \int_0^T R(s)c(s)\,\mathrm{d}s \right) \;=\; x\,.$$

Proof. (of Proposition 4.4.3)

Note that if there exists a portfolio π such that the pair (π, c) is admissible and such that the associated wealth $X^{\pi,c}$ has final value $X_T^{\pi,c} = Z$, then the local Q-martingale

$$M_t \;=\; x + \int_0^t R(s)\,\pi^T(s)\,\sigma(s)\,\mathrm{d}\tilde{B}(s) \tag{4.13}$$

is positive, since it can be written as

$$M_t \;=\; X^{\pi,c}(t)\,R(t) + \int_0^t R(s)\,c(s)\,\mathrm{d}s\,.$$

Hence M is a supermartingale. Moreover, by assumption $E_Q(M_T) = x = E_Q(M_0)$. We can easily deduce that the supermartingale M is in fact a martingale, and satisfies

$$M_t \;=\; E_Q\left(M_T | \mathcal{F}_t \right) \;=\; E_Q\left(X_T^{\pi,c}R(T) + \int_0^T R(s)c(s)\,\mathrm{d}s \,\big|\, \mathcal{F}_t \right).$$

Let us now show the existence of a π satisfying (4.13), by using the predictable representation theorem (see annex).

For a given pair (Z, c), the process

$$M_t = E_Q\left(ZR(T) + \int_0^T R(s)c(s)\,\mathrm{d}s \mid \mathcal{F}_t\right)$$

is a martingale under Q. According to the predictable representation theorem, there exists a predictable process φ, such that $\int_0^T \|\varphi(s)\|^2\mathrm{d}s < \infty$ a.s. and

$$M_t = M_0 + \int_0^t \varphi(s)\,\mathrm{d}\tilde{B}_s .$$

As the pair (Z, c) satisfies (4.12), we have $M_0 = E_Q(M_T) = x$.

Thus, let π be the process defined by

$$\pi(t) = [R(t)]^{-1}\left[\sigma^T(t)\right]^{-1}\varphi^T(t) \qquad t \in [0, T]$$

and let $X^{\pi, c}(t)$ be defined by

$$X^{\pi, c}(t)\, R(t) = x + \int_0^t \varphi(s)\,\mathrm{d}\tilde{B}(s) - \int_0^t R(s)\,c(s)\,\mathrm{d}s .$$

Let us check that $X^{\pi, c}$ is the wealth process associated with the pair (π, c), and that the pair is admissible. By construction,

$$X^{\pi, c}(t)\, R(t) = x + \int_0^t R(s)\pi^T(s)\sigma(s)\,\mathrm{d}\tilde{B}(s) - \int_0^t R(s)\,c(s)\,\mathrm{d}s ,$$

hence $X^{\pi, c}$ is the wealth process associated with (π, c). Moreover,

$$
\begin{aligned}
X^{\pi, c}(t)\, R(t) &= M_t - \int_0^t R(s)\,c(s)\,\mathrm{d}s \\
&= E_Q\left(ZR(T) + \int_t^T R(s)c(s)\,\mathrm{d}s \mid \mathcal{F}_t\right) , \qquad (4.14)
\end{aligned}
$$

which is positive. We also get $X_T^{\pi, c} = Z$.

Finally, as the measurability and integrability conditions placed on the pair (π, c) are satisfied, the pair is admissible. $\square$

Exercise 4.4.5. Show that if a strategy finances a positive final wealth and a positive consumption, then the associated wealth always remains positive.

We can use the fact that

$$R_t X_t = E_Q\left(R_T X_T + \int_t^T R(s)c(s)\,\mathrm{d}s \mid \mathcal{F}_t\right) .$$

Remark 4.4.6. Under P, constraint (4.11) can be written as

$$E_P\left(L_T R_T X_T + \int_0^T L(s)R(s)c(s)\,\mathrm{d}s\right) = x .$$

4.5 Existence of an Optimal Pair

The optimization problem involves maximizing $J(x; \pi, c)$ over the set of admissible pairs (π, c), that is, over the set of pairs such that the associated final wealth $X^{\pi,c}(T)$ is positive and such that the pair $(c, X_T^{\pi,c})$ satisfies the constraint

$$E_P \left\{ \int_0^T L(t)R(t)c(t)\,dt + L(T)R(T)X^{\pi,c}(T) \right\} \leq x . \tag{4.11}$$

Write $V(x) := \sup\{J(x; \pi, c)\,;\,(\pi, c) \in \mathcal{A}(x)\}$. We are first going to determine a pair (c^*, X_T^*) that maximizes

$$E\left\{ \int_0^T U_1(c(t))\,dt + U_2(X_T) \right\}$$

and satisfies constraint (4.11). From there we will deduce a pair (c^*, π^*) that is admissible, according to the workings of Sect. 4.4.

Though it is an abuse of notation, we will write

$$J(x; \pi, X_T^{\pi,c}) := E_P \left\{ \int_0^T U_1(c(t))\,dt + U_2(X_T^{\pi,c}) \right\} .$$

We will say that the pair (X_T, c) is admissible when (4.11) is satisfied.

First, we establish an elementary property of utility functions, which will prove to be very important later on.

Proposition 4.5.1. *If U satisfies* **U1** *and* **U2**, *then*

$$U(I(y)) - y\,I(y) = \max\{U(c) - cy\,,\,c \geq 0\} \tag{4.15}$$

Proof. The result follows directly from the concavity of U. Indeed,

$$U(I(y)) - U(c) \geq U'(I(y))(I(y) - c) .$$

If $I(y) > 0$, $U'(I(y)) = y$, and hence

$$U(I(y)) - U(c) \geq y(I(y) - c) .$$

If $I(y) = 0$, $y \geq U'(0)$, and hence

$$U(0) - U(c) \geq -U'(0)c \geq -yc .$$

$\square$

4.5.1 Construction of an Optimal Pair

A Candidate to the Title

To construct an optimal pair, we will use the method of Lagrange multipliers, which will guide us in making an intuitive choice for the optimal pair. It will then remain to check the results of our intuition. We introduce the Lagrangian for the constrained problem. For $\lambda \in \mathbb{R}_+$, we define

$$\mathcal{L}(c, X_T\, ; \lambda) \ = \ E\left\{ \int_0^T U_1(c(t))\mathrm{d}t + U_2(X_T)\right\}$$
$$+ \lambda E\left\{ x - \left(\int_0^T L_t R_t c_t \,\mathrm{d}t + L_T R_T X_T \right)\right\}.$$

A sufficient condition for (c^*, X_T^*) to be optimal is for there to exist a Lagrange multiplier $\lambda^* \in \mathbb{R}_{++}$ such that (c^*, X_T^*, λ^*) is a saddle point of $\mathcal{L}$, that is to say that for all (c, X_T, λ), (c, X_T) satisfying (4.11) and $\lambda \in \mathbb{R}_+$,

$$\mathcal{L}(c, X_T\, ; \lambda^*) \ \leq \ \mathcal{L}(c^*, X_T^*\, ; \lambda^*) \ \leq \ \mathcal{L}(c^*, X_T^*\, ; \lambda) \,.$$

The second inequality implies that the pair (c^*, X_T^*) saturates the constraint. The first inequality implies that the pair is optimal. Thus, we can determine λ^*.

To satisfy the first inequality, we look for a (c_t^*, X_T^*) such that for all (t, ω), $c_t^*(\omega)$ maximizes $U_1(c(t, \omega)) - \lambda^* L_t R_t c(t, \omega)$, and for all ω, $X_T^*(\omega)$ maximizes $U_2(X(T, \omega)) - \lambda^* L_T R_T X(T, \omega)$.

According to Proposition 4.5.1, this brings us to study the pair

$$c_t^* \ = \ I_1(\lambda^* \zeta_t) \,; \ X_T^* \ = \ I_2(\lambda^* \zeta_T), \tag{4.16}$$

where

$$\zeta_t = R(t) L_t \,. \tag{4.17}$$

As we remarked earlier, λ^* must be such that the constraint (4.11) is saturated, i.e., such that

$$E\left\{ \int_0^T \zeta_t I_1(\lambda^* \zeta_t) \,\mathrm{d}t + \zeta_T I_2(\lambda^* \zeta_T)\right\} \ = \ x \,. \tag{4.18}$$

We will later prove the existence of λ^*.

Checking the Optimality of the Solution

Assume that there exists λ^* satisfying (4.18). Let c^* and X^* be defined as in (4.16).

We need to check that this pair is optimal. Using property 4.5.1 for utility functions, it is clear that the gain

$$J(x; c, X(T)) := E\left[\int_0^T U_1[c(s)]\,\mathrm{d}s + U_2\left[X(T)\right]\right]$$

satisfies

$$J(x; c, X(T)) \leq E\left[\int_0^T U_1[I_1(y_s)]\,\mathrm{d}s + U_2\left[I_2(y_T)\right]\right]$$

$$+ E\left[\int_0^T y_s\left[c(s) - I_1(y_s)\right]\mathrm{d}s + y_T\left[X_T - I_2(y_T)\right]\right]$$

for any positive $\mathcal{F}_s$-adapted process y_s and for any pair (c, X_T).

By choosing $y_s := \lambda^* R(s)\,L_s = \lambda^*\zeta_s$, we get for any admissible pair (c, π),

$$J(x; c, X(T)) \leq E\left\{\int_0^T U_1(c_s^*)\,\mathrm{d}s + U_2\left(X_T^*\right)\right\} = J(x; c^*, X_T^*)$$

where the pair (c^*, X_T^*) is admissible by the choice of λ^*. Hence we obtain the optimality of (c^*, X^*).

Existence of the Optimal Pair

In order to obtain the existence of λ^*, we now introduce a new assumption on the utility functions U_1 and U_2.

U3
$$\begin{cases} LRI_1(\lambda\zeta) \in L^1(\Omega \times [0, T]\,;\ \mathrm{d}P \times \mathrm{d}t) & \forall \lambda > 0\,, \\ L_T R_T I_2(\lambda\zeta_T) \in L^1(\Omega; \mathrm{d}P) & \forall \lambda > 0\,. \end{cases}$$

This assumption is given in terms of I_1 and I_2; L, R and ζ are taken as given.

Lemma 4.5.2. *Under assumptions* **U1.2.3**, *the function* $\mathcal{X}$ *defined on* $\mathbb{R}_+$ *by*

$$\mathcal{X}(y) = E\left\{\int_0^T \zeta_t\, I_1(y\zeta_t)\,\mathrm{d}t + \zeta_T\, I_2(y\zeta_T)\right\} \qquad (4.19)$$

is strictly decreasing on $[0, \bar{y}]$ *for* $\bar{y} = \inf\{y \mid \mathcal{X}(y) = 0\}$, *is continuous and satisfies*

$$\lim_{y \to 0} \mathcal{X}(y) = +\infty\,, \qquad \lim_{y \to +\infty} \mathcal{X}(y) = 0\,.$$

Therefore, $\mathcal{X}$ *has a continuous inverse* $\mathcal{Y}$. *There exists* λ^* *satisfying* (4.18).

Proof. The proof does not present any difficulties. We show separately that $\mathcal{X}$ is left-continuous and right-continuous, by monotone convergence and by dominated convergence[7]. Then we need only define $\lambda^* = \mathcal{Y}(x)$. $\square$

---

[7] See Karatzas et al. [231], Lemma 4.2.

4.5.2 The Value Function

It is then easy to obtain the value function for the problem. By substituting in for c^* and X_T^*, we obtain

$$V(x) = G(\mathcal{Y}(x)) ,$$

where G is defined on $\mathbb{R}_+$ by

$$G(y) = E \left\{ \int_0^T U_1 \left[I_1(y\zeta_t) \right] \mathrm{d}t + U_2 \left[I_2(y\zeta_T) \right] \right\} . \qquad (4.20)$$

Under the assumption

U4 U_i if of class C^2 and U_i'' is increasing, $i = 1, 2$,

we can show[8] that G and $\mathcal{X}$ are of class C^1 and that $G'(y) = y\,\mathcal{X}'(y)$. It remains to justify the differentiation of G under the integral sign.

Note that $\mathcal{X}$ and G are of the form

$$E \left\{ \int_0^T \zeta_t h_1(y\zeta_t) \, \mathrm{d}t + \zeta_T h_2(y\zeta_T) \right\}$$

for chosen functions h_i. We have already encountered functions of this type in Chap. 3, and we know that they are associated with the solutions to parabolic equations. We will be exploiting this feature in the next section.

Thus we have used techniques based on martingales and on concavity, to show the existence of an optimal pair (c^*, X_T^*). By martingale techniques again, we have associated to this pair, a portfolio π^* such that the pair (c^*, π^*) is admissible.

Let bring together the results obtained.

Theorem 4.5.3. *Under assumptions* **U1.2.3** *on utility functions U_1 and U_2, there exists an optimal pair $(\pi^*, c^*) \in \mathcal{A}(x)$ such that*

$$J(x; \pi^*, c^*) = \sup \{ J(x; \pi, c), (\pi, c) \in \mathcal{A}(x) \} .$$

Let $\mathcal{Y}$ be the inverse of the function $\mathcal{X}$ defined by (4.19) and let $\zeta_t = L_t R_t$. The pair (π^*, c^*) is determined by

$$c_t^* = I_1(\mathcal{Y}(x)\zeta_t),$$

π_t^* *is the portfolio associated with c_t^* and with the final wealth X_T^*, with*

$$X_T^* = I_2(\mathcal{Y}(x)\zeta_T)$$

where the functions I_i are the inverse functions of the U_i'. Moreover, we have

$$J(x; \pi^*, c^*) = G(\mathcal{Y}(x))$$

where G is defined by (4.20).

---

[8] See Karatzas et al. [231], Proposition 4.4.

Exercise 4.5.4. Returning to Exercise 4.3.6, show that the previous results remain true when $\zeta_t = L_t R_t \Gamma_t^{-1}$.

Let us show, by means of a few calculations, how this method can be used to produce an explicit expression for the pair (c^*, π^*).

4.5.3 A Special Case

We return to the example given in Sect. 4.3.3, with $U_1(x) = x^\alpha$, $0 < \alpha < 1$, $U_2(x) = 0$, where the coefficients r, b and σ are constants. We are going to show how we can use stochastic calculus to give a more precise solution to the problem.

Let us write U for the function U_1, and I for the inverse of U'. An expression for I is immediate:

$$I(y) = \left(\frac{y}{\alpha}\right)^{1/\alpha - 1} .$$

The optimal consumption is given by $I(\mathcal{Y}(x)\zeta_t)$ where $\mathcal{Y}(x)$ satisfies

$$x = E\left[\int_0^T \zeta_t I[\mathcal{Y}(x)\zeta_t]\,\mathrm{d}t\right] = \left(\frac{\mathcal{Y}(x)}{\alpha}\right)^{1/(\alpha-1)} E\left[\int_0^T \zeta_t^{\alpha/(\alpha-1)}\,\mathrm{d}t\right] .$$

Setting

$$K^{-1} := E\left[\int_0^T \zeta_t^{\alpha/(\alpha-1)}\,\mathrm{d}t\right] ,$$

we get $\mathcal{Y}(x) = \alpha(Kx)^{\alpha-1}$ and

$$c_t^* = I(\mathcal{Y}(x)\zeta_t) = xK\zeta_t^{1/(\alpha-1)} .$$

To obtain the optimal portfolio, we need the predictable representation of the martingale

$$M_t = E_Q\left[\int_0^T R(s)\,c^*(s)\,\mathrm{d}s \mid \mathcal{F}_t\right] ,$$

for if $U_2(x) = 0$, then the optimal final wealth is zero. More precisely, we need to determine φ such that

$$M_t = x + \int_0^t \varphi(s)\,\mathrm{d}\tilde{B}_s ,$$

where, as before, $\tilde{B}$ denotes the Q-$\mathcal{F}_t$-Brownian motion

$$\tilde{B}_t = B_t - \eta t \quad \text{with} \quad \eta = -(b - r)\sigma^{-1} .$$

We will then have

$$\pi^*(t) = [R(t)]^{-1}\,\varphi(t)\sigma^{-1} = \mathrm{e}^{rt}\sigma^{-1}\varphi(t) .$$

As the coefficients are constants, we obtain

$$\zeta_t = e^{-rt} \exp\left(\eta B_t - \eta^2 t/2\right),$$

and hence, setting $\beta = 1/(\alpha - 1)$ and $\nu = r(1 + \beta) - \frac{1}{2}\eta^2\beta(\beta + 1)$ (the same ν as in 4.3.3) and introducing the martingale exponential

$$\tilde{L}_t = \exp\left(\eta\beta\tilde{B}_t - \frac{1}{2}\eta^2\beta^2 t\right),$$

we see that

$$R_t c_t^* = xK R_t \zeta_t^\beta = xK\tilde{L}_t e^{-\nu t},$$

and that the martingale M can be given explicitly as a function of $\tilde{L}$:

$$M_t = E_Q\left(\int_0^T R_s c_s^* \, ds | \mathcal{F}_t\right)$$

$$= xK\int_0^t e^{-\nu s}\tilde{L}_s \, ds + xK\int_t^T e^{-\nu s} E_Q(\tilde{L}_s | \mathcal{F}_t) \, ds.$$

Since $\tilde{L}$ is a Q-$\mathcal{F}_t$-martingale, we obtain

$$M_t = xK\left[\int_0^t e^{-\nu s}\tilde{L}_s \, ds + \frac{e^{-\nu t} - e^{-\nu T}}{\nu}\tilde{L}_t\right].$$

Using $d\tilde{L}_t = \eta\beta\tilde{L}_t \, d\tilde{B}_t$, and the rules of stochastic calculus,

$$dM_t = xK\left[e^{-\nu t}\tilde{L}_t dt + \frac{\tilde{L}_t}{\nu}d\left(e^{-\nu t} - e^{-\nu T}\right) + \frac{e^{-\nu t} - e^{-\nu T}}{\nu}d\tilde{L}_t\right]$$

$$= \frac{xK\eta\beta}{\nu}\tilde{L}_t\left(e^{-\nu t} - e^{-\nu T}\right)d\tilde{B}_t.$$

It follows that

$$\pi_t^* = \frac{xK\eta\beta}{\nu\sigma}\left(e^{-\nu t} - e^{-\nu T}\right)e^{rt}\tilde{L}_t = \frac{xK(b - r)}{\sigma^2(\alpha - 1)\nu}\left(\exp -\nu(T-t)-1\right)\zeta_t^{1/\alpha-1}.$$

We also obtain $\nu K^{-1} = 1 - e^{-\nu T}$ and

$$G(y) = \left(\frac{y}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\frac{1 - e^{-\nu T}}{\nu}$$

$$V(x) = (x)^\alpha\left(\frac{1 - e^{-\nu T}}{\nu}\right)^{1-\alpha}.$$

Thus we recover the formulae obtained in Sect. 4.3.3.

Exercise 4.5.5. Study the case $U_1(x) = U_2(x) = x^\delta$, $0 < \delta < 1$. Show that

$$I(x) = \left(\frac{x}{\delta}\right)^{1/\delta - 1}$$

$$\mathcal{X}(y) = E\left\{\int_0^T \zeta_t \left(\frac{y\zeta_t}{\delta}\right)^{1/\delta - 1} dt + \zeta_T \left(\frac{y\zeta_T}{\delta}\right)^{1/\delta - 1}\right\} = y^{1/\delta - 1} K$$

with

$$K = E\left\{\int_0^T \zeta_t^{\delta/\delta - 1} dt + \zeta_T^{\delta/\delta - 1}\right\} \delta^{1/1 - \delta}.$$

Check that we also get $G(y) = y^{\delta/\delta - 1} \delta K$ and $V(x) = \delta x^\delta K^{1-\delta}$, and that the optimal pair is given by

$$c_t^* = \frac{x}{K} \left(\frac{\zeta_t}{\delta}\right)^{1/\delta - 1} \quad ; \quad X_T^* = \frac{x}{K} \left(\frac{\zeta_T}{\delta}\right)^{1/\delta - 1}.$$

Note that c^* and X_T^* are proportional to the initial wealth x. This is also true in the case of $\ln x$. It can be shown that $\ln$ and x^δ are the only utility functions that have this property.

4.6 Solution in the Case of Deterministic Coefficients

In this section, we assume that the coefficients r, b and σ are deterministic. Let us find the optimal portfolio at each point in time.

We retrace the steps of our study of Sect. 4.2, and make use of the dynamic programming principle.

Let us use the notation

$$R_t^\alpha := \exp\left[-\int_\alpha^t r(s) \, ds\right]$$

$$L_t^\alpha := \exp\left[\int_\alpha^t \eta(s) \, dB_s - \frac{1}{2} \int_\alpha^t \|\eta(s)\|^2 \, ds\right],$$

where the process η is defined by (4.9). The process $(L_t^\alpha, \alpha \leq t \leq T)$ is a P-martingale satisfying $dL_t^\alpha = \eta_t L_t^\alpha \, dB_t$. The condition for admissibility between times α and T is written

$$E_{Q_\alpha} \left(\int_\alpha^T R_t^\alpha c_t \, dt + R_T^\alpha X_T^\alpha\right) \leq x,$$

where $Q_\alpha = L_T^\alpha P$.

As in Sect. 4.3.1, we write $\mathcal{A}(\alpha, x)$ for the set of pairs (π, c) that are admissible between times α and T for a level of wealth x at time α. Let

$$V(\alpha, x) := \sup_{(\pi, c) \in \mathcal{A}(\alpha, x)} J(\alpha, x; \pi, c) .$$

By transposing the workings of the previous section, we obtain an optimal pair: setting, for $s \geq \alpha$ (α and x are fixed)

$$c_s^{\alpha *} = I_1(\mathcal{Y}(\alpha, x) \zeta_s^{\alpha}) , \quad X_T^{\alpha *} = I_2(\mathcal{Y}(\alpha, x) \zeta_T^{\alpha}) ,$$

where $\mathcal{Y}(\alpha, \cdot)$ is the inverse function of $\mathcal{X}(\alpha, \cdot)$ with

$$\mathcal{X}(\alpha, y) = E\left\{ \int_{\alpha}^{T} L_t^{\alpha} R_t^{\alpha} I_1(y\zeta_t^{\alpha}) \, dt + L_T^{\alpha} R_T^{\alpha} I_2(y\zeta_T^{\alpha}) \right\}$$

for

$$\zeta_t^{\alpha} = L_t^{\alpha} R_t^{\alpha} .$$

Using Itô's lemma, it can easily be shown that

$$d\zeta_t^{\alpha} = \zeta_t^{\alpha} \left[-r_t dt + \eta_t \, dB_t \right] . \tag{4.21}$$

4.6.1 The Value Function and Partial Differential Equations

The value function $V(\alpha, x)$ can be determined as in the previous section, by means of a function G defined by

$$G(\alpha, y) = E\left\{ \int_{\alpha}^{T} U_1 \left[I_1(y\zeta_t^{\alpha}) \right] \, dt + U_2 \left[I_2(y\zeta_T^{\alpha}) \right] \right\} .$$

Let $\mathcal{H}$ be the set of class $C^{1,2}$ functions on $[0, T] \times \mathbb{R}_+^*$, such that there exist K and $\alpha > 0$ with

$$\sup_{0 \leq t \leq T} |H(t, y)| \leq K(1 + y^{\alpha} + y^{-\alpha}) , \quad y > 0 . \tag{4.22}$$

The results established in the previous chapter, along with equation (4.21), satisfied by ζ^{α}, yield the following result.

Theorem 4.6.1. *If h_1 and h_2 are in $\mathcal{H}$,*

$$H(\alpha, y) = E\left\{ \int_{\alpha}^{T} h_1\left(y\zeta_t^{\alpha}\right) \, dt + h_2\left(y\zeta_T^{\alpha}\right) \right\}$$

is the unique solution in $\mathcal{H}$ to the partial differential equations with boundary conditions

$$\begin{cases} \dfrac{\partial H}{\partial t} - r_t y \dfrac{\partial H}{\partial y} + \dfrac{1}{2} \|\eta_t\|^2 y^2 \dfrac{\partial^2 H}{\partial y^2} = -h_1 , & t \in [\alpha, T] \quad y > 0 , \\[2mm] H(T, y) = h_2(y) & y > 0 . \end{cases} \tag{4.23}$$

We are led to make a new assumption on the utility functions.

U5   The functions $U_i \circ I_i$ and $y \to y I_i(y)$ belong in $\mathcal{H}$.

We then obtain

Proposition 4.6.2. *Under the assumptions **U1** to **U5**, the value function* $V(\alpha, x)$ *is given by*
$$V(\alpha, x) \; = \; G\left(\alpha, \mathcal{Y}(\alpha, x)\right) \; ,$$
where G is the unique solution to (4.23) *that is associated with $h_i = U_i \circ I_i$, and where $\mathcal{Y}(\alpha, \cdot)$ is the inverse function of $\mathcal{X}(\alpha, \cdot)$; $\mathcal{X}(\alpha, y)$ is defined by*

$$\mathcal{X}(\alpha, y) \; = \; \frac{\mathcal{T}(\alpha, y)}{y} \; ,$$

where $\mathcal{T}(\alpha, y)$ is the unique solution to (4.23) *associated with $h_i(y) = y\, I_i(y)$. Moreover, $\dfrac{\partial V}{\partial x}(\alpha, x) = \mathcal{Y}(\alpha, x)$.*

Thus we have reduced the problem to two Cauchy problems. We can use these results to determine the optimal wealth.

4.6.2 Optimal Wealth

Is is also possible to give an explicit expression for the optimal wealth and for the optimal portfolio. We saw in (4.14) that the wealth associated with a final wealth X_T and with a consumption $(c_t, 0 \le t \le T)$, is given by

$$X(t)R(t) \; = \; E_Q\left[X_T R(T) + \int_t^T R(s)c(s)\,\mathrm{d}s \mid \mathcal{F}_t\right].$$

The results of Sect. 4.5 show that the optimal wealth at time t is given by:

$$X^*(t)R(t) \; = \; E_Q\left[R(T)\,I_2(\mathcal{Y}(x)\zeta_T) + \int_t^T R(s)\,I_1(\mathcal{Y}(x)\zeta_s)\,\mathrm{d}s \mid \mathcal{F}_t\right].$$

Using the fact that
$$E_Q\left[Z \mid \mathcal{F}_t\right] \; = \; \frac{E_P\left[ZL_T \mid \mathcal{F}_t\right]}{E_P\left[L_T \mid \mathcal{F}_t\right]}$$

we obtain

$$X^*(t) \; = \; \zeta_t^{-1}\,E_P\left[\zeta_T\,I_2(\mathcal{Y}(x)\zeta_T) + \int_t^T \zeta_s\,I_1(\mathcal{Y}(x)\zeta_s)\,\mathrm{d}s \mid \mathcal{F}_t\right].$$

The process ζ is Markov with respect to $\mathcal{F}_t$, since its coefficients are deterministic. Hence

$$X^*(t) \;=\; \zeta_t^{-1}\, E_P\left[\zeta_T\, I_2(\mathcal{Y}(x)\zeta_T) + \int_t^T \zeta_s\, I_1(\mathcal{Y}(x)\zeta_s)\, \mathrm{d}s \,\Big|\, \zeta_t \right] ,$$

and ζ satisfies

$$\zeta_s \;=\; \zeta_t\, \zeta_s^t, \qquad t < s ,$$

where ζ_s^t is independent of ζ_t. Using the same notation as in Proposition 4.6.2, where

$$\mathcal{T}(t,y) \;=\; E_P\left(y\zeta_T^t\, I_2\left(y\zeta_T^t \right) + \int_t^T y\zeta_s^t\, I_1\left(y\zeta_s^t \right)\, \mathrm{d}s \right) ,$$

we can see that

$$X^*(t) \;=\; (\mathcal{Y}(x)\zeta_t)^{-1}\, \mathcal{T}\left(t,\, \mathcal{Y}(x)\zeta_t \right) \;=\; \mathcal{X}\left(t,\, \mathcal{Y}(x)\zeta_t \right) .$$

In addition, the function $\mathcal{T}$ satisfies the partial differential equation (cf. (4.23))

$$\frac{\partial \mathcal{T}}{\partial t} - r_t y\, \frac{\partial \mathcal{T}}{\partial y} + \frac{1}{2}\|\eta_t\|^2\, y^2\, \frac{\partial^2 \mathcal{T}}{\partial y^2} \;=\; -y\, I_1(y)$$

and the boundary condition

$$\mathcal{T}(T,y) \;=\; y\, I_2(y) .$$

Remark 4.6.3. Cox and Huang [68] applied similar kinds of working to the case of Markovian coefficients, that is to coefficients of the form $r(S_t, t)$, $\sigma(S_t, t)$.

We can also give the partial differential equation satisfied by $\mathcal{X}(t,y) = \frac{\mathcal{T}(t,y)}{y}$. We find that

$$\frac{\partial \mathcal{X}}{\partial t} - r_t \mathcal{X} + y\left(-r_t + \|\eta_t\|^2 \right) \frac{\partial \mathcal{X}}{\partial y} + \frac{1}{2}\|\eta_t\|^2\, y^2\, \frac{\partial^2 \mathcal{X}}{\partial y^2} \;=\; -I_1(y) , \qquad (4.24)$$

with the boundary condition

$$\mathcal{X}(T,y) \;=\; I_2(y) .$$

4.6.3 Obtaining the Optimal Portfolio

In all generality, the optimal portfolio is obtained by applying a predictable representation theorem. We return to the dynamic programming principle, and to the workings of Sect. 4.2.

The HJB equation has led us to a candidate to the title of optimal pair:

$$c_t^* \;=\; I_1\left\{ \frac{\partial V}{\partial x}(t, X_t^*) \right\} ,$$

$$\pi_t^* \;=\; -[a(t)]^{-1}\, (b_t - r_t \mathbf{1})\, \frac{\partial V}{\partial x}(t, X_t^*)\, \left\{ \frac{\partial^2 V}{\partial x^2}(t, X_t^*) \right\}^{-1} .$$

We saw (Proposition 4.6.2) that $\frac{\partial V}{\partial x}(t, x) = \mathcal{Y}(t, x)$. We fall back on

$$c_t^* = I_1\left(\mathcal{Y}(t, X_t^*)\right) , \qquad (4.25\text{i})$$

and we obtain

$$\pi_t^* = -[a(t)]^{-1} (b_t - r_t\mathbf{1}) \, \mathcal{Y}(t, X_t^*) \left\{\frac{\partial \mathcal{Y}}{\partial x}(t, X_t^*)\right\}^{-1} . \qquad (4.25\text{ii})$$

It remains to check that this pair is indeed optimal. To do this, it is enough to verify that X_t^* is associated with the pair (c_t^*, π_t^*) expressed in its feedback form.

Since

$$X^*(t) = \mathcal{X}(t, \mathcal{Y}(x)\zeta_t) ,$$

it follows from Itô's lemma and from (4.24) that

$$\mathrm{d}X^*(t) = (r_t X_t^* - c_t^*) \, \mathrm{d}t + \pi_t^{*T}\left[(b_t - r_t\mathbf{1})\mathrm{d}t + \sigma_t\mathrm{d}W_t\right] , \qquad (4.26)$$

and the result follows.

Remark 4.6.4. Cox and Huang [68] first determine the optimal wealth process and from there deduce the optimal portfolio. They find an analogous expression to (4.25ii), even though they use the function $\mathcal{X}$ rather than the function $\mathcal{Y}$ (it is enough to notice that $\left(\frac{\partial \mathcal{Y}}{\partial x}\right)^{-1} = \mathcal{X}'(y)$).

4.7 Market Completeness and NAO

We work under the hypotheses given in (**H**).
An arbitrage opportunity is a portfolio π such that

(i) $(\pi, 0) \in \mathcal{A}(0)$

(ii) the wealth process $X^{\pi,0}$ satisfies $P(X_T^{\pi,0} > 0) > 0$.

Proposition 4.7.1. *In the model described in Sect. 4.1, and under the hypotheses (**H**), the market is complete and there are no arbitrage opportunities.*

It is straightforward to check that there are no arbitrage opportunities in our model: as (4.11) is satisfied, we have for $(\pi, 0) \in \mathcal{A}(0)$

$$E_P(L_T R_T X_T^{\pi,0}) \leq 0 .$$

The approach developed in the proof of Proposition 4.4.3 enables us to valuate an asset. We have seen that if Z is a positive $\mathcal{F}_T$-measurable random variable such that $E_Q(ZR(T))$ is known and is equal to x, then there exists

a portfolio π such that the strategy $(\pi, 0)$ attains the final wealth Z, with an initial wealth equal to x. Thus with an initial capital of x, we could build a portfolio π enabling us to attain Z at time T. The value of this portfolio at time t would then be defined by $V(t)$ with

$$R(t)V_t = E_Q\left(ZR(T)\,|\,\mathcal{F}_t\right)$$

(using (4.14) with $c = 0$, as in this case the agent has no consumption). Taking into account (4.16), this equation can also be written as

$$V_t = \frac{E_P\left(\zeta_T Z\,|\,\mathcal{F}_t\right)}{\zeta_t} = \frac{E_P\left(ZU_2'(X_T^*)\,|\,\mathcal{F}_t\right)}{U_1'(c_t^*)} = \frac{E_P\left(ZU_1'(c_T^*)\,|\,\mathcal{F}_t\right)}{U_1'(c_t^*)} ,$$

where we are using the marginal utilities.

As any positive $\mathcal{F}_T$-measurable random variable can be attained with an admissible strategy by choosing an initial capital v, the market is complete.

Notice that, when we restrict ourselves to square-integrable portfolios, the portfolio that attains Z is unique. To see this, it is enough to check that if π_1 and π_2 both belong to $\mathcal{A}(x)$ and finance Z, then the process

$$(M_1 - M_2)(t) = \int_0^t R(s)\left(\pi_1(s) - \pi_2(s)\right)^T \sigma(s)\,\mathrm{d}\tilde{B}_s$$

is a Q-martingale that is zero at time T, and hence is itself zero. Taking its second moment, and observing that σ is invertible, we obtain for all t: $\pi_1(t) = \pi_2(t)$ a.s..

Notes

The problem of finding an optimal portfolio was first solved in Merton [272], (1971), using the methods of dynamic programming. The methods of stochastic calculus, and in particular the martingale representation theorem, have made it possible to prove the existence of an optimal strategy in a very general framework. We can consult Karatzas et al. [231], (1987), Cox and Huang [68], (1988). All these authors place themselves in a complete market framework. These theories then enable us to exhibit a hedging portfolio. The case of deterministic coefficients is particularly interesting, as the optimal portfolio–consumption pair can then be given in feedback form.

The method presented here applies to the case of complete markets with dynamics driven by processes to which the predictable representation theorem can be applied. In particular, this is the case when prices are driven by processes with jump components. Merton [275], (1976), studied the case where stock prices display such discontinuities. His work was later taken up by Aase and Øksendal [1], (1988), and then by Jeanblanc-Picqué and Pontier [215], (1990), and by Shirakawa [334] (1991).

The notion of a viscosity solution is used to prove the existence and uniqueness of the solution to the HJB equation. (Shreve and Soner [339], (1991), Zariphopoulou [378], (1994)). The main reference on the HJB equation is Fleming and Soner [156] (1993).

The case in which transaction costs or constraints intervene is much more difficult. It has been studied by Constantinides [63], (1986), and later by Davis and Normann [88], (1990), and Shreve and Soner [340], (1994). A recent approach is given in Cvitanić [32, 74], (1996, 2001).

The issue of optimal hedging in an incomplete market has been studied by Föllmer and Sondermann [163], (1986), Bouleau and Lamperton [41], (1989), Schweizer [329, 330], (1988, 1991). He and Pearson [185, 186], (1991), Shreve and Xu [342, 343], (1992), Fleming and Zariphopoulou [157], (1991), Duffie, Fleming, Soner and Zariphopoulou [117], (1997), and Cvitanić and Karatzas [75], (1993), have worked on the case where there are additional constraints on the portfolio.

More recently, the problem of optimizing the final wealth or consumption has been addressed in incomplete markets. The first results of note were obtained by Pagès [297], (1989), who characterized the set of equivalent measures to P making discounted prices into martingales, and studied the case of an optimal consumption. Subsequent results have been obtained by He and Pearson [185, 186], (1991), and then Karatzas et al. [232], (1990). He and Pagès [184], (1993), El Karoui and Jeanblanc [138], (1998) looked at the case where the agent has revenues.

Optimization under an infinite horizon and with an asymptotic criterion has been studied by Morton and Pliska [282] (1995), Foldes [160], (1990), Huang and Pagès [198], (1992), and by Konno, Suzuki and Pliska [239], (1993).

Problems involving a recursive utility function, were first touched upon by Duffie and Epstein [116], (1992), Duffie and Skiadias [124], (1994) and, using the concept of backward stochastic differential equation, by El Karoui, Peng and Quenez [143], (1997). Detemple and Zapatero [102], (1992), introduced utility functions that depend on earlier consumption (habit formation).

Finally, we point out the numerical methods developed for dealing with such problems (Sulem [349], (1992), Fitzpatrick and Fleming [154] (1991)).

Lecture notes by Cvitanić [32], (1996), Karatzas [230], (1997), Korn [240] (1998) and Karatzas and Shreve [233] give the most recent results, along with detailed and complete bibliographies.

ANNEX 4

1 The Predictable Representation Property

a) The Brownian Motion Case

Let $(B_t)_{t \geq 0}$ be a d-dimensional Brownian motion. We denote by $\mathcal{F}_t$ the filtration obtained by completing $\sigma(B_s , s \leq t)$, that is by adding the null sets.

Theorem [9] *Let $(M_t, t \geq 0)$ be a continuous $\mathcal{F}_t$-martingale that is zero at time 0. Then there exists a unique predictable process $(\phi_t, t \geq 0)$ such that*

$$M_t = \sum_{i=1}^{d} \int_0^t \phi_i(s) \, \mathrm{d}B_s^i = \int_0^t \phi(s) \, \mathrm{d}B_s$$

and

$$\int_0^t \phi^2(s) \, \mathrm{d}s < \infty \qquad P\text{-a.s.} \quad t \in [0, T] \,.$$

If moreover $E(M_T^2) < \infty$, then $E \int_0^T \phi^2(s) \, \mathrm{d}s < \infty$.

In particular, the predictable representation theorem enables us to define the stochastic integral with respect to a $\mathcal{F}_t$-martingale:

$$\int_0^t \psi(s) \mathrm{d}M_s := \int_0^t \psi(s) \, \phi(s) \, \mathrm{d}B_s \,.$$

Let $(B_t)_{t \geq 0}$ be a P-Brownian motion, and let $\mathcal{F}_t$ be its filtration. Let L_t be a Girsanov density $L_t = \exp\left\{ \int_0^t h(s) \, \mathrm{d}B_s - \frac{1}{2} \int_0^t h^2(s) \, \mathrm{d}s \right\}$, where h is bounded. We then know that $\tilde{B}_t = B_t - \int_0^t h(s) \, \mathrm{d}s$ is a Q-Brownian motion. In general, the filtration $\tilde{\mathcal{F}}_t = \sigma(\tilde{B}_s , s \leq t)$ is not equal to $\mathcal{F}_t$. However, we can show [10] that the predictable representation theorem still holds under Q:

Theorem   *Any continuous Q-$\mathcal{F}_t$ martingale can be written in the form $\int_0^t \phi(s) \, \mathrm{d}\tilde{B}_s$, where ϕ is a predictable process satisfying:*

$$\int_0^t \phi^2(s) \, \mathrm{d}s < \infty \qquad P\text{-a.s. (or } Q\text{-a.s.)} \,.$$

b) The General Case

If $(M, t \geq 0)$ is a (local) martingale, not necessarily adapted to the Brownian filtration, we can develop the definition of a stochastic integral with respect to M.

---

[9] See Karatzas and Shreve [233].
[10] Revuz and Yor [307] Exercise 1.27, Chap. VIII, Sec. 1.

Let M be a $\mathcal{G}_t$-P-martingale. We say that M has the predictable representation property for $(\mathcal{G}_t$-$P)$ if any P-$\mathcal{G}_t$-martingale that is zero at time zero, can be written as $\int_0^t \phi(s)\, dM_s$ where ϕ is a $\mathcal{G}_t$-predictable process.

Example Let B be a d-dimensional Brownian motion, and let $\mathcal{F}_t$ be its filtration.

If $M_t = \int_0^t \sigma_s\, dB_s$ where rank $\sigma(s) = d$ a.s. on $\Omega \times [0,T]$, then M is a d-dimensional process that has the predictable representation property for $(\mathcal{F}_t, P)$. Indeed, if $(X_t, t \geq 0)$ is a $\mathcal{F}_t$-martingale, then $dX_t = \phi_t\, dB_t$ where ϕ_t is predictable. As we have $dM_t = \sigma_t\, dB_t$, we can deduce that $dX_t = \phi_t \sigma_t^{-1}\, dM_t$.

2 Dynamic Programming

We will content ourselves with an intuitive approach to the results concerning dynamic programming. The interested reader can refer to Fleming and Rishel [155] or to Krylov [245] for the full proofs.

The principle behind dynamic programming is a very general principle from the theory of stochastic control, and holds under very general assumptions. The idea is as follows:

- if we use one strategy on the interval $[\alpha, t]$ and another strategy on the interval $[t, T]$, we obtain a strategy on $[\alpha, T]$,
- if we are given a strategy on the interval $[\alpha, T]$, we can decompose it into one strategy on $[\alpha, t]$ and one strategy on $[t, T]$.

Of course, the strategies need to be "glued back together" together "by continuity".

The first remark leads to

$$E\left\{ \int_\alpha^t U_1(c_s)\, ds + V(t, X_t^\alpha) \right\} \leq V(\alpha, x)\,,$$

since $V(t, X_t^\alpha)$ corresponds to using an optimal strategy on $[t, T]$, and since combining a strategy on $[\alpha, t]$ and a strategy on $[t, T]$ (in our case, the optimal strategy) produces a strategy on $[\alpha, T]$.

The equality

$$V(\alpha, x) = \sup_{(\pi, c) \in \mathcal{A}(\alpha, x)} E\left\{ \int_\alpha^t U_1(c_s)ds + V(t, X_t^{\alpha, x; \pi, c}) \right\}$$

follows from the fact that an optimal strategy on $[\alpha, T]$ yields an optimal strategy on $[t, T]$.

If we accept the principle of dynamic programming, it is just as easy to have the intuition for the HJB equation.

Firstly, we apply Itô's lemma to $V(t, X_t^{\alpha,x})$ between times α and $\alpha + h$

$$
\begin{aligned}
V(\alpha + h, X_{\alpha+h}^{\alpha,x}) - V(\alpha, x) &= \int_\alpha^{\alpha+h} \frac{\partial V}{\partial t}(s, X_s^\alpha) ds \\
&+ \int_\alpha^{\alpha+h} \frac{\partial V}{\partial x}(s, X_s^{\alpha,x}) dX_s^{\alpha,x} + \frac{1}{2} \int_\alpha^{\alpha+h} \frac{\partial^2 V}{\partial x^2}(s, X_s^{\alpha,x}) \|\pi_s^T \sigma_s\|^2 ds .
\end{aligned}
$$

We then use the principle of dynamic programming under the form

$$
V(\alpha, x) \geq E\left[\int_\alpha^{\alpha+h} U_1(c_s) ds + V(\alpha + h, X_{\alpha+h}^{\alpha,x}) \right] .
$$

If follows that

$$
\begin{aligned}
0 \geq E\Bigg\{ \int_\alpha^{\alpha+h} \Bigg[U_1(c_s) + \frac{\partial V}{\partial t}(s, X_s^{\alpha,x}) + \left\{ X_s^{\alpha,x} r_s - c_s + \pi_s^T(b_s - r_s \mathbf{1}) \right\} \\
\times \frac{\partial V}{\partial x}(s, X_s^{\alpha,x}) + \frac{1}{2} \|\pi_s^T \sigma_s\|^2 \frac{\partial^2 V}{\partial x^2}(s, X_s^{\alpha,x}) \Bigg] ds \\
+ \int_\alpha^{\alpha+h} \pi_s^T \sigma_s \frac{\partial V}{\partial x}(s, X_s^{\alpha,x}) dB_s \Bigg\} .
\end{aligned}
$$

These workings are purely **formal**: we would need to check the integrability conditions.

Next, we divide through by h, and let h tend to 0. Still only **formally**, it follows that

$$
\begin{aligned}
0 \geq U_1(c_\alpha) + \frac{\partial V}{\partial t}(\alpha, x) \\
+ \left\{ x r_\alpha - c_\alpha + \pi_\alpha(b_\alpha - r_\alpha \mathbf{1}) \right\} \frac{\partial V}{\partial x}(\alpha, x) + \frac{1}{2} \|\pi_\alpha^T \sigma_\alpha\|^2 \frac{\partial^2 V}{\partial x^2}(\alpha, x) .
\end{aligned}
$$

Hence the relationship that holds for all π, c:

$$
\begin{aligned}
0 \geq \frac{\partial V}{\partial t}(\alpha, x) \\
+ \sup_{\pi, c} \left\{ U_1(c) + \left\{ x r_\alpha - c + \pi(b_\alpha - r_\alpha \mathbf{1}) \right\} \frac{\partial V}{\partial x}(\alpha, x) + \frac{1}{2} \|\pi^T \sigma_\alpha\|^2 \frac{\partial^2 V}{\partial x^2}(\alpha, x) \right\} .
\end{aligned}
$$

The HJB equation follows as a result: if an optimal pair exists, then the inequalities above become equalities.

We are now going to state without proof the conditions under which the value function is a solution to the HJB equation.

We take a special case: that of constant coefficients, with $U_2 = 0$. We set $U = U_1$.

Let $\mathcal{L}$ be the operator defined by

$$\mathcal{L}H(t,y) \;=\; -\frac{\partial H}{\partial t}(t,y) + ry\,\frac{\partial H}{\partial y}(t,y) - \frac{1}{2}\|\eta\|^2\,y^2\,\frac{\partial^2 H}{\partial y^2}(t,y)$$

with $\eta = -\sigma^{-1}(b - r\mathbf{1})$.

We suppose that there exist functions G and S in $C^{1,3}([0,T] \times \mathbb{R}_{++}, \mathbb{R}_+)$ such that

$$\begin{cases} \mathcal{L}G(t,y) \;=\; U(I(y)) & t \in [0,T]\,, \quad y \in \mathbb{R}_{++} \\ G(T,y) \;=\; 0 & y \in \mathbb{R}_{++} \end{cases}$$

$$\begin{cases} \mathcal{L}S(t,y) \;=\; y\,I(y) & t \in [0,T]\,, \quad y \in \mathbb{R}_{++} \\ S(T,y) \;=\; 0 & y \in \mathbb{R}_{++} \end{cases}$$

and such that G, $\frac{\partial G}{\partial y}$, S and $\frac{\partial S}{\partial y}$ satisfy polynomial growth conditions of the type

$$\max_{0 \le t \le T} H(t,y) \;\le\; M(1 + y^{-\lambda} + y^{\lambda})\,, \qquad y \in \mathbb{R}_{++}$$

where M and λ are strictly positive constants. We also assume that U is of class C^2.

It can then be shown[11] that the following result holds.

Theorem *Under the previous assumptions, the value function is of class $C^{1,2}([0,T] \times \mathbb{R}_{++})$ and satisfies the HJB equation (4.5) as well as the boundary conditions (4.6).*

---

[11] See Karatzas and Shreve [233], Chap. 5 Theorem 8.22.

5

The Yield Curve

The uncertainty attached to the future movements of interest rates is an important part of the theory of financial decision making. Most agents are risk-averse, and risk is linked in particular to interest rates. Investment decisions and asset/liability management are often very sensitive to perturbations of the yield curve. To hedge interest rate risk, the markets use increasingly complicated financial products (forward contracts, futures contracts, options on contracts). These constitute the forward markets.

It is therefore important to understand the factors that drive interest rates, to model the yield curve, to analyze financial instruments such as options on zero coupon bonds, and to develop strategies for hedging interest rate risk.

5.1 Discrete-Time Model

We adopt a model, in which transactions take place at set times indexed by integers, and entail no cost. We work under the assumption of no arbitrage, and of the existence of a martingale measure.

Someone who borrows one euro at time n, will have to pay back $F(n, N)$ euros at time N, when the loan is due to be repaid. We call $F(n, N)$ the forward price of 1 euro.

If $S(n)$ is the price of a financial product given in units of time n, the forward price of S is expressed in units of time N as $S_F(n) = S(n)F(n, N)$.

Definition 5.1.1. *A zero coupon bond with maturity N is a security that pays one euro at time N and does not generate any cash flows before N.*

The price at time n of a zero coupon bond with maturity N ($n \leq N$) is denoted by $P(n, N)$. It is the price of one euro paid out at time N. We have $P(N, N) = 1$. The forward price of 1 euro and the price of a zero coupon bond are linked by the relationship $F(n, N) = [P(n, N)]^{-1}$ (from the assumption of NAO).

We define the instantaneous forward rate by $f(n, N) := \ln \dfrac{P(n, N)}{P(n, N+1)}$,
and the spot rate by $r(n) := f(n, n)$. We have $P(n, n+1) = e^{-r(n)}$.

The term structure of rates is given by the study either of the family $P(n, .)$, or of the family $f(n, .)$. One approach involves studying the dynamics of the yield curve, that is, expressing $f(n, N)$ as a function of today's curve $f(0, N)$.

In a deterministic model, the relationship between rates of different maturities must be such that $P(n, n+1)P(n+1, N) = P(n, N)$, for all n and N, so as to avoid arbitrage opportunities between the zero coupon bonds of different maturities: to obtain 1 euro at time N, we must pay $P(n, N)$ at time n; we could also pay $P(n+1, N)$ at time $n+1$, which comes down to paying $P(n, n+1)P(n+1, N)$ at time n. In this case, we have in particular $f(0, N) = f(n, N) = r(N)$, $\forall n \le N$. In a deterministic model, the instantaneous forward rate depends only on its maturity.

To study a model in an uncertain world, Ho and Lee [193] make the following assumptions:

- the price of zero coupon bonds depends only on the number of up-movements in rates between times 0 and n,
- the price at time $n+1$ of a zero coupon with maturity N differs from the price obtained in the deterministic case by the addition of a random perturbation function h,
- the perturbation function h is only a function of time to maturity and of the behaviour of the price between times n and $n+1$.

At time n, there are $n+1$ states of the world. For each of these states at time n, there are two possible states at time $n+1$, depending on whether the price increases or decreases. The price at time n of a zero coupon bond with maturity N, is $P(n, N; j)$, where j refers to the number of increases before time n. The condition imposed on the prices corresponding to different maturities is

$$P(n+1, N) = \frac{P(n, N)}{P(n, n+1)}\, h(n+1, N) \,; \quad n+1 \le N \,,$$

where h is a perturbation function. In more detail:

$$P(n+1, N; j) = \frac{P(n, N; j)}{P(n, n+1; j)}\, h(0\,; n+1, N)\,,$$

$$P(n+1, N; j+1) = \frac{P(n, N; j)}{P(n, n+1; j)}\, h(1\,; n+1, N)\,.$$

We suppose that $h(1; n+1, N) \ge h(0; n+1, N)$.

The random nature of the perturbation becomes clearer if we introduce the random variable Y_{n+1} that is worth 1 if the prices increases between times n and $n+1$, and 0 otherwise. We assume the $(Y_n, n \geq 1)$ to be independent and identically distributed[1] under the risk-neutral probability measure.

The move in $P(n, N)$ is given by

$$P(n+1, N) = \frac{P(n, N)}{P(n, n+1)} h(Y_{n+1}; n+1, N) \tag{5.1}$$

where $h(.\,; n+1, N)$ depends only on $N - n$. We have $h(.\,; N, N) = 1$.

As was noted previously, in world of certainty, the assumption of no arbitrage implies that h is identically 1. In our model of an uncertain world, we need to impose a condition that precludes all arbitrage opportunities between zero coupon bonds of different maturities. This will translate into a specification of the perturbation function.

We will now prove the following result.

Theorem 5.1.2. *Under the assumption of no arbitrage opportunities, there exists $\delta > 1$ and $\pi \in [0, 1]$ such that*

$$\pi h(0; n, N) + (1 - \pi) h(1; n, N) = 1 \tag{5.2}$$

and

$$h(0; n, N) = \frac{1}{\pi + (1 - \pi) \delta^{N-n}} \;;\;\; h(1; n, N) = \frac{\delta^{N-n}}{\pi + (1 - \pi) \delta^{N-n}} \,. \tag{5.3}$$

Proof. We saw in Chaps. 1 and 2, that when the state space is finite, the assumption of no arbitrage is equivalent to the existence of a risk-neutral probability measure, under which discounted prices are martingales.

The discount factor can be expressed as a function of the prices of zero coupon bonds: the factor that discounts prices given at time n corresponds to today's value for one euro at time n, and is given by:

$$A(0, n) = \prod_{k=1}^{n} P(k-1, k) \,.$$

Hence the sequence $(A(0, n)P(n, N), n \leq N)$ is a martingale, for any value of N.

The relationship (5.1) can written

$$A(0, n+1)P(n+1, N) = A(0, n)P(n, N)h(Y_{n+1}; n+1, N) \,.$$

Taking conditional expectations relative to $\mathcal{F}_n$, we find that the conditional expectation of $h(Y_{n+1}; n+1, N)$ is equal to 1, i.e.,

---

[1] See El Karoui and Saada [148] for a generalization.

$$\pi_n h(0\,;n+1,N) + (1-\pi_n)h(1\,;n+1,N) \;=\; 1\,,$$

where π_n is equal to $E(Y_{n+1}=0\,|\mathcal{F}_n)$. The number π_n does not depend on n, as the Y_n are independent and identically distributed.

In our model, we have assumed that the price at time n does not depend on the path taken. We can calculate $P(n+2,N)$, assuming that the price rose between n and $n+1$, and then fell. Then we find that

$$
\begin{aligned}
P(n+2,N;j+1) &= \frac{P(n,N;j)\,h(1\,;n+1,N)\,h(0\,;n+2,N)}{P(n,n+1;j)\,P(n+1,n+2;j+1)}\\
&= \frac{P(n,N;j)\,h(1\,;n+1,N)\,h(0;n,N)}{P(n,n+2;j)\,h(1\,;n+1,n+2)}\;.
\end{aligned}
$$

Similarly, we can carry out the calculations, assuming that the price first fell and then rose:

$$P(n+2,N;j+1) = \frac{P(n,N;j)\,h(1\,;n+2,N)\,h(0\,;n+1,N)}{P(n,n+2;j)\,h(0\,;n+1,n+2)}\;.$$

By equating two results, we find that:

$$
\begin{aligned}
h(1\,;n+1,N)&h(0\,;n+2,N)\,h(0\,;n+1,n+2)\\
&= h(1\,;n+2,N)\,h(0\,;n+1,N)\,h(1\,;n+1,n+2)\,.
\end{aligned}
$$

Using (5.2) and the assumption that $h(.\,;n,N)$ depends only on $(N-n)$, we find, setting $s=N-n-1$ and $h(.\,;s)=h(.\,;t,t+s)$,

$$
\begin{aligned}
(1-\pi)\,(1-\pi h(0\,;s))\,&h(0;s-1)h(0\,;1)\\
&= (1-\pi h(0\,;s-1))\,h(0\,;s)\,(1-\pi h(0\,;1))\,.
\end{aligned}
$$

We can write this expression as $\quad \dfrac{1}{h(0\,;s+1)} = \dfrac{\delta}{h(0\,;s)}+\gamma\,,\quad$ where δ and γ are defined by

$$h(0\,;1) = \frac{1}{\pi+(1-\pi)\,\delta}\quad\text{and}\quad \gamma = \frac{\pi(h(0\,;1)-1)}{(1-\pi)\,h(0\,;1)}\;.$$

We obtain the expression (5.3) by solving this difference equation, using the condition $h(0\,;0)=1$. The inequality $\delta>1$ then follows from $h(1;s)\geq h(0;s)$. $\qquad\square$

From here, we will deduce the movements in the spot rate

$$r(n) = -\ln P(n,n+1)\,.$$

We will show that the number of up-moves in the price is an explanatory variable, and we will express the price of zero coupons and the forward and spot rates as functions of this variable. More precisely, we will prove:

Theorem 5.1.3. *Let $P(0, N)$ be today's curve of zero coupon bond prices, and let $f(0, N)$ be the spot instantaneous forward rate curve. We have*

$$P(n, N) = \frac{P(0, N)}{P(0, n)} \prod_{j=1}^{n} \frac{h(Y_j \,; j, N)}{h(Y_j \,; j, n)}$$

and

$$f(n, N) = f(0, N) + \ln(1/\delta) \sum_{j=1}^{n} Y_j + \ln \frac{\pi + (1 - \pi)\, \delta^N}{\pi + (1 - \pi)\, \delta^{N-n}}$$

$$r(n) = f(0, n) + \ln(1/\delta) \sum_{j=1}^{n} Y_j + \ln(\pi + (1 - \pi)\, \delta^n) \,.$$

Proof. Our choice of model entails

$$P(n, N) = \frac{P(0, N)}{P(0, n)} \prod_{j=1}^{n} \frac{h(Y_j \,; j, N)}{h(Y_j \,; j, n)} \,.$$

Let us set $\psi(n, N) = \dfrac{h(1 \,; n, N)}{h(0 \,; n, N)} = \delta^{N-n}$. Noticing that

$$h(Y_j; j, N) = \psi(j, N)^{Y_j} h(0; j, N) \,,$$

we obtain an expression for the instantaneous forward rates:

$$f(n, N) = \ln \frac{P(n, N)}{P(n, N + 1)} = \ln \frac{P(0, N)}{P(0, N + 1)}$$

$$+ \sum_{j=1}^{n} Y_j \left(\ln \psi(j, N) - \ln \psi(j, N + 1) \right)$$

$$+ \sum_{j=1}^{n} \left(\ln h(0 \,; j, N) - \ln h(0 \,; j, N + 1) \right) \,.$$

After simplification, this equals

$$\ln \frac{P(0, N)}{P(0, N + 1)} + \ln(1/\delta) \sum_{j=1}^{n} Y_j + \ln \frac{\pi + (1 - \pi)\, \delta^N}{\pi + (1 - \pi)\, \delta^{N-n}} \,.$$

$\square$

Remark 5.1.4. The variance of the spot rate is given by $(\ln \delta)^2 nq(1-q)$, where q is the expectation of Y_j, and it converges to infinity with n. This is an important drawback to the model. Furthermore, the forward and spot rates can become negative[2]. Therefore this model is not satisfactory.

---

[2] Sandmann and Sonderman [321] have studied a model that precludes this possibility.

5.2 Continuous-Time Model

Two main approaches are used in continuous time. The first involves modeling the prices of zero coupon bonds in a way that is consistent with the assumption of NAO, and thence deducing an expression for the spot rate. The second approach uses the spot rate as an explanatory variable. We present these two approaches and we show, on a few examples, how these models lead to the valuation of interest rate products.

5.2.1 Definitions

We give the same definitions as in discrete time. A *zero coupon bond with maturity* T is a security that pays one euro at time T, and provides no other cash flows between times t and T. We assume that for all T, there exists a zero coupon bond with maturity T[3].

The price at time t of a zero coupon bond with maturity T is denoted by $P(t, T)$. We have $P(T, T) = 1$.

If $S(t)$ is the price of a financial asset in units of time t, we call the *forward price* of S, its price expressed in units of time T, i.e., $S_F(t) = \dfrac{S(t)}{P(t, T)}$.

We introduce the *yield to maturity* at time t, $Y(t, T)$, defined by

$$P(t, T) = \exp\left[-(T - t)Y(t, T)\right] .$$

The *forward spot rate* at time t with maturity T is

$$f(t, T) = -\left[\frac{\partial \ln P(t, \theta)}{\partial \theta}\right]_{\theta = T} .$$

Thus we have

$$Y(t, T) = \frac{1}{T - t} \int_t^T f(t, u)\, du \quad \text{and} \quad P(t, T) = \exp\left(-\int_t^T f(t, u)\, du\right) .$$

The *instantaneous spot rate* is

$$r(t) = \lim_{T \to t} Y(t, T) := -\left[\frac{\partial \ln P(t, T)}{\partial T}\right]_{T = t} = f(t, t) .$$

The yield curve is given by the function $\theta \to Y(t, \theta)$.

The discount factor is

---

[3] In practice, this assumption does not hold.

$$R(t) := \exp\left(-\int_0^t r(s)\,\mathrm{d}s \right) .$$

At each time t, we can observe a range of rates: the family $s \to Y(t, s+t)$ of interest rates with maturities $s + t$, observed at time t. We would like to study the behaviour of the curve $Y(t, \theta)$ as a function of today's yield curve, which is $Y(0, \theta)$.

In a deterministic model, we must have

$$P(t, T) = P(t, u) P(u, T) , \quad \forall t \le u \le T ,$$

to preclude arbitrage opportunities. Thence, under the assumption of differentiability, we deduce the existence of a real-valued function r such that $P(t, T) = \exp\left(-\int_t^T r(s)\,\mathrm{d}s \right)$. As in the discrete model, we can check that $f(t, T) = f(0, T) = r(T) , \quad \forall t \le T$ and that $Y(t, T) = \frac{1}{T-t} \int_t^T r(u)\,\mathrm{d}u$. The yield to maturity is thus the average value of the spot rate.

In a stochastic model, as usual we take a probability space equipped with a filtration $\mathcal{F}_t$, which we assume to be a Brownian filtration. We assume that at time t, the price $P(t, .)$ of zero coupon bonds is known, i.e., the $P(t, .)$ are $\mathcal{F}_t$-measurable variables. To give the assumption of NAO an explicit form, we assume that the processes $P(., T)$ are positive, adapted, and continuous, and that $P(t, T)$ is continuously differentiable with respect to T.

We assume that there exists a probability measure Q under which discounted prices are square-integrable martingales: in particular, under Q, the process $(R(t)P(t, T),\ t \ge 0)$ is a martingale. This property holds for all T, in order to preclude arbitrage opportunities between products of different maturities.

The property leads to some interesting results. First of all, since $P(T, T) = 1$, it must be the case that $P(0, T) = E_Q(R(T))$, and that

$$P(t, T) = E_Q\left[\exp\left(-\int_t^T r(u)\,\mathrm{d}u \right) \Big| \mathcal{F}_t \right] . \tag{5.4}$$

Remark 5.2.1. Using the notation $L_T = \frac{\mathrm{d}Q}{\mathrm{d}P}\big|_{\mathcal{F}_t}$ and $L_t = E_P(L_T | \mathcal{F}_t)$, the predictable representation theorem implies[4] that there exists an adapted process q such that

$$L_t = \exp\left(\int_0^t q(s)\,\mathrm{d}B_s - 1/2 \int_0^t q^2(s)\mathrm{d}s \right) .$$

Using the predictable representation theorem again, we deduce that for each maturity T, there is an adapted process $\sigma(t, T)$ such that under P,

---

[4] See Heath–Jarrow–Morton [189] or Lamberton–Lapeyre [250].

$$\mathrm{d}P(t,T) = P(t,T)\left(\left(r(t) - q(t)\sigma(t,T)\right)\mathrm{d}t + \sigma(t,T)\,\mathrm{d}\tilde{B}_t\right),$$

where $\tilde{B}$ is a P-$\mathcal{F}_t$-Brownian motion. The quantity $q(t)\sigma(t,T)$ is the difference between the riskless rate and the average rate of return on the zero coupon bond. The process q is called the risk premium.

5.2.2 Change of Numéraire

The Forward Measure

The value at time t of a deterministic cash flow F received at time T is

$$FP(t,T) = F\,E_Q\left[\exp\left(-\int_t^T r(u)\,\mathrm{d}u\right)\Big|\mathcal{F}_t\right].$$

If the cash flow is random, its value at time t is

$$E_Q\left[F\,\exp\left(-\int_t^T r(u)\,\mathrm{d}u\right)\Big|\mathcal{F}_t\right].$$

We can give an interpretation of this formula, by introducing the notation $F_c P(t,T)$ for the *certainty equivalent* of F, defined by

$$F_c = \frac{1}{P(t,T)}E_Q\left[F\,\exp\left(-\int_t^T r(u)\,\mathrm{d}u\right)\Big|\mathcal{F}_t\right].$$

We will re-write this last equality using a change of probability measure.

By the assumption of NAO, the process $R(t)P(t,T)$ is a Q-martingale, so its expectation is constant and equal to $P(0,T)$.

For all T, the process $\left(\zeta_t^T := \dfrac{R(t)P(t,T)}{P(0,T)}, t \geq 0\right)$ is a positive Q-martingale with expectation 1. Therefore, we can use ζ_t^T as the density of change of measure or Radon-Nicodym density. Let Q_T be the probability measure defined[5] on $(\Omega, \mathcal{F}_T)$ by $Q_T(A) = E_Q(\zeta_t^T 1_A)$, for all $A \in \mathcal{F}_t$. When T is fixed, we will use the notation: $\zeta_t = \zeta_t^T$.

Definition 5.2.2. *The probability measure* Q_T *defined on* $\mathcal{F}_t$ *by* $\dfrac{\mathrm{d}Q_T}{\mathrm{d}Q} = \zeta_t^T$ *is called the T-forward measure.*

With this notation,
$$F_c = E_{Q_T}(F|\mathcal{F}_t).$$

---

[5] As described in Sect. 3 of the annex to Chap. 3, we check that Q_T is well-defined.

When F is the value of a security, the certainty equivalent F_c is called the forward price of F.

When r is deterministic, $Q_T = Q$.

The measure Q_T is the martingale measure that corresponds to choosing the zero coupon bond with maturity T as numéraire, as the property below makes explicit.

Property 5.2.3. *If $(X_t, t \geq 0)$ is a price process, its forward price $(X_t/P(t,T), t \geq 0)$ is a martingale under Q_T.*

Proof. Take T to be fixed. Let $(X_t, t \leq T)$ be a price process. By definition of the martingale measure Q, the discounted price process $X_t R(t)$ is a Q-martingale. We want to show that $(X_t/P(t,T); t \leq T)$ is a Q_T-martingale. According to the formula for conditional expectations under a change of measure[6], we have

$$E_{Q_T}\left[\frac{X_t}{P(t,T)}\,|\mathcal{F}_s\right] \;=\; \frac{E_Q\left[\frac{X_t\zeta_t}{P(t,T)}\,|\mathcal{F}_s\right]}{E_Q[\zeta_t\,|\mathcal{F}_s]} \;=\; \frac{E_Q\left[X_t R_t\,|\mathcal{F}_s\right]}{P(0,T)\zeta_s} \;=\; \frac{X_s}{P(s,T)}\;.$$

$\square$

Forward and Futures Contracts

We now define these two financial products.

A *forward contract* with maturity T and with as underlying, an asset whose price at time t is V_t, is a contract that entitles its holder to buy or to sell the asset at time T, at a price that is set when the contract is signed (at time t). This contract does not entail any cash flows when it reaches maturity.

The price of the contract refers to the price G_t at which it was agreed (at t, when the contract was signed) that the asset would traded at time T.

A *futures contract* is a forward contract with continuous readjustment. More precisely, a futures contract with expiry T, written on an asset whose price is V_t at time t, is a contract which sets a price (the price of the contract) that provides the basis for "margin calls". Each player gives a guarantee, in the form of a deposit that is placed in a current account to his name. At each day's close of trade, each player's position is readjusted. If there is a loss, the player must finance it; if there is a gain, his account will be credited. The underlying is delivered at time T, at the price of time T, rather than at the price agreed upon in the contract. Let us give an example taken from Aftalion and Poncet [2].

---

[6] Sect. 4 of the annex to Chap. 3.

Example 5.2.4. On October 15 1990, an investor sold a contract on a notional bond[7], with maturity in December 1990, at a rate of 97.52. As the contract's nominal value was 500 000 FRF, and as a change in rate of 0.01% corresponded to 50 FRF, the following table gives the margin changes generated by this position, up until the day before the position was brought to a close (a minus sign indicates a loss, that is a margin contribution, and a plus sign indicates a gain, hence a margin restitution).

Date	Closing rate	Margin
15/10	97.86	$-1\,700$ FRF
16/10	97.70	$+\,800$ FRF
17/10	97.10	$+3\,000$ FRF
18/10	96.74	$+1\,800$ FRF
19/10	96.22	$+2\,600$ FRF
22/10	96.50	$-1\,400$ FRF
23/10	96.06	$+2\,200$ FRF

The player bought back the contract on October 24, at the rate of 96.00. The margin changes left him a net sum of 7 300 FRF, to which we must add the results of his trades on October 24, that is 300 FRF – the difference between the rate at which the contract was bought back (96.00) and the closing rate on October 23 (96.06). Of course, the total effect of margin changes and of the gain (or loss) that occurs on the day that the position is closed out (in this case a total of 7 600 FRF) corresponds to the difference between the initial ask price[8] and the final bid price[9], i.e., to $500\,000(97.52-96.00)/100 = 7\,600$ FRF.

Proposition 5.2.5. *The price at time t of a forward contract with maturity T, on an asset whose price process is given by V(s), is*

$$G(t) = E_{Q_T}(V(T)|\mathcal{F}_t) \,.$$

The price of a futures contract (the futures price) is

$$H(t) = E_Q(V(T)|\mathcal{F}_t) \,.$$

If r is deterministic, the forward and futures prices are equal.

Proof. To obtain the price of a forward contract, we use the fact that the process X_t defined by

---

[7] Notional bonds provide the basis for contracts on the MATIF, the French futures exchange, which is the context of this example.

[8] The price that a seller in the market asks for.

[9] The price that a buyer in the market is prepared to bid.

$$X_t = 0 \quad \text{for } t < T$$
$$X_T = V_T - G_t \quad \text{at T,}$$

is a price process.

A futures contract can be characterized as an asset whose dividend process is given by the process H, and whose price process is zero. We seek to calculate the futures price associated with obtaining V_T, that is, the value at time t of a dividend process H such that $H_T = V_T$. Let $R_s^t = \exp\left(-\int_t^s r(u)\,\mathrm{d}u\right)$. The cumulative dividend process (see Sect. 3.4.4) $H_s^R = \int_t^s R_u^t \mathrm{d}H_u$ is a Q-martingale. If we assume r to be positive and bounded, then the process R is bounded above and below, and from the equality $\mathrm{d}H_s^R = R_s^t \mathrm{d}H_s$, we deduce that H is a Q-martingale. Hence $H_t = E_Q(H_T|\mathcal{F}_t)$, as required. □

The Spot Rate

Proposition 5.2.6. *Let T be fixed. The forward spot rate $(f(t,T), t \leq T)$ is a Q_T-martingale*

$$f(t,T) = E_{Q_T}[r_T\,|\mathcal{F}_t]\ , \quad t \leq T\ , \tag{5.5}$$

that equals the price of a forward contract written on the spot rate.

In particular, $f(0,T) = E_{Q_T}(r(T))$ is the price at time 0 of a forward contract written on the spot rate with the same maturity T.

The price of a zero coupon bond can be expressed as a function of the spot rate by

$$P(t,T) = \exp\left(-\int_t^T E_{Q_s}[r_s\,|\mathcal{F}_t]\,\mathrm{d}s\right)\ , \quad t \leq T\ . \tag{5.6}$$

Proof. By definition, the forward rate $f(t,T)$ is equal to

$$-\lim_{h \to 0} \frac{P(t,T+h) - P(t,T)}{h\,P(t,T)}\ .$$

The process $P(t,T+h)$ is a price process, therefore from Property 5.2.3, $\dfrac{P(t,T+h)}{P(t,T)}$ is a Q_T-martingale. As a result

$$f(t,T) = -\lim_{h \to 0} \frac{1}{h} E_{Q_T}\left\{(P(T,T+h) - 1)\,|\mathcal{F}_t\right\}\ ,$$

so that $f(t,T) = E_{Q_T}[r_T|\mathcal{F}_t]$. Under Q_T, $f(t,T)$ is the best L^2 estimator of $r(T)$, for $\mathcal{F}_t$ given. By definition, $\ln P(t,T) = -\int_t^T f(t,s)\mathrm{d}s$, hence

$$P(t,T) = \exp\left(-\int_t^T E_{Q_s}[r_s\,|\mathcal{F}_t]\,\mathrm{d}s\right)\ .$$

□

Exercise 5.2.7. Let $P^f(t, T, T^*) := P(t, T^*)/P(t, T)$ be the forward price at time t, in units of time T, of a zero coupon bond with maturity $T^* > T$. Using Jensen's inequality, show that the forward spot rate $Y^f(t, T, T^*)$ defined by

$$P^f(t, T, T^*) = \exp\left(-(T^* - T)Y^f(t, T, T^*)\right)$$

is a Q_T-submartingale.

Forward Price, Futures Price

We can specify the relationship between the forward price and the futures price. By definition of Q_T, for any product $Z \in L^2(\Omega, \mathcal{F}_T, Q)$,

$$E_{Q_T}[Z \,|\mathcal{F}_t] \;=\; E_Q\left[Z \frac{\zeta(T)}{\zeta(t)} \,|\mathcal{F}_t\right] .$$

Moreover, Proposition 5.2.6 and the properties of the exponential show that

$$\frac{\zeta_T}{\zeta_t} \;=\; \frac{R_T}{R_t P(t, T)} = \exp\left[\int_t^T (f(t, u) - r(u))\, \mathrm{d}u\right]$$

$$= 1 + \int_t^T (f(t, u) - r(u)) \exp\left\{\int_t^u (f(t, v) - r(v))\mathrm{d}v\right\} \mathrm{d}u$$

$$= 1 + \int_t^T \frac{\zeta_u^u}{\zeta_t^u} (f(t, u) - r(u))\, \mathrm{d}u .$$

(the equality between the third and fourth terms is obtained using the formula

$$\exp\left(\int_t^T g(u)\, \mathrm{d}u\right) = 1 + \int_t^T g(u) \left(\exp\int_t^u g(v)\, \mathrm{d}v\right) \mathrm{d}u ,$$

which is established by differentiating with respect to T).

When the variables Z_i are $\mathcal{F}_u$-measurable, we use the notation

$$\mathrm{Cov}_{Q_u}(Z_1, Z_2 \,|\mathcal{F}_s) = E_{Q_u}(Z_1 Z_2 \,|\mathcal{F}_s) - E_{Q_u}(Z_1 \,|\mathcal{F}_s) E_{Q_u}(Z_2 \,|\mathcal{F}_s)$$

$$= E_{Q_u}\left((Z_1 - E_{Q_u}(Z_1 \,|\mathcal{F}_s)) Z_2 \,|\mathcal{F}_s\right) ,$$

for their conditional covariance with respect to $\mathcal{F}_s$, under Q_u.

We write $Z_u = E_Q(Z|\mathcal{F}_u)$. Using

$$E_Q\left(Z \frac{\zeta_u^u}{\zeta_t^u} (f(t, u) - r(u)) \,|\mathcal{F}_t\right) = E_{Q_u}(Z_u (f(t, u) - r(u)) \,|\mathcal{F}_t) ,$$

and taking into account the fact that $f(t, u) = E_{Q_u}(r(u)|\mathcal{F}_t)$, we obtain

$$E_{Q_T}(Z \,|\mathcal{F}_t) = E_Q(Z \,|\mathcal{F}_t) - \int_t^T \text{Cov}_{Q_u}(Z_u, r(u) \,|\mathcal{F}_t) \, du \qquad (5.7)$$

In particular,

$$E_{Q_T}(Z) = E_Q(Z) - \int_0^T \text{Cov}_{Q_u}(Z_u, r(u)) \, du \ .$$

Proposition 5.2.8. *The price at time 0 of a forward contract with maturity T, written on Z, is the price at time 0 of a futures contract with the same characteristics, minus a covariance bias.*

5.2.3 Valuation of an Option on a Coupon Bond

The price of a European option with payoff $h(T)$ at time T is given by

$$C(t) \ = \ R_t^{-1} E_Q\left[h(T)R_T \,|\mathcal{F}_t\right] \ .$$

Let us consider an option with maturity T on a product that makes deterministic payments F_n at times T_n, $\quad T < T_n < T_{n+1}$, and let $V(t) = \sum_{n=1}^N F_n P(t, T_n)$.

Theorem 5.2.9. *The price of a European option with strike K and maturity T written on an asset that makes the payouts F_n at times T_n is*

$$C(0) \ = \ \sum_{n=1}^N F_n P(0, T_n) Q_n\left[V(T) > K\right] - K P(0, T) Q_T\left[V(T) > K\right] \ ,$$

where Q_n is the T_n-forward measure.

Proof. By definition,

$$C(0) = E_Q(R_T(V(T) - K)^+) = E_Q\left[R_T\left(\sum_{n=1}^N F_n P(T, T_n) - K\right)^+\right] ,$$

which can be written

$$C(0) = \sum_{n=1}^N F_n E_Q\left[R_T P(T, T_n) 1_{\{V(T) > K\}}\right] - K E_Q\left[R_T 1_{\{V(T) > K\}}\right] \ .$$

By definition of Q_n, we have

$$E_Q\left[R_T P(T, T_n) 1_{\{V(T) > K\}}\right] = P(0, T_n) E_{Q_n}\left[1_{\{V(T) > K\}}\right] \ ,$$

which produce the result required. $\qquad\square$

Exercise 5.2.10. Show that if r is constant, then the price of a European option with maturity T, strike K and with, as the underlying, the forward contract F_t on an asset with constant volatility σ, is:

$$C_t = e^{-r(T-t)} \left(F_t \Phi(d_1) - K \Phi(d_2) \right) ,$$

with $d_1 = \dfrac{\ln(F_t K^{-1}) + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}$, $d_2 = d_1 - \sigma\sqrt{T-t}$. This is known as Black's formula.

5.3 The Heath–Jarrow–Morton Model

This model specifies the dynamics of zero coupon bonds under the assumption of NAO. We then study the evolution of the yield curve.

5.3.1 The Model

We suppose that for any maturity, the dynamics of the zero coupon bond with maturity T are given by

$$\mathrm{d}P(t,T) = P(t,T) \left((r(t) - q(t)\sigma(t,T))\mathrm{d}t + \sigma(t,T)\mathrm{d}\tilde{B}_t \right) ,$$

where $\tilde{B}$ is a Brownian motion under the historical probability measure P, and where σ is a matrix of adapted, bounded coefficients, which are continuous in t and continuously differentiable with respect to T.

The process q is the risk premium (see Remark 5.2.1), and is assumed to be bounded. The instantaneous return of the zero coupon bond is $r(t) - q(t)\sigma(t,T)$.

As $P(T,T) = 1$, we assume that $\sigma(T,T) = 0$.

Under the risk-neutral measure Q, the price dynamics of the zero coupon bond are:

$$\mathrm{d}P(t,T) = P(t,T) \left(r(t)\mathrm{d}t + \sigma(t,T)\mathrm{d}B_t \right) \tag{5.8}$$

where B is defined by $\mathrm{d}B_t = \mathrm{d}\tilde{B}_t - q(t)\mathrm{d}t$, and is a Q-$\mathcal{F}_t$-Brownian motion.

Let $I(t,T) := \sigma(t,T)\sigma^*(t,T)$ (where $*$ denotes the transpose), and let us assume that $E_Q \left(\exp \frac{1}{2} \int_0^T I(s,T)\mathrm{d}s \right) < \infty$ and that $E_Q \left| \frac{\partial}{\partial t}\sigma(s,t) \right|^2 < \infty$. We then have:

$$P(t,T) = P(0,T) \exp \left\{ \int_0^t \sigma(s,T)\mathrm{d}B_s - \frac{1}{2}\int_0^t I(s,T)\mathrm{d}s + \int_0^t r(s)\mathrm{d}s \right\} \tag{5.9}$$

and

$$\zeta_t^T = \exp \left\{ \int_0^t \sigma(s,T)\mathrm{d}B_s - \frac{1}{2}\int_0^t I(s,T)\mathrm{d}s \right\} . \tag{5.10}$$

Theorem 5.3.1. *The forward spot rates are given by*

$$f(t, T) = f(0, T) - \int_0^t \frac{\partial \sigma}{\partial T}(s, T) \, dB_s + \int_0^t \frac{\partial \sigma}{\partial T}(s, T) \sigma^*(s, T) \, ds$$

and the spot rate satisfies

$$r(t) = f(0, t) - \int_0^t \frac{\partial \sigma}{\partial T}(s, t) \, dB_s + \int_0^t \frac{\partial \sigma}{\partial T}(s, t) \sigma^*(s, t) \, ds \; .$$

Proof. Formula (5.9) can be written as

$$\ln P(t, T) = \ln P(0, T) + \int_0^t r(s) \, ds - \frac{1}{2} \int_0^t I(s, T) \, ds + \int_0^t \sigma(s, T) \, dB_s \; ,$$
$$t \leq T \; . \quad (5.11)$$

We differentiate this expression with respect to T, to get

$$\frac{\partial}{\partial T} \ln P(t, T) = \frac{\partial}{\partial T} \ln P(0, T) - \frac{1}{2} \int_0^t \frac{\partial}{\partial T} I(s, T) \, ds + \int_0^t \frac{\partial}{\partial T} \sigma(s, T) \, dB_s \; ,$$

and hence the expression for the forward spot rate.

It now remains to use

$$r(t) = \left[-\frac{\partial}{\partial T} \ln P(t, T) \right]_{T=t} \; ,$$

to obtain the expression for the spot rate. The forward spot rate is a biased estimator of the spot rate, where the bias is independent of the volatility. In the historical world (under P), the risk premium must be taken into account. □

Remark 5.3.2. Using (5.10), we can see that $d\zeta_t = \sigma(t, T) \zeta_t \, dB_t$. Applying Girsanov's theorem, we find that the vector $B_t^{Q_T} := B_t - \int_0^t \sigma(s, T) \, ds$ is a Q_T-Brownian motion: once again we find that under Q_T, $f(t, T) = f(0, T) - \int_0^t \frac{\partial \sigma}{\partial T}(s, T) \, dB_s^{Q_T}$ is a martingale.

Exercise 5.3.3. Using the fact that

$$E_Q\left(r(T) \,|\, \mathcal{F}_t\right) = f(0, T) - \int_0^t \frac{\partial \sigma}{\partial T}(s, T) \, dB_s + \int_0^t \frac{\partial \sigma}{\partial T}(s, T) \sigma^*(s, T) \, ds$$
$$+ \frac{1}{2} \int_t^T E_Q\left(\frac{\partial I}{\partial T}(s, T) \,|\, \mathcal{F}_t\right) ds \; ,$$

show that

$$f(t,T) = E_Q\left(r(T)\,|\mathcal{F}_t\right) + \frac{1}{2}\int_t^T E_Q\left(\frac{\partial I}{\partial T}(s,T)\,|\mathcal{F}_t\right)\,\mathrm{d}s\ .$$

Using covariance calculations, establish the fact that

$$f(t,T) = E_Q\left(r(T)\,|\mathcal{F}_t\right) + \mathrm{Cov}_Q\left(r(T),\int_t^T \sigma(s,T)\,\mathrm{d}B_s\,|\mathcal{F}_t\right)\ ,$$

and next, using (5.11), show that

$$f(t,T) = E_Q\left(r(T)\,|\mathcal{F}_t\right) + \int_t^T \mathrm{Cov}_Q\left(r(s),r(T)\,|\mathcal{F}_t\right)\,\mathrm{d}s$$
$$-\frac{1}{2}\int_t^T \mathrm{Cov}_Q\left(I(s,T),r(T)\,|\mathcal{F}_t\right)\,\mathrm{d}s\ .$$

5.3.2 The Linear Gaussian Case

The Gaussian model is a model in which the volatility of the zero coupon bond is deterministic, and the spot rate is a Gaussian diffusion (Jamshidian [208], El Karoui et al [141, 142]). In this framework, we obtain a Black–Scholes type formula for the valuation of options on zero coupon bonds.

The Model

As before, we take as given for any maturity, the dynamics of zero coupon bond prices under the historic measure:

$$\mathrm{d}P(t,T) = P(t,T)\left((r(t) - q(t)\sigma(t,T))\mathrm{d}t + \sigma(t,T)\mathrm{d}\tilde{B}_t\right)\ ,$$

where $\tilde{B}$ is a one-dimensional Brownian motion under P, and where σ is a bounded deterministic function of class C^1 with respect to its second variable. We continue to assume that $\sigma(T,T) = 0$.

Under the risk-neutral measure, we have

$$\mathrm{d}P(t,T) = P(t,T)\left(r(t)\mathrm{d}t + \sigma(t,T)\,\mathrm{d}B_t\right)$$

where B is a Brownian motion under Q.

The results that were proved in the previous section here take the form:

Theorem 5.3.4. *The price at time t of a zero coupon bond is*

$$P(t,T)$$
$$= \frac{P(0,T)}{P(0,t)} \exp\left[\int_0^t (\sigma(s,T) - \sigma(s,t))\,\mathrm{d}B_s + \frac{1}{2}\int_0^t (\sigma^2(s,t) - \sigma^2(s,T))\,\mathrm{d}s\right],$$
$$t \leq T.$$

The forward spot rate is given by

$$f(t,T) = f(0,T) - \int_0^t \frac{\partial}{\partial T}\sigma(s,T)\,\mathrm{d}B_s + \int_0^t \sigma(s,T)\frac{\partial}{\partial T}\sigma(s,T)\,\mathrm{d}s.$$

The spot rate satisfies

$$r(t) = f(0,t) - \int_0^t \frac{\partial}{\partial T}\sigma(s,t)\,\mathrm{d}B_s + \int_0^t \sigma(s,t)\frac{\partial}{\partial T}\sigma(s,t)\,\mathrm{d}s, \quad \forall t. \quad (5.12)$$

Because σ is deterministic, the spot rate, as well as the forward spot rate, are Gaussian processes. (They can therefore take negative values, though after parameter calibration, this only occurs with a small probability).

These formulae give the dynamics of zero coupon bond prices, and of the spot and forward rates. These dynamics depend only on today's yield curve and on the volatility. The forward spot rate $f(0,t)$ is a biased estimator of the forward spot rate $f(t,T)$ and of the spot rate $r(t)$. The bias term depends only on the volatility.

By definition of Y, we obtain:

$$Y(t,T) = Y^f(0,t,T) - \int_0^t \frac{\sigma(s,T) - \sigma(s,t)}{T-t}\,\mathrm{d}B_s$$
$$+ \frac{1}{2}\int_0^t \frac{\sigma^2(s,T) - \sigma^2(s,t)}{T-t}\,\mathrm{d}s, \quad t \leq T \quad (5.13)$$

where

$$Y^f(0,t,T) = \frac{1}{T-t}\int_t^T f(0,s)\mathrm{d}s.$$

The yield curve at time t can be obtained from the initial curve by means of a random term and of the deterministic risk premium that depends only on the volatility.

Change of Numéraire

Let us return to the issue of a changing numéraire. We would like to interpret the biases in the previous formulae in terms of covariances. The variances and covariances are not altered by a change of measure of the Girsanov kind; and,

the volatility σ being deterministic, it has the same distribution under Q as under Q_T.

As commented on earlier,

$$\mathrm{Cov}_{Q_s}\left(E_Q(r(T)|\mathcal{F}_s), r(s)\right) = \mathrm{Cov}_Q\left(r(T), r(s)\right)$$

and, using (5.7),

$$f(t,T) = E_{Q_T}\left(r(T)|\mathcal{F}_t\right) = E_Q\left(r(T)|\mathcal{F}_t\right) - \int_t^T \mathrm{Cov}_Q\left(r(s), r(T)|\mathcal{F}_t\right)\,\mathrm{d}s \ .$$

In particular,

$$f(0,t) = E_Q(r(t)) - \int_0^t \mathrm{Cov}_Q(r(s), r(t))\,\mathrm{d}s \ .$$

According to (5.12), $r(t) = E_Q(r(t)) - \int_0^t \dfrac{\partial}{\partial T}\sigma(s,t)\,\mathrm{d}B_s$, hence we get

$$r(t) = f(0,t) - \int_0^t \frac{\partial}{\partial T}\sigma(s,t)\,\mathrm{d}B_s + \int_0^t \mathrm{Cov}_Q\left(r(s), r(t)\right)\,\mathrm{d}s \ .$$

Using (5.13) we also obtain:

$$\begin{aligned}
Y(t,T^*) &= Y^f(0,t,T^*) - \int_0^t \frac{\sigma(s,T^*) - \sigma(s,t)}{T^* - t}\,\mathrm{d}B_s^{Q_T} \\
&\quad + \frac{1}{2}\int_0^t \frac{\sigma^2(s,T^*) - \sigma^2(s,t)}{T^* - t}\,\mathrm{d}s \\
&\quad - \int_0^t \frac{\sigma(s,T^*) - \sigma(s,t)}{T^* - t}\,\sigma(s,T)\,\mathrm{d}s
\end{aligned}$$

$$\begin{aligned}
&= Y^f(0,t,T^*) - \int_0^t \frac{\sigma(s,T^*) - \sigma(s,t)}{T^* - t}\,\mathrm{d}B_s^{Q_T} \\
&\quad + \frac{1}{2}(T^* - t)\,\mathrm{Var}_Q\,Y_{t,T^*} - (T - t)\,\mathrm{Cov}_Q(Y_{t,T}, Y_{t,T^*}) \ .
\end{aligned}$$

The instantaneous rate contains a drift term that depends on the correlation between rates at earlier dates. Under Q_t, the instantaneous rate has expectation

$$E_{Q_t}(Y_{t,T}) = Y^f(0,t,T) + \frac{1}{2}(T - t)\,\mathrm{Var}_Q\,Y_{t,T} \ .$$

It is the price at time 0 of a forward contract with maturity t on the rate between times t and T. At time t, the expectation under Q_t of $Y_{t,T}$ equals

the forward rate plus a positive risk premium that is proportional to $\operatorname{Var} Y_{t,T}$. Under Q,

$$E_Q(Y_{t,T}) = Y^f(0,t,T) + \frac{1}{2}(T-t)\operatorname{Var} Y_{t,T} + \int_0^t \operatorname{Cov}(Y_{t,T}, r(s))\,\mathrm{d}s\ .$$

Thus it is also the price of a futures contract on the rate for the period (t, T). If the correlation between times 0 and t is positive, the price of the futures contract is higher than the price of the forward contract. Under Q_T, by exploiting the relationship between forward and futures prices, we can write:

$$\begin{aligned}
E_{Q_T}(Y_{t,T^*}) &= E_{Q_t}(Y_{t,T^*}) - \int_0^t \operatorname{Cov}(Y_{t,T^*}, r(s))\,\mathrm{d}s \\
&= E_{Q_t}(Y_{t,T^*}) - (T-t)\operatorname{Cov}(Y_{t,T^*}, Y_{t,T}) \\
&= Y^f(0,t,T^*) + \frac{T-t}{2}\operatorname{Var}_Q Y_{t,T} - (T-t)\operatorname{Cov}(Y_{t;T^*}, Y_{t,T})\ .
\end{aligned}$$

Exercise 5.3.5. Let $P^f(t,T,T^*) := \dfrac{P(t,T^*)}{P(t,T)},\ t \le T$ be the forward price of a zero coupon bond with maturity T^*. Show that

$$\begin{aligned}
P^f(t,T,T^*) = P^f(0,T,T^*) \\
\times \exp\left[\int_0^t [\sigma(s,T^*) - \sigma(s,T)]\,\mathrm{d}B_s - \frac{1}{2}\int_0^t [\sigma^2(s,T^*) - \sigma^2(s,T)]\,\mathrm{d}s\right]\ .
\end{aligned}$$

Using Property 5.2.3, check that $P^f(t,T,T^*)$ is a martingale under Q_T. Thence deduce that

$$\begin{aligned}
P^f(t,T,T^*) = P^f(0,T,T^*) \\
\times \exp\left[\int_0^t [\sigma(s,T^*) - \sigma(s,T)]\,\mathrm{d}B_s^{Q_T} - \frac{1}{2}\int_0^t [\sigma(s,T^*) - \sigma(s,T)]^2\,\mathrm{d}s\right]\ .
\end{aligned}$$

A Special Case

Let us consider the case $\sigma(s,t) = \sigma(t-s)$ where σ is a constant. We then have

$$f(t,T) = f(0,T) + \sigma^2 t\left(T - \frac{t}{2}\right) - \sigma B_t\ .$$

As remarked upon earlier, the forward spot rates are Gaussian, and can become negative with a positive probability.

$$r(t) = f(0,t) + \frac{\sigma^2 t^2}{2} - \sigma B_t\ .$$

The spot rate is equal to the forward rate between times 0 and t, plus a random perturbation and an adjustment factor. It too can be negative with a positive probability. From the expression $P(t,T) = \exp\left(-\int_t^T f(t,s)\mathrm{d}s\right)$, we also obtain

$$P(t,T) = \exp\left(-\int_t^T f(0,s)\mathrm{d}s - \sigma^2 \frac{Tt}{2}(T-t) + \sigma(T-t)B_t\right) \ .$$

This formula has a drawback: B_t is not observable. We remedy the situation by using the spot rate

$$P(t,T) = \exp\left(-\int_t^T (f(0,s) - f(0,t))\mathrm{d}s - \sigma^2 t \frac{(T-t)^2}{2} - (T-t)r(t)\right) \ .$$

This last formula depends only on observables, and only one parameter remains to be estimated.

Exercise 5.3.6. We suppose that $\sigma(s,t) = \sigma \dfrac{1 - \mathrm{e}^{-\lambda(t-s)}}{\lambda}$. Show that $r_t = E_Q(r_T) - \sigma Z_t$ where $Z_t = \int_0^t \mathrm{e}^{-\lambda(t-s)}\mathrm{d}B_s$. Next show that $f(t,T) = E_Q(f(t,T)) - \sigma \mathrm{e}^{-\lambda(T-t)} Z_t$ and that $Y(t,T) = E_Q(Y(t,T)) - \sigma \dfrac{1 - \mathrm{e}^{-\lambda t}}{\lambda t} Z_t$.

Valuation of a Call on a Zero Coupon Bond

We would like to valuate a call with strike K and maturity T, on a zero coupon bond with maturity T^*. The value of the call at maturity (that is at time T) is

$$(P(T,T^*) - K)^+ \ ,$$

hence $C(t) = E_Q\left((P(T,T^*) - K)^+ \exp\left(-\int_t^T r(s)\mathrm{d}s\right) | \mathcal{F}_t\right)$. If we choose the zero coupon bond of maturity T as numéraire, the associated probability measure is Q_T, and as we saw (in Exercise 5.3.5),

$$\mathrm{d}P_{t,T,T^*}^f = P_{t,T,T^*}^f \{\sigma(t,T^*) - \sigma(t,T)\} \, \mathrm{d}B_t^{Q_T} \ .$$

By definition of Q_T, we have

$$C(0) = E_Q\left((P(T,T^*) - K)^+ \exp\left(-\int_0^T r(s)\,\mathrm{d}s\right)\right)$$

$$= P(0,T)E_{Q_T}\left(\left(P_{T,T,T^*}^f - K\right)^+\right) \ .$$

Let $\Sigma(t) = \sigma(t,T^*) - \sigma(t,T)$ and $\Sigma^2 = \int_0^T \Sigma^2(s)\mathrm{d}s$. Analogous calculations to those of the Black–Scholes formula show that:

$$C(0) = P(0,T)\left(P^f_{0,T,T^*}\Phi(d) - K\Phi(d - \Sigma)\right)$$

where

$$d = \left(\ln\frac{P^f(0,T,T^*)}{K} + \frac{1}{2}\Sigma^2\right)/\Sigma \ .$$

5.4 When the Spot Rate is Given

Let us now assume that the spot rate is given.

We suppose that under the historic measure P, the spot rate $r(t)$ follows the Itô process defined by

$$\mathrm{d}r(t) \ = \ f(t,r_t)\,\mathrm{d}t + \rho(t,r_t)\,\mathrm{d}\tilde{B}_t \tag{5.14}$$

where $\tilde{B}$ is a one-dimensional P-Brownian motion, and f and ρ are continuous functions, which satisfy growth and Lipschitz conditions such that equation (5.14) admits a unique solution[10].

As with the Black–Scholes formula, we assume that the value $P(t,T)$ of a zero coupon bond is a function of $r(t)$, which we write $P(t,T;r(t))$, where $P(t,T;r)$ belongs to $C^{1,1,2}(\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R})$.

Itô's formula leads to (T is considered to be fixed, and we write $P(t,T)$ instead of $P(t,T;r(t))$):

$$\mathrm{d}P(t,T) = \left(\frac{\partial P}{\partial t} + f\frac{\partial P}{\partial r} + \frac{1}{2}\rho^2\frac{\partial^2 P}{\partial r^2}\right)(t,T)\mathrm{d}t + \rho\frac{\partial P}{\partial r}(t,T)\mathrm{d}\tilde{B}_t$$
$$= P(t,T)(\mu_t\mathrm{d}t + \sigma_t\mathrm{d}\tilde{B}_t)$$

with

$$\mu(t,T) \ = \ \frac{1}{P(t,T)}\left(\frac{\partial P}{\partial t} + f\frac{\partial P}{\partial r} + \frac{1}{2}\rho^2\frac{\partial^2 P}{\partial r^2}\right)(t,T) \tag{5.15i}$$

and

$$\sigma(t,T) \ = \ \frac{\rho(t,r(t))}{P(t,T)}\frac{\partial P}{\partial r}(t,T) \ . \tag{5.15ii}$$

The assumption of no arbitrage leads us to use the measure Q under which discounted prices all have the same expected return. Here, using Girsanov's theorem, we get $\left.\dfrac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathcal{F}_t} \ = \ L_t$, with $\ \mathrm{d}L_t = q_t L_t \mathrm{d}B_t$ and where $\ q(s) =$

---

[10] In practice, we choose simple expressions for f, ρ and q, depending on parameters. These parameters are then estimated by calibrating or fitting the theoretical yield curve to the observed yield curve.

$\dfrac{\mu(s,T) - r(s)}{\sigma(s,T)}$. Thus we obtain (by noting that as Q does not depend on T, nor does q) that all the discounted prices $P(s,t)$ are Q-martingales, i.e., that

$$P(s,t) = E_Q\left(P(t,t)\exp\left(-\int_s^t r(u)\mathrm{d}u\right)|\mathcal{F}_s\right); \quad s \le t \tag{5.16}$$

or, under the historic probability measure, that

$$P(s,t) = E_P\left(\exp\left\{\int_s^t q(u)\mathrm{d}B_u - 1/2\int_s^t q_u^2\mathrm{d}u - \int_s^t r(u)\mathrm{d}u\right\}|\mathcal{F}_s\right).$$

We exploit the Q-martingale property of $P(s,t;r(s))\exp\left[-\int_0^s r(u)\,\mathrm{d}u\right]$, by saying that the infinitesimal generator associated with the diffusion is zero. Hence

$$\frac{\partial P}{\partial t} + (f + \rho q)\frac{\partial P}{\partial r} + \frac{1}{2}\rho^2\frac{\partial^2 P}{\partial r^2} - rP = 0. \tag{5.17}$$

This type of equation is known as an evolution equation. Thus we recover

$$\begin{aligned}
\mathrm{d}P(t,T) &= P(t,T)\left(\left(r(t) - q(t)\sigma(t,T)\right)\mathrm{d}t + \sigma(t,T)\mathrm{d}\tilde{B}_t\right) \\
&= \left(P(t,T)r(t) - \rho q\frac{\partial P}{\partial r}\right)\mathrm{d}t + \rho\frac{\partial P}{\partial r}\mathrm{d}\tilde{B}_t. \tag{5.18}
\end{aligned}$$

This evolution equation applies to any security linked to interest rates, as long as the security does not pay out coupons. For a numerical solution to the equation, we need to assign it a terminal condition. For a zero coupon bound, this terminal condition is $P(T,T) = 1$.

Exercise 5.4.1. Let us show how, following the approach of Vasicek [360], (1977), we can recover the result above, reasoning by no arbitrage between zero coupon bonds of different maturities. We suppose that the price of a zero coupon bond has the dynamics

$$\mathrm{d}P(t,T) = P(t,T)(\mu_t\mathrm{d}t + \sigma_t\mathrm{d}\tilde{B}_t),$$

and that $P(t,t) = 1$. At time t, the agent sells an amount $\pi_1(t)$ of zero coupon bonds of maturity s_1, and buys an amount $\pi_2(t)$ of zero coupon bonds of maturity s_2. We have

$$\mathrm{d}\pi_i(t) = \pi_i(t)\left\{\mu(t,s_i)\mathrm{d}t + \sigma(t,s_i)\mathrm{d}\tilde{B}_t\right\}.$$

The resulting portfolio has value $\pi := \pi_2 - \pi_1$. Show that if π were an arbitrage portfolio, then we would have

$$\pi_2(t)\mu(t,s_2) - \pi_1(t)\mu(t,s_1) = \pi(t)r(t),$$

$$\pi_2(t)\sigma(t,s_2) - \pi_1(t)\sigma(t,s_1) = 0,$$

and that the existence of an arbitrage portfolio implies that $\dfrac{\mu(t,s) - r(t)}{\sigma(t,s)}$ must be constant and independent of s. We denote this term by $q(t,r)$; it is the market price of risk. Check that we recover equation (5.17).

5.5 The Vasicek Model

We now study a special case, by specifying the dynamics of the spot rate.

5.5.1 The Ornstein–Uhlenbeck Process

In this model, we suppose that under P the spot rate satisfies the stochastic differential equation

$$\mathrm{d}r_t = a(b - r(t))\mathrm{d}t + \rho\,\mathrm{d}\tilde{B}_t\,; \quad r(0) = r_0\,, \tag{5.19}$$

where a, b and ρ are strictly positive constants. This process is known as the Ornstein–Uhlenbeck process. The instantaneous mean is proportional to the difference between the value of b and the value of $r(t)$. A pull–back force tends to bring $r(t)$ closer to the value of b.

Proposition 5.5.1. *The explicit form of the solution to* (5.19) *is*

$$r(t) = (r_0 - b)\mathrm{e}^{-at} + b + \rho \int_0^t \mathrm{e}^{-a(t-u)}\,\mathrm{d}\tilde{B}_u\,. \tag{5.20}$$

Proof. It is enough to note that applying Itô's lemma to (5.19) yields

$$\mathrm{d}(\mathrm{e}^{at}r_t) = \mathrm{e}^{at}(ab\,\mathrm{d}t + \rho\,\mathrm{d}\tilde{B}_t)\,,$$

and to then integrate this last equation. $\square$

If r_0 is a constant, then $r(t)$ is a Gaussian variable with mean $(r_0 - b)\mathrm{e}^{-at} + b$ and variance $\dfrac{\rho^2}{2a}(1 - \mathrm{e}^{-2at})$. In particular, it is not a positive random variable.

More generally, if $r(0)$ is a Gaussian random variable that is independent of the Brownian motion B, then the family $r(t)$ is a random Gaussian function with expectation and variance (carry out the workings using (5.20))

$$E_P(r(t)) = b\left(1 - \mathrm{e}^{-at}\right) + \mathrm{e}^{-at}E_P(r(0))\,,$$

$$\mathrm{Var}_P\, r(t) = \rho_0\,\mathrm{e}^{-2at} + \rho^2\,\frac{1 - \mathrm{e}^{-2at}}{2a}\,,$$

$$\mathrm{Cov}_P(r(t), r(s)) = \rho_0\mathrm{e}^{-a(t+s)} + \rho^2 \int_0^s \mathrm{e}^{-a(s-u)}\mathrm{e}^{-a(t-u)}\mathrm{d}u$$

$$= \mathrm{e}^{-a(t+s)}\left(\rho_0 + \rho^2\,\frac{\mathrm{e}^{2as} - 1}{2a}\right)\,,$$

for $s \leq t$ and where ρ_0 denotes the variance of $r(0)$. We can also calculate the conditional expectation and conditional variance of $r(t)$:

Proposition 5.5.2. *If $(r(t), t \geq 0)$ is a process as in (5.19), we have for $s \leq t$*

$$E_P\left(r(t)\,|\mathcal{F}_s\right) \;=\; b + (r(s) - b)\,\mathrm{e}^{-a(t-s)} \,, \qquad\qquad (5.21)$$

and

$$\mathrm{Var}_P\left(r(t)\,|\mathcal{F}_s\right) = \frac{\rho^2}{2a}\left(1 - \mathrm{e}^{-2a(t-s)}\right) \,,$$

where

$$\mathrm{Var}_P\left(r(t)\,|\mathcal{F}_s\right) = E_P\left(r^2(t)\,|\mathcal{F}_s\right) - \left(E_P\left(r(t)\,|\mathcal{F}_s\right)\right)^2 \,.$$

Proof. These results can be obtained directly from the expression for $r(t)$

$$r(t) = (r(s) - b)\,\mathrm{e}^{-a(t-s)} + b + \rho\int_s^t \mathrm{e}^{-a(t-u)}\mathrm{d}\tilde{B}_u \,.$$

The workings for the variance are carried out in the same way, using the fact that

$$\mathrm{Var}_P(r(t)\,|\mathcal{F}_s) = \mathrm{Var}_P\,\rho\int_s^t \mathrm{e}^{-a(t-u)}\,\mathrm{d}\tilde{B}_u \,.$$

Thence, we deduce the expressions for the expectation and variance of $r(t)$, as well as

$$E_P\left(\int_s^t r(u)\,\mathrm{d}u\,|\mathcal{F}_s\right) \;=\; \int_s^t E_P(r(u)\,|\mathcal{F}_s)\,\mathrm{d}u$$

$$= b(t - s) + (r(s) - b)\,\frac{1 - \mathrm{e}^{-a(t-s)}}{a} \,.$$

Similar calculations, involving this time the covariances, show that

$$\mathrm{Var}_P\left(\int_s^t r(u)\,\mathrm{d}u\,|\mathcal{F}_s\right) \;=\; -\frac{\rho^2}{2a^3}(1 - \mathrm{e}^{-a(t-s)})^2 + \frac{\rho^2}{a^2}\left((t - s) - \frac{1 - \mathrm{e}^{-a(t-s)}}{a}\right) \,.$$

$\square$

Proposition 5.5.3. *The variable $\int_0^t r(s)\mathrm{d}s$ is a Gaussian variable with mean*
$bt + (r_0 - b)\dfrac{1 - \mathrm{e}^{-at}}{a}$ *and variance* $-\dfrac{\rho^2}{2a^3}\left(1 - \mathrm{e}^{-at}\right)^2 + \dfrac{\rho^2}{a^2}\left(t - \dfrac{1 - \mathrm{e}^{-at}}{a}\right)$.

Moreover,

$$E_P\left(\int_s^t r(u)\,\mathrm{d}u\,|\mathcal{F}_s\right) = b(t - s) + (r(s) - b)\,\frac{1 - \mathrm{e}^{-a(t-s)}}{a} \,, \qquad (5.22\mathrm{i})$$

and

$$\mathrm{Var}_P\left(\int_s^t r(u)\,\mathrm{d}u\,|\mathcal{F}_s\right) =$$

$$-\frac{\rho^2}{2a^3}\left(1 - \mathrm{e}^{-a(t-s)}\right)^2 + \frac{\rho^2}{a^2}\left((t - s) - \frac{1 - \mathrm{e}^{-a(t-s)}}{a}\right) \,. \qquad (5.22\mathrm{ii})$$

5.5.2 Determining $P(t,T)$ when q is Constant

We give two methods for determining $P(s,t)$ explicitly. The first uses the valuation partial differential equation, and the second involves using (5.4) and the distribution of r. We suppose that q is constant.

The Valuation Equation

We look for a solution to (5.17) of the form $P(t,T) = \exp(\alpha r(t) + \beta)$, where the coefficients (α, β) depend on $\theta = T - t$. They must then satisfy

$$-\alpha'(\theta)rP - \beta'(\theta)P + (a(b-r) + \rho q)\alpha(\theta)P + \frac{\rho^2}{2}\alpha^2(\theta)P - rP = 0 .$$

Hence α and β are solutions to

$$\begin{cases} \alpha'(\theta) + a\alpha(\theta) + 1 = 0 \\[2mm] -\beta'(\theta) + a(b + \rho q)\alpha(\theta) + \frac{\rho^2}{2}\alpha^2(\theta) = 0 , \end{cases}$$

with initial conditions $\alpha(0) = 0$ and $\beta(0) = 0$ coming from $P(T,T) = 1$. This yields

$$\begin{cases} \alpha(\theta) = \dfrac{1}{a}\left(e^{-a\theta} - 1\right) \\[3mm] \beta(\theta) = -\left(b + \dfrac{\rho q}{a} - \dfrac{\rho^2}{2a^2}\right)\theta - \left(b + \dfrac{\rho q}{a} - \dfrac{\rho^2}{a^2}\right)\dfrac{e^{-a\theta}}{a} - \dfrac{\rho^2}{4a^3}e^{-2a\theta} - K , \end{cases}$$

where the constant K is chosen in such a way that $\beta(0) = 0$.

We find

$$\beta(\theta) = -Y(\infty)\theta + \frac{1 - e^{-a\theta}}{a}\left(Y(\infty) - \frac{\rho^2}{2a^2}\right) + \frac{\rho^2}{4a^3}\left(1 - e^{-2a\theta}\right) ,$$

where $Y(\infty) = b + \dfrac{\rho q}{a} - \dfrac{\rho^2}{2a^2}$ represents the return on a zero coupon bond with infinite maturity, as appears in the following formulae. Thus we obtain:

$$P(t,T) = \exp\Big(-Y(\infty)(T-t) + \left(1 - e^{-a(T-t)}\right)\frac{Y(\infty) - r(t)}{a}$$
$$-\frac{\rho^2}{4a^3}\left(1 - e^{-a(T-t)}\right)^2 . \quad (5.23)$$

Using (5.18) (or differentiating (5.23)), we get:

$$dP(t,T) = P(t,T)\left((r(t) + \frac{\rho q}{a}(1 - e^{-a(T-t)})\,dt - \frac{\rho}{a}(1 - e^{-a(T-t)})\,d\tilde{B}_t\right) . \quad (5.24)$$

Theorem 5.5.4. *In the Vasicek model, the dynamics of the forward rate are given by*

$$f(t,T) = Y(\infty) - e^{-a(T-t)}\left(Y(\infty) - r(t)\right) + \frac{\rho^2}{2a^2}\left(1 - e^{-a(T-t)}\right)e^{-a(T-t)} \,,$$

therefore

$$f(t,T) - E\left(r(T)\,|\mathcal{F}_t\right) = \left(Y(\infty) - b + \frac{\rho^2}{2a^2}e^{-a(T-t)}\right)\left(1 - e^{-a(T-t)}\right) \,.$$

This model can be criticized on several grounds: the coefficients are constant over time, and $f(t,\infty)$, the long rate with infinite maturity, is constant, which does not occur in practice.

We have that $\sigma(t,T) = \frac{\rho}{a}\left(1 - e^{-a(T-t)}\right)$: the further out the maturity, the greater the volatility. We also have $\mu(t,\infty) = r(t) + \frac{\rho q}{a}$ and $\sigma(t,\infty) = \frac{\rho}{a}$. The rate $Y(t,T) = -\frac{1}{T-t}\ln P(t,T)$ is easily calculated:

$$Y(t,T) = Y(\infty) + (r(t) - Y(\infty))\frac{1 - e^{-a(T-t)}}{a(T-t)} + \frac{\left(1 - e^{-a(T-t)}\right)^2\rho^2}{4a^3(T-t)} \,.$$

If we study the function $T \to Y(t,T)$ (the so-called yield curve), we see that

$$Y(t,t) = r(t) \quad \text{and} \quad Y(t,T) \underset{T\to\infty}{\longrightarrow} Y(\infty) \,.$$

Moreover,

- if $r(t) \le Y(\infty) - \frac{\rho^2}{4a^2}$, the curve is strictly increasing,
- if $Y(\infty) - \frac{\rho^2}{4a^2} \le r(t) \le Y(\infty) + \frac{\rho^2}{2a^2}$, it is increasing and then decreasing,
- and if $Y(\infty) + \frac{\rho^2}{2a^2} < r(t)$, the curve is strictly decreasing.

If we define $R(t,\theta) = Y(t,t+\theta)$, we see that $R(t,\theta) \underset{\theta\to\infty}{\longrightarrow} Y(\infty)$ (which is independent of t).

The different shapes of this yield curve correspond to many of the curves observed in the markets. Nevertheless, some of the observed curves cannot be obtained in this model. Moreover, the problem of calibrating the parameters has not been satisfactorily solved, and $r(t)$ and $R(\infty)$ are not truly observables. In addition, the rates can become negative.

Calculation of the Conditional Expectation

Under Q, the process for r follows

$$\mathrm{d}r(t) = a\left(b_q - r(t)\right)\mathrm{d}t + \rho\mathrm{d}B_t$$

where $b_q = b + \dfrac{\rho q}{a}$ and where B is a Q-Brownian motion. The price of a zero coupon bond with maturity T is

$$P(t,T) = E_Q\left[\exp - \int_t^T r(s)\mathrm{d}s \,|\mathcal{F}_t\right] .$$

As r is Gaussian, it follows (see Annex 3) that $P(t,T)$ is a function of $r(t)$ of the form $\exp(\alpha r(t) + \beta) = \exp\left[E_Q(X\,|\mathcal{F}_t) - 1/2\,\mathrm{Var}_t\,X\right]$, where $X = -\int_t^T r(u)\,\mathrm{d}u$. Using the results of Proposition 5.5.3, we recover formula (5.23).

5.6 The Cox–Ingersoll–Ross Model

5.6.1 The Cox–Ingersoll–Ross Process

Cox–Ingersol–Ross [70] introduced a model where, in the risk-neutral world[11], the spot rate is driven by the equation

$$\mathrm{d}r_t = a(b - r_t)\,\mathrm{d}t + \rho\sqrt{r_t}\,\mathrm{d}B_t, \qquad r(0) = r_0 \tag{5.25}$$

with a, b and ρ positive.

We can show that this equation admits a unique solution[12] that is positive, but we do not have an explicit form for it. The solution does not reach 0 if $2ab \geq \rho^2$.

Theorem 5.6.1. *Let $r(t)$ be the process satisfying*

$$\mathrm{d}r_t = a(b - r_t)\mathrm{d}t + \rho\sqrt{r_t}\mathrm{d}B_t .$$

Its conditional expectation and conditional variance are given by

$$E_Q(r(t)\,|\mathcal{F}_s) = r(s)\mathrm{e}^{-a(t-s)} + b\left(1 - \mathrm{e}^{-a(t-s)}\right) ,$$

$$\mathrm{Var}_Q(r(t)\,|\mathcal{F}_s) = r(s)\frac{\rho^2\left(\mathrm{e}^{-a(t-s)} - \mathrm{e}^{-2a(t-s)}\right)}{a} + \frac{b\rho^2\left(1 - \mathrm{e}^{-a(t-s)}\right)^2}{2a} .$$

---

[11] If we were to work under the historic measure, we would assume that the spot rate satisfies (5.25) under P and that $q(t) = \alpha\sqrt{r(t)}$, where α is a constant.

[12] See Ikeda and Watanabe [204] p. 222 or Karlin [234].

Proof. By definition, for $s \leq t$, we have

$$r(t) = r(s) + a \int_s^t (b - r(u)) \mathrm{d}u + \rho \int_s^t \sqrt{r(u)} \, \mathrm{d}B_u , \qquad (5.26)$$

and, applying Itô's formula,

$$
\begin{aligned}
r^2(t) &= r^2(s) + 2a \int_s^t (b - r(u)) r(u) \mathrm{d}u + 2\rho \int_s^t (r(u))^{3/2} \mathrm{d}B_u + \rho^2 \int_s^t r(u) \mathrm{d}u \\
&= r^2(s) + (2ab + \rho^2) \int_s^t r(u) \mathrm{d}u - 2a \int_s^t r^2(u) \mathrm{d}u + 2\rho \int_s^t (r(u))^{3/2} \mathrm{d}B_u .
\end{aligned}
$$

$$ (5.27) $$

Assuming that the stochastic integrals that appear in the equalities above have zero expectation, we obtain for $s = 0$

$$E_Q(r_t) = r_0 + a \left(bt - \int_0^t E_Q(r_u) \mathrm{d}u \right) ,$$

and

$$E_Q(r^2(t)) = r^2(0) + (2ab + \rho^2) \int_0^t E_Q(r(u)) \mathrm{d}u - 2a \int_0^t E_Q(r^2(u)) \mathrm{d}u .$$

Solving the equation $\Phi(t) = r_0 + a \left(bt - \int_0^t \Phi(u) \mathrm{d}u \right)$, which can be transformed into the differential equation $\Phi'(t) = a(b - \Phi(t))$, we obtain

$$E[r(t)] = b + (r(0) - b)\mathrm{e}^{-at} .$$

Similarly, we calculate

$$\mathrm{Var}[r(t)] = \frac{\rho^2}{a} \left(1 - \mathrm{e}^{-at} \right) \left[r(0)\mathrm{e}^{-at} + \frac{b}{2}(1 - \mathrm{e}^{-at}) \right] ,$$

and the conditional expectation and variance of r:

$$E_Q(r(t) \,|\, \mathcal{F}_s) = r(s)\mathrm{e}^{-a(t-s)} + b \left(1 - \mathrm{e}^{-a(t-s)} \right) ,$$

$$\mathrm{Var}_Q(r(t) \,|\, \mathcal{F}_s) = r(s) \frac{\rho^2 \left(\mathrm{e}^{-a(t-s)} - \mathrm{e}^{-2a(t-s)} \right)}{a} + \frac{b\rho^2 \left(1 - \mathrm{e}^{-a(t-s)} \right)^2}{2a} .$$

□

5.6.2 Valuation of a Zero Coupon Bond

Proposition 5.6.2. *The price of a zero coupon bond with maturity T is of the form*

$$P(t,T) \;=\; \Phi(T-t)\mathrm{e}^{-r(t)\Psi(T-t)} \tag{5.28}$$

with

$$\Phi(s) \;=\; \left[\frac{2\gamma \mathrm{e}^{(a+\gamma)s/2}}{(a+\gamma)(\mathrm{e}^{\gamma s}-1)+2\gamma}\right]^{\frac{2ab}{\rho^2}}$$

$$\Psi(s) \;=\; \frac{2(\mathrm{e}^{\gamma s}-1)}{(a+\gamma)(\mathrm{e}^{\gamma s}-1)+2\gamma}\,, \qquad \gamma = (a^2+2\rho^2)^{1/2}\,.$$

Proof. Once again, we present two methods.

The Valuation Equation

The valuation equation is given by

$$\frac{\partial P}{\partial t} + a(b-r)\frac{\partial P}{\partial r} + \frac{1}{2}\rho^2 r \frac{\partial^2 P}{\partial r^2} - rP = 0\,, \tag{5.29}$$

with $P(T,T)=1$. If we look for solutions to (5.29) in the form (5.28), we find that Φ and Ψ are solutions to $\frac{1}{2}\rho^2\Psi^2 + a\Psi + \Psi' = 1$, $\quad \Psi(0)=0$ and $\Phi' = -ab\Psi\Phi$, $\quad \Phi(0)=0$. It is straightforward to check that there is a solution to the valuation equation, under the required form.

Calculation of the Conditional Expectation

Another method involves using probabilistic results to calculate

$$P(t,T) = E_Q\left[\exp\left(-\int_t^T r(s)\mathrm{d}s\right) \,\Big|\, \mathcal{F}_t\right]\,.$$

To do this, we need the conditional distribution of the variable

$$\exp\left(-\int_t^T r(s)\mathrm{d}s\right)\,.$$

We can check[13] that

---

[13] See Revuz–Yor [307].

$$E_Q \left[\exp \left(- \int_t^T r(s) \mathrm{d}s \right) \bigg| \mathcal{F}_t \right]$$

$$= E_Q \left[\exp \left(- \int_t^T r(s) \mathrm{d}s \right) \bigg| r_t \right] = F(T - t, r_t)$$

with $F(s, x) = E_Q \left[\exp \left(- \int_0^s r_x(u) \mathrm{d}u \right) \right]$, r_x being the process that is a solution to (5.25), with initial value x.

We show how to obtain the Laplace transform of $\int_0^s r_x(u) \mathrm{d}u$.

Let $G(s, x) = E_Q \left[\exp \left(- \mu \int_0^s r_x(u) \mathrm{d}u \right) \right]$ be that Laplace transform. Using results on partial differential equation (see Sect. 3 of Annex 3), we look for G as a solution to

$$\frac{\partial G}{\partial t} = \frac{\rho^2}{2} x \frac{\partial^2 G}{\partial x^2} + a(b - x) \frac{\partial G}{\partial x} - \mu x G$$

subject to the initial condition $G(0, x) = 1$. This is in fact equation (5.29). □

Exercise 5.6.3. This is a generalization of the previous models (Hull and White [202], 1990).
We suppose that under P,

$$\mathrm{d}r(t) = (\theta(t) + a(t)(b - r(t)) \, \mathrm{d}t + \sigma(t) \, \mathrm{d}\tilde{B}_t \ .$$

Show that P must satisfy

$$\frac{\partial P}{\partial t} + (\varPhi(t) - a(t)r) \frac{\partial P}{\partial r} + \frac{1}{2} \sigma^2(t) \frac{\partial^2 P}{\partial r^2} - rP = 0$$

where $\varPhi(t) = a(t)b + \theta(t) - q(t)\sigma(t)$.

Exercise 5.6.4. The two-factor model of Schaeffer and Schwartz [327], (1984).
The explanatory variables are the long rate $\ell(t)$ and the difference between the long rate and the spot rate, $e(t) = r(t) - \ell(t)$. We assume that under P,

$$de_t = m(\mu - e_t)\mathrm{d}t + \gamma \, \mathrm{d}\tilde{B}_1(t)$$

$$d\ell_t = a(b - \ell_t)\mathrm{d}t + c\sqrt{\ell_t} \, \mathrm{d}\tilde{B}_2(t)$$

where $\tilde{B}_1$ and $\tilde{B}_2$ are independent Brownian motions. We assume that the risk premia are given by q_1 for e and by $q_2\sqrt{\ell}$ for ℓ. The zero coupon bond has the value

$$P(t, T; e, \ell) = E_Q \left(\exp \left(- \int_t^T r(u) \, \mathrm{d}u \right) \bigg| e_t = e, \ell_t = \ell \right) \ .$$

Show that P satisfies

$$\frac{1}{2} \gamma^2 P''_{ee} + \frac{1}{2} c^2 \ell P''_{\ell\ell} + m(\mu + q_1\gamma - e)P'_e + a(b + q_2\ell c - \ell)P'_\ell + P'_t - (e + \ell)P = 0 \ .$$

Notes

The results of the first section (the discrete-time model) have been generalized by Jensen and Nielsen [217], (1998), who assume that the perturbation function depends on the number of up-moves, and by Sandmann and Sondermann [321], (1992). By considering the spot rate as a state variable, El Karoui and Saada [148], (1992), have shown that the forward spot rate depends on the spot rate in a linear fashion. The continuous time model can be introduced as a limit of the (discrete) Ho and Lee model. This approach was developed Heath–Jarrow–Morton [187], (1990). A detailed carried out by El Karoui and Saada [148], (1992), looks closely at the assumption of path-independence. Their calculations provided the inspiration for our exposition of the results concerning expressions for the yield curve.

The section in continuous time owes a great deal to Nicole El Karoui ([141, 142], (1992), and [136], (1993)). We give her our heartfelt thanks for allowing us to follow her lecture notes and papers in this way.

Presentations of the futures contract can by found in Duffie [113], (1989) and Hull [200], (2000).

Over the last few years, a new approach to studying the term structure of interest rates has been used. By means of the martingale measure, this method gives no arbitrage prices that depend only on the market price of risk. The starting point is a model of the dynamics of the zero coupon bonds, rather than of the interest rates. Today's yield curve is taken to be exogenous to the model and a model for the evolution of the yield curve is then developed. Martingale methods produce, in most cases, necessary and sufficient conditions for no arbitrage. The absence of arbitrage between the prices of zero coupon bonds of different maturities translates into the existence of a probability measure Q, under which discounted prices are martingales (Heath–Jarrow–Morton [188], (1990), Jamshidian [207, 208] (1989, 1991), El Karoui et al. [141, 142] (1992)). We model the price dynamics under Q, and study the dynamics of the yield curve.

The first articles to use the concept of the martingale measure in interest rate modeling, are Artzner and Delbaen [15], (1989), and Heath–Jarrow–Morton [187, 188], (1990). The approach makes it possible to valuate interest rate products: thus El Karoui and Rochet, [146], (1990) priced options on zero coupon bonds. The linear Gaussian case has been studied in detail by Jamshidian [208], (1991), and by El Karoui et al. [148] (1992). The sections concerning the Markovian model and the linear Gaussian case are directly inspired by El Karoui et al. [141, 142] (1992). These papers also contain numerous applications, along with a study of the quadratic Gaussian case. The forward measure has used by Geman [169], (1989), El Karoui–Rochet [146] (1990) and Geman–El Karoui–Rochet [170], (1995) to study the valuation of a variety of options.

Brennan and Schwartz [44], (1979), Jamshidian [206, 207], (1989), Longstaff and Schwartz [258], (1992), El Karoui and Lacoste [140], (1992), Duffie and Kan [120], (1993), Frachot and Lesne [165], (1993) also develop multi-factor models.

We have not addressed the issue of parameter calibration, or of the statistics of processes linked to the yield curve. The theory enables us to calculate the yield curve explicitly, once we have chosen a model, and identified the risk premium and the parameters. However, in practice, estimating the parameters from historical data for the short rate, and then building a yield curve in a Vasicek or Cox–Ingersol–Ross model, does not produce satisfactory results. Moreover, the risk premium is a function of the spot rate and of time (rather than of maturity), yet most models assume it to be constant, which is not satisfactory. Another method involves adjusting the risk premium parameter in such a way that the associated yield curve is as close as possible to the observed rates. We then find parameters that are stable over time, which is in contradiction to the underlying model. The reader is referred to Brigo and Mercurio [45], (2001).

Martellini and Priaulet's book [265], (2000), is an excellent introduction to interest rates. Lecture notes by Björk (published in [32], (1997)), Björk's paper [35], (2001), Musiela and Rutkowski's book [285], (1997), and Rebonato [306], (1996), present more elaborate models, and constitute of quasi-exhaustive study of yield curve modeling.

6

Equilibrium of Financial Markets in Discrete Time

What are the effects of introducing financial markets into an economy? How are stock prices determined? These are some of the questions that the current theory of equilibrium in financial markets seeks to address.

As we showed in the previous chapters, necessary conditions that prices (or gains) must satisfy follow from the NAO assumption: for example, there must exist a probability measure under which prices (or gains), discounted by the riskless rate, are martingales. However the compatibility of investors' choices has not been taken into account. The theory of equilibrium in financial markets set out here, on the other hand, takes investor preferences as a basis for explaining how stock prices are determined. The theory originated with Arrow [12] in 1953, and takes as its starting point the "theory of general equilibrium".

Recall that the theory of general equilibrium, which was initiated by Walras, explains the prices of economic goods using the equality of supply and demand. The first proofs of the existence (and uniqueness) of a Walras equilibrium are owed to Wald [361], (1936), however it was only at the beginning of the fifties that the theory was formalized in all generality, and that a body of existence proofs was given by Arrow–Debreu [13], (1954), McKenzie [270], (1959), Gale [168], (1955), Kuhn [246, 247], (1956), Nikaido [293], (1956) and Uzawa [356], (1956).

During the same period, Arrow [12], (1953) in "Le rôle des valeurs boursières pour la repartition la meilleure des risques", and then Debreu, [90], (1953) in "Economie de l'incertain", showed that the theory of general equilibrium, originally a static and deterministic theory, could be extended to the case where the future is uncertain, by introducing the concept of contingent goods. In the same article, Arrow [12] noted that as the introduction of these new concepts assumed a great number of markets to be open, so it required agents to have a huge computational capacity. Thus he suggested creating financial markets in order to lessen the number of open markets. The modern theory of equilibrium in financial markets is built on Arrow's idea.

Radner's 1972 paper [305] "Existence of Equilibrium of Plans, Prices and Price Expectations in a Sequence of Markets" plays a central part in the theory, for several reasons. On the one hand, Radner transposes Arrow's model into a dynamic framework, and introduces a more general class of assets than was considered by Arrow. Because he shows that, even when there are only a few assets, an exchange economy with financial markets can have an equilibrium, his article is the starting point for the "theory of incomplete markets". On the other hand, Radner introduces the concept of "rational expectations" into the Arrow model. The concept, which was first introduced by Muth [286] in 1961, has gained considerable importance in economic theory over the last thirty years, thus giving Radner a precursory role.

In the following, we will be using the notation below. Let x and y be two vectors in $\mathbb{R}^h$.

$x \geq y$ if and only if $x_i \geq y_i$ for all $i = 1, \ldots, h$.

$x > y$ if and only if $x \geq y$ and $x \neq y$.

$x \gg y$ if and only if $x_i > y_i$ for all $i = 1, \ldots, h$.

$x \cdot y$ denotes the scalar product of the two vectors.

$\mathbb{R}^h_+$ denotes the set of vectors x in $\mathbb{R}^h$ such that $x \geq 0$, and $\mathbb{R}^h_{++}$ denotes the set of vectors x of $\mathbb{R}^h$ such that $x \gg 0$.

6.1 Equilibrium in a Static Exchange Economy

We first recall the concept of a pure exchange economy within Arrow–Debreu theory. A l-good and m-consumer exchange economy is described by the data

1. the agents' sets of consumptions, which are here assumed to equal $\mathbb{R}^l_+$,
2. the agents' sets of endowments $e_i \in \mathbb{R}^l_{++}$ $(i = 1, \ldots, m)$,
3. the agents' sets of preferences, which are represented by utility functions $u_i : \mathbb{R}^l_+ \to \mathbb{R}$, $(i = 1, \ldots, m)$. Here we assume that

   U1   the functions u_i are continuous, strictly concave, and increasing, for all $(i = 1, \ldots, m)$.

Under "perfect competition", consumers cannot influence prices. Given a set of prices $p = (p^1, \ldots, p^l) \in \mathbb{R}^l_{++}$, the set of all the goods that agent i can buy at price p, given his endowment

$$B_i(p) = \{c \in \mathbb{R}^l_+ \mid p \cdot c \leq p \cdot e_i\},$$

is called the budget set of agent i.

When $p \gg 0$, it can easily be shown that $B_i(p)$ is a convex compact set. We suppose that agent i maximizes his utility function over his budget set.

As his utility function is assumed to be continuous and strictly concave, the agent chooses a unique vector of goods, denoted by $d_i(p) \in \mathbb{R}^l_+$, which is called the agent's demand at price p. As his preferences are increasing, the budget constraint is binding at $d_i(p)$, and we have

$$p \cdot d_i(p) \; = \; p \cdot e_i \qquad \text{for all} \quad i \, .$$

Definition 6.1.1. *A collection* $(p\, , d_i(p)\, , i = 1, \dots, m)$ *is an equilibrium if*

$$
\begin{aligned}
&i) \quad p \gg 0 \\
&ii) \quad \sum_{i=1}^{m} d_i(p) \; = \; \sum_{i=1}^{m} e_i \; := \; e \, .
\end{aligned}
\tag{6.1}
$$

The name "aggregate excess demand function" is given to the function $z = \sum_{i=1}^{m} (d_i - e_i)$. The price system p is an equilibrium price if and only if:

$$p \gg 0 \qquad and \qquad z(p) = 0 \, . \tag{6.1'}$$

Thus, an equilibrium price has the remarkable property of containing all the information in the economy that the individual agents require. Indeed, if each individual reacts to the price system according to his own endowments and preferences, without knowing those of others, then their individual demands are globally coherent.

From the mathematical point of view, proofs of the existence of an equilibrium come down to searching for the zeros of the aggregate excess demand function. Barring exceptional cases, existence proofs use a fixed point argument. The proofs fall into three categories. In the first, we find proofs that use either a fixed point theorem, such as those of Brouwer or Kakutani, or analogous arguments. In this category we find for example proofs carried out in the fifties, and proofs of the existence of an equilibrium in infinite dimensions. Proofs of this kind do not required any assumptions of differentiability. The second category uses combinatorial algorithms for calculating fixed points, based on Sperner's lemma (cf. Scarf [323, 324]), and dates back to the early seventies. The third category uses differential topology (the work of Debreu [92, 93], Dierker[103, 104], Balasko [20], Mas-Colell [267] and the references therein, and Smale); and is the most recent. The approach was developed in order to study the qualitative properties of equilibrium. In particular, this type of proof has been applied in the theory of incomplete markets. Problems related to the existence of an equilibrium have led to a great deal of progress in the mathematical theory of fixed points.

We present here two proofs of the existence of an equilibrium, both using Brouwer's or Kakutani's theorems.

The first proof is purely topological, and is the more classic approach. It is carried out in the space of goods $\mathbb{R}^l_+$. The second proof, the Negishi method,

is carried out in the space $\mathbb{R}_+^m$, where m is the number of agents. The proof is particularly interesting when the space of goods is infinite dimensional and the number of agents is finite, as we will see in Chap. 7, in continuous-time models.

6.2 The Demand Approach

For this method, we work directly with the excess demand function z. We can show[1] that the function $z : \mathbb{R}_{++}^l \to \mathbb{R}^l$ has the following properties:

Proposition 6.2.1.

1. z is homogeneous of degree zero, i.e., $z(\alpha p) = z(p)$ for all $p \gg 0$ and $\alpha > 0$.
2. z is continuous on $\mathbb{R}_{++}^l$.
3. z satisfies Walras' law, i.e., $p \cdot z(p) = 0$ for all $p \gg 0$.
4. If $p_n \xrightarrow[n\to\infty]{} p$ and $p^j = 0$, then $\|z(p_n)\| \xrightarrow[n\to\infty]{} \infty$.
5. z is bounded below: $z(p) \geq -e$ for all p.

As the excess demand function is positively homogeneous, we assume that

$$p \in \Delta^{l-1} = \left\{ p \in \mathbb{R}_+^l \ , \ \sum_{k=1}^l p^k = 1 \right\} .$$

Recall that a correspondence F (also called a many-valued function) from X into Y, is a mapping from X into $\mathcal{P}(Y)$ the set of subsets of Y. In other words, a correspondence differs from a mapping in that $F(x)$ can contain more than a single point. The graph of $F : X \to Y$ is the set

$$\text{graph} \, F = \{(x, y) \in X \times Y \ , \ y \in F(x)\} .$$

First, let us recall the following theorem:

Theorem 6.2.2 (Brouwer's Theorem). *Any continuous mapping from the simplex Δ^{l-1} into itself admits a fixed point.*

The theorem admits the following extension to correspondences:

Theorem 6.2.3 (Kakutani's Theorem[2]). *Let S be a non-empty compact convex subset of $\mathbb{R}^l$, let φ be a convex non-empty valued correspondence from S into S, and whose graph is closed. Then φ has a fixed point. That is, there exists $x \in S$ such that $x \in \varphi(x)$.*

---

[1] Mas-Colell et al. [268] pp. 581-582.
[2] Aliprantis and Border [4].

This leads us to deduce the following result.

Lemma 6.2.4 (Gale–Nikaido–Debreu Lemma). *Let S be a convex closed subset of the unit simplex Δ^{l-1}. Let f be a continuous function from S into $\mathbb{R}^l$ such that $p \cdot f(p) = 0$ for all p. Then there exists $p^* \in S$ such that*

$$p \cdot f(p^*) \ \leq \ 0 \qquad \text{for all} \qquad p \in S \ . \tag{6.2}$$

Proof. Let us consider the correspondence $\mu : f(S) \to S$ defined by

$$\mu(z) = \{p \in S \mid p \cdot z = \max\{q \cdot z \mid q \in S\}\} \ .$$

The correspondence μ is convex compact valued, and we can easily show that the graph of μ is closed. Let us consider the correspondence from $S \times f(S)$ into itself, defined by: $(p, z) \mapsto (\mu(z), f(p))$. Its values are convex and non-empty, and its graph is closed. It follows from Kakutani's theorem, that there exists (p^*, z^*) such that $p^* \in \mu(z^*)$ and $z^* = f(p^*)$. Hence

$$p \cdot f(p^*) \ = \ p \cdot z^* \ \leq \ p^* \cdot z^* \ = \ p^* \cdot f(p^*) = 0 \qquad \text{for all } \ p \in S \ .$$

$\square$

We can now finish our proof of the existence of an equilibrium.

Theorem 6.2.5. *Under the assumption* **U1***, there exists an equilibrium.*

Proof. If we could apply the Gale–Nikaido–Debreu lemma to the simplex Δ^{l-1} and to the function z, the proof of the existence of an equilibrium would be immediate. Indeed, if $p \cdot z(p^*) \leq 0$ for all $p \in \Delta^{l-1}$ then $z(p^*) \leq 0$, but as $p^* \cdot z(p^*) = 0$ (Walras' law) then necessarily $z(p^*) = 0$. Unfortunately the Gale–Nikaido–Debreu lemma does not apply to the aggregate excess demand function z, as it is not continuous on the boundary of the simplex. This leads us to truncate the simplex, and to work by taking limits.

For $n \in \mathbb{N}$, let

$$\Delta_n^{l-1} \ = \ \left\{p \in \Delta^{l-1} \mid p^j \geq \frac{1}{n} \ , \ j \in \{1, \ldots, l\}\right\} \ .$$

As the restriction of z to Δ_n^{l-1} is continuous, it then follows from the GND lemma that there exists $p_n^* \in \Delta_n^{l-1}$ such that

$$p \cdot z(p_n^*) \ \leq \ 0 \quad , \quad p \in \Delta_n^{l-1} \ . \tag{6.3}$$

Since the sequence (p_n^*) is in Δ^{l-1}, it has a limit point p^*. Let us show that $p^* \gg 0$. According to Proposition 6.2.1 (4), it is enough to show that the sequence $z(p_n^*)$ is bounded. From Proposition 6.2.1 (5), it is bounded below by $-e$. Moreover, if $z(p_n^*) = (z^k(p_n^*))_{k=1}^l$, by applying (6.3) to $p^j = \frac{1}{l}$, for all j and for n large enough, we obtain:

$$z^1(p_n^*) \ \leq \ -\sum_{k=2}^{l} z^k(p_n^*) \ \leq \ \sum_{k=2}^{l} e^k \ .$$

The sequence $z^1(p_n^*)$ is therefore bounded above, and a similar reasoning can be used to show that the sequences $z^k(p_n^*)$ for $k = 2, \ldots, l$ are also bounded above.

Therefore $p^* \gg 0$, and from Proposition 6.2.1 (2), $z(p^*)$ is a limit point of the sequence $z(p_n^*)$.

As the sequence of truncated simplices Δ_n^{l-1} is increasing, we have $p \cdot z(p^*) \leq 0$ for all $p \in \Delta_n^{l-1}$, and by taking limits as $n \to \infty$, we have $p \cdot z(p^*) \leq 0$, $p \in \Delta^{l-1}$. Hence $z(p^*) \leq 0$. Since $p^* \cdot z(p^*) = 0$, we have $z(p^*) = 0$, and hence the existence of an equilibrium.   □

6.3 The Negishi Method

Although purely topological proofs do exist, here we give an exposition of the Negishi method under fairly restrictive assumptions of differentiability on the utility functions u_i, so as to emphasize formulae that also appear in continuous time.

In addition to the previous assumptions, we suppose

U2   For all i, u_i is C^2 on $\mathbb{R}_{++}^l$.

U3   For all i, u_i satisfies the "Inada conditions": $\dfrac{\partial u_i}{\partial x^j}(x) \to \infty$ if $x^j \to 0$,
    where the other components of x are fixed.

6.3.1 Pareto Optima

Definition 6.3.1. *An allocation $(c_i)_{i=1}^{m} \in (\mathbb{R}_+^l)^m$ is a Pareto optimum if there do not exist $(c_i')_{i=1}^{m} \in (\mathbb{R}_+^l)^m$ with $\sum_{i=1}^{m} c_i' \leq e$ such that $u_i(c_i') \geq u_i(c_i)$ for all i, and $u_j(c_j') > u_j(c_j)$ for at least one j.*

Definition 6.3.2. *A pair $(\overline{p}, (\overline{c}_i)_{i=1}^{m} \in (\mathbb{R}_+^l)^m)$ is an equilibrium with transfer payments, if for all i,*

$$\begin{cases} \overline{c}_i \text{ maximizes } u_i(c_i) \text{ under the constraint} \\ \overline{p} \cdot c_i \ \leq \ \overline{p} \cdot \overline{c}_i \end{cases}$$

and if $\sum_{i=1}^{m} \overline{c}_i = e$.

Thus, we can see that $(\overline{p}, (\overline{c}_i)_{i=1}^{m} \in (\mathbb{R}_+^l)^m)$ would be an equilibrium if agent i had $\overline{c}_i$ as his initial endowment. We would need to "transfer" $\overline{p} \cdot (e_i - \overline{c}_i)$ to him in order to attain an equilibrium.

The Negishi method rests on the first and second welfare theorems, which state that any equilibrium is a Pareto optimum, and that any Pareto optimum is an equilibrium with transfer payments. An equilibrium is therefore a Pareto optimum whose transfer payments are zero. Thus we start by characterizing Pareto optima.

6.3.2 Two Characterizations of Pareto Optima

Let $\alpha \in \Delta^{m-1}$. We will call α a utility weight vector, as we will construct an aggregate utility function by attributing the weight α_i to agent i. Let us consider the problem P_α for a given e:

$$
P_\alpha \quad
\begin{cases}
\text{maximize} \quad \{\alpha_1 u_1(c_1) + \cdots + \alpha_m u_m(c_m)\} \\
\text{under the constraints} \\
c_i \geq 0 \quad \text{for all } i \text{ and } \sum_{i=1}^m c_i \leq e .
\end{cases}
$$

We obtain the following result:

Proposition 6.3.3. $(\bar{c}_i)_{i=1}^m$ *is a Pareto optimum if and only if there exists a utility weight vector* $\alpha \in \Delta^{m-1}$ *such that* $(\bar{c}_i)_{i=1}^m$ *is the optimal solution to problem* P_α.

Proof. It is easy to show that the solution to P_α is Pareto optimal. Conversely, let us show that we can associate with any Pareto optimum $(\bar{c}_i)_{i=1}^m$, a utility weight vector $\alpha \in \Delta^{m-1}$ such that $(\bar{c}_i)_{i=1}^m$ is the optimal solution to the associated problem P_α.

Let us consider the following sets:

$$
A = \left\{ (c_i)_{i=1}^m \mid c_i \geq 0 \text{ for all } i \,,\, \sum_{i=1}^m c_i \leq e \right\} ,
$$

$$
U = \{ (u_i(c_i))_{i=1}^m \mid (c_i)_{i=1}^m \in A \} ,
$$

and

$$
V = \left\{ z \in \mathbb{R}^m \mid z^i \geq u_i(\bar{c}_i) \text{ for all } i \,, \right.
$$

$$
\left. z^j > u_j(\bar{c}_j) \text{ for at least one } j \right\} .
$$

It is straightforward to show that U is convex and compact, and that V is non-empty and convex, and by definition of the Pareto optimum, $U \cap V = \emptyset$. It follows from Minkowski's Theorem[3], that there exists a family of coefficients $(\alpha_1, \ldots, \alpha_m)$ that are all non-zero, and satisfy

---

[3] See Chap. 1, Sect. 1.2.

$$\sum_{i=1}^{m} \alpha_i w_i \ \leq \ \sum_{i=1}^{m} \alpha_i z^i \qquad w \in U \ , \ z \in V \ . \tag{6.4}$$

Let us apply (6.4) to the pair $w = [u_i(\bar{c}_i)]$ and $z = [u_1(\bar{c}_1) + t \, , \, u_2(\bar{c}_2)$ $\ldots, \, u_m(\bar{c}_m)]$ for $t > 0$. By cancelling out terms, we obtain $t\alpha_1 \geq 0$ for all $t > 0$, and hence $\alpha_1 \geq 0$, and by symmetry, $\alpha_i \geq 0$ for all i. As (6.4) is homogeneous in α, and as the α_i are not all zero, we can assume that $\alpha = (\alpha_1, \ldots, \alpha_m) \in \Delta^{m-1}$. Finally, as $[u_i(\bar{c}_i)]_{i=1}^{m} \in \overline{V}$ (closure of V), (6.4) implies that

$$\sum_{i=1}^{m} \alpha_i u_i(c_i) \ \leq \ \sum_{i=1}^{m} \alpha_i u_i(\bar{c}_i) \tag{6.5}$$

for all $(c_i)_{i=1}^{m}$ such that $c_i \geq 0$ for every i and $\sum_{i=1}^{m} c_i \leq e$. Hence $(\bar{c}_i)_{i=1}^{m}$ is a solution to problem P_α. $\qquad \square$

Let us use Proposition 6.3.3 to obtain another characterization of Pareto optima.

Let us fix $\alpha \in \Delta^{m-1}$, and consider problem P_α. It can easily be seen that if $\alpha_i = 0$, then $\bar{c}_i(\alpha) = 0$. Under Inada's condition, we can show that if $\alpha_i > 0$, then $\bar{c}_i(\alpha) \gg 0$. (The property does not hold without the condition.) Therefore there exists a vector of Lagrange multipliers $\lambda = (\lambda_1, \ldots, \lambda_l)$, $\lambda \gg 0$ such that, for all i such that $\alpha_i > 0$:

$$\alpha_i \ \mathrm{grad} \ u_i(\bar{c}_i) = \lambda \ . \tag{6.6}$$

This relationship implies that the utility weight vector that we attribute to a Pareto optimum, is unique. Indeed, assume that $\bar{c}_{p+1} = \cdots = \bar{c}_m = 0$. Then $\alpha_{p+1} = \cdots = \alpha_m = 0$, and if $[\, \mathrm{grad} \ u_i(\bar{c}_i)]^1$ denotes the first component of $\mathrm{grad} \ u_i(\bar{c}_i)$, we have

$$\alpha_1 [\, \mathrm{grad} \ u_1 \ (\bar{c}_1)]^1 \ = \ \alpha_2 [\, \mathrm{grad} \ u_2 \ (\bar{c}_2)]^1 \ = \ \ldots \ = \ \alpha_p [\, \mathrm{grad} \ u_p \ (\bar{c}_p)]^1 \ .$$

As $\sum_{i=1}^{p} \alpha_i = 1$, the α_i are uniquely determined. $\qquad \square$

Let us define the aggregate utility:

$$\begin{cases} u(\alpha, e) \ = \ \max \ \{\alpha_1 u_1(c_1) + \cdots + \alpha_m u_m(c_m)\} \\ \text{under the constraints} \\ c_i \geq 0 \, , \quad \text{for all } i \quad \text{and} \quad \sum_i c_i \leq e \ . \end{cases} \tag{6.7}$$

We now prove the second welfare theorem.

Proposition 6.3.4. *An allocation $(\bar{c}_i)_{i=1}^{m}$ is a Pareto optimum if and only if, when α is the associated utility weight vector, $\big(p(\alpha), (\bar{c}_i)_{i=1}^{m}\big)$ with $p(\alpha) = \mathrm{grad} \ u(\alpha, e)$ is an equilibrium with transfer payments.*

Proof. Let us first show that if $(p(\alpha), (\bar{c}_i)_{i=1}^m)$ is an equilibrium with transfer payments, then $(\bar{c}_i)_{i=1}^m$ is a Pareto optimum. If not, there would exist an allocation $(c_i')_{i=1}^m \in (\mathbb{R}_+^l)^m$ with $\sum_{i=1}^m c_i' \leq e$ such that $u_i(c_i') \geq u_i(\bar{c}_i)$ for all i, and with a strict inequality for some j. We would then have $p(\alpha) \cdot c_i' \geq p(\alpha) \cdot \bar{c}_i$ for all i, and $p(\alpha) \cdot c_j' > p(\alpha) \cdot \bar{c}_j$ for at least one j. Hence we would have $p(\alpha) \cdot \sum_{i=1}^m c_i' > p(\alpha) \cdot e$, which contradicts the inequality $\sum_{i=1}^m c_i' \leq e$.

Now let $(\bar{c}_i)_{i=1}^m$ be a Pareto optimum, and let α be the associated utility weight vector. Assume further that $\alpha_1 > 0, \; \alpha_2 > 0, \; \ldots, \alpha_p > 0$ and that $\alpha_{p+1} = \cdots = \alpha_m = 0$. Consider the following system of $(p+1)l$ equations with $(p+1)l$ unknowns $(\bar{c}_i, \; \lambda) \in (\mathbb{R}_+^l)^p \times \mathbb{R}_{++}^l$:

$$\begin{cases} \alpha_1 \; \text{grad} \; u_1(\bar{c}_1) = \lambda \\ \vdots \\ \alpha_p \; \text{grad} \; u_p(\bar{c}_p) = \lambda \\ \sum_{i=1}^p \bar{c}_i = e \, . \end{cases} \tag{6.8}$$

Let G be the matrix below (I denotes the unit matrix of $L(\mathbb{R}^l, \mathbb{R}^l)$)

$$\begin{bmatrix} \alpha_1 \dfrac{\partial^2 u_1}{\partial x_1^2}(\bar{c}_1) & 0 & & 0 & I \\ 0 & \alpha_2 \dfrac{\partial^2 u_2}{\partial x_2^2}(\bar{c}_2) & & & I \\ & & \ddots & & \\ 0 & 0 & & \alpha_p \dfrac{\partial^2 u_p}{\partial xp^2}(\bar{c}_p) \, I & \\ I & I & & I & 0 \end{bmatrix}$$

G can then be written

$$G = \begin{bmatrix} A & B \\ B^T & 0 \end{bmatrix}$$

where $A \in L(\mathbb{R}^{pl}, \mathbb{R}^{pl})$ is negative definite and $B \in L(\mathbb{R}^l, \mathbb{R}^{pl})$ has rank l.

Let us show that $\text{Ker} \; G = \{0\}$. Let $(X, Y) \in \mathbb{R}^{pl} \times \mathbb{R}^l$ be such that $G\binom{X}{Y} = 0$. Then we get $AX + BY = 0$ and $B^T X = 0$, and hence $-X = A^{-1}BY$ and $Y^T B^T A^{-1} BY = 0$. Thus $BY = 0$. As B is injective, $Y = 0$ and hence $X = 0$.

It follows from the local inversion theorem that $((\bar{c}_i)_{i=1}^m, \; \lambda)$ is a differentiable function of e. The function $u(\alpha, \cdot)$ is then also a differentiable function of e. Using differential notation, we obtain from (6.6),

$$du = \sum_{i : \alpha_i > 0} \alpha_i \; \text{grad} \; u_i(\bar{c}_i) \, d\bar{c}_i = \lambda \sum_{i : \alpha_i > 0} d\bar{c}_i = \lambda \, de$$

and hence

$$\lambda = \text{grad } u(\alpha, e) . \tag{6.9}$$

We set:

$$p(\alpha) = \text{grad } u(\alpha, e) = \alpha_i \text{ grad } u_i(\bar{c}_i) \text{ for all } i \text{ such that } \alpha_i > 0 . \tag{6.10}$$

As a result of (6.10), for all i such that $\alpha_i > 0$, $\bar{c}_i$ is a solution to the optimization problem P_i:

$$P_i \quad \begin{cases} \max \ u_i(c_i) \\ \text{under the constraints} \\ p(\alpha) \cdot c_i \ \leq \ p(\alpha) \cdot \bar{c}_i . \end{cases}$$

If $\alpha_i = 0$, then trivially $\bar{c}_i = 0$ is the optimal solution to P_i. Hence $(p(\alpha), \bar{c}_i ; i = 1, \ldots, m)$ is an equilibrium with transfer payments. □

6.3.3 Existence of an Equilibrium

We deduce the following results:

Theorem 6.3.5. *Under the assumptions* **U1**, **U2** *and* **U3**, *there exists an equilibrium.*

Proof. Let $\alpha \in \Delta^{m-1}$ and let $[\bar{c}_i(\alpha)]_{i=1}^m$ be the optimal solution to the associated problem P_α. Let $(p(\alpha) \cdot (\bar{c}_i(\alpha) - e_i))_{i=1}^m$ be the associated transfers. Let us prove the existence of an α^*, called the "equilibrium weight" such that the transfers are zero for all of the agents.

Let $\Phi : \Delta^{m-1} \to \mathbb{R}^l$ be the transfer function defined as follows:

$$\Phi_i(\alpha) = p(\alpha) \cdot (\bar{c}_i(\alpha) - e_i) , \qquad \text{for all } i . \tag{6.11}$$

By definition, α^* is an equilibrium weight if and only if it is a zero of Φ.

To show that Φ admits a zero, we first show that Φ is continuous. To do this, notice that it follows from the theorem of the maximum (see annex) that $[\bar{c}_i(\alpha)]_{i=1}^m$, the solution to P_α, is a continuous function of α. As $p(\alpha) = \alpha_i \text{ grad } u_i[\bar{c}_i(\alpha)]$ for all i such that $\alpha_i > 0$, the mapping $\alpha \to p(\alpha)$ is also continuous, and hence Φ is continuous. Moreover, Φ has the following properties:

$$\sum_{i=1}^m \Phi_i(\alpha) = p(\alpha) \cdot \left(-e + \sum_{i=1}^m \bar{c}_i(\alpha) \right) = 0 . \tag{6.12}$$

If $\alpha_i = 0$, then $\bar{c}_i(\alpha) = 0$ and hence $\Phi_i(\alpha) = -p(\alpha) \cdot e_i < 0 . \tag{6.13}$

Next we use the theorem below, which is a generalization of Brouwer's theorem. Let $H = \{c \in \mathbb{R}^m , \ \sum_{i=1}^m c_i = 0\}$.

Theorem 6.3.6. *Let $\Phi : \Delta^{m-1} \to H$ be continuous and satisfy the boundary condition*

$$if \quad \alpha_i = 0 \quad , \quad \Phi_i(\alpha) < 0 \; .$$

Then there exists $\alpha_0 \gg 0$ such that $\Phi(\alpha_0) = 0$.

Proof. Let $\Phi_i^- = \max(-\Phi_i, 0)$. Under this notation, $\Phi_i^-(\alpha) = 0$ is equivalent to $\Phi_i(\alpha) \geq 0$.

Let us consider the continuous mapping from Δ^{m-1} into itself, defined by:

$$G_i(\alpha) \;=\; \frac{\alpha_i + \Phi_i^-(\alpha)}{1 + \sum_{i=1}^{m} \Phi_i^-(\alpha)} \; .$$

It follows from Brouwer's theorem that G has a fixed point α_0. There are two possible cases. Either $\Phi_i(\alpha_0) = 0$ for all i and our proof is finished, or there exists i such that $\Phi_i(\alpha_0) > 0$ (since $\sum_{i=1}^{m} \Phi_i(\alpha_0) = 0$ we cannot have $\Phi_i(\alpha_0) \leq 0$ for all i). From the boundary condition, $\alpha_{0i} \neq 0$.

Therefore we have $\alpha_{0i} = \dfrac{\alpha_{0i}}{1 + \sum_{i=1}^{m} \Phi_i^-(\alpha_0)}$. Hence $\sum_{i=1}^{m} \Phi_i^-(\alpha_0) = 0$.

This implies that $\Phi_i^-(\alpha_0) = 0$, so that $\Phi_i(\alpha_0) \geq 0$ for all i. As $\sum_{i=1}^{m} \Phi_i(\alpha_0) = 0$, we have $\Phi_i(\alpha_0) = 0$ for all i. The boundary condition implies that $\alpha_0 \gg 0$. $\square$

Remark 6.3.7. The existence of a zero for Φ is equivalent to the existence of a fixed point for the mapping $\Phi + \mathrm{Id}$.

Moreover, as $\sum_{i=1}^{m} \Phi_i = 0$, the mapping $\Phi + \mathrm{Id}$ has values in $\{x \in \mathbb{R}^m \mid \sum_{i=1}^{m} x_i = 1\}$, which is the plane containing the unit simplex Δ^{m-1}. Instead of looking for a fixed point of a continuous mapping from the simplex into itself, as in Brouwer's theorem, we look for the fixed point of a continuous mapping from the simplex into the plane containing the simplex. We have a boundary condition

$$if \, \alpha_i = 0 \, , \quad (\Phi + \mathrm{Id})_i(\alpha) \;=\; \Phi_i(\alpha) \;<\; 0 \; .$$

This condition is called "outward," as at the boundary of the simplex, the vector field $\Phi + \mathrm{Id}$ points towards the outside of the simplex. For this reason, the result is considered to be a generalized form of Brouwer's theorem.

Remark 6.3.8. Assume $\Phi : \Delta^{m-1} \to H$ to be continuous, and to satisfy the "inward" condition at the edge: when $\alpha_i = 0$, $\Phi_i(\alpha) > 0$. Then Φ admits a strictly positive zero. Obviously, it is enough to change Φ into $-\Phi$.

6.4 The Theory of Contingent Markets

As mentioned in the introduction, Arrow [12], (1953) in "Le rôle des valeurs boursières pour la répartition la meilleure des risques" and then Debreu [90],

(1953) in "Economie de l'incertain" showed how the static and deterministic theory of equilibrium could be generalized to the multi-period case and to the case with uncertainty, on condition that a good is defined not only by its physical characteristics, but also by the state of the world, the date of its use, and on the condition that there are open markets for all these goods. Let us specify these ideas in a simple case.

We consider an exchange economy with two dates, m agents and l goods. A time 1, the future is uncertain, and there are k possible states of the world. There are open markets for all the goods in all states of nature. That is to say that an agent can buy a contract for the delivery of a given merchandise in a given state of the nature. The contract is paid for, even though the delivery does not take place unless the specified event occurs. Agent i can therefore make consumption plans $c_i(j) \in \mathbb{R}^l_+$ for state j. The vector $c_i = (c_i(1), c_i(2), \ldots, c_i(k)) \in (\mathbb{R}^l_+)^k$ is called the contingent consumption plan. Let us assume that agent i has preferences over the set of contingent consumption plans, and that they are represented by a utility function $u_i : (\mathbb{R}^l_+)^k \to \mathbb{R}$. Finally, let $e_i = (e_i(1), \ldots, e_i(k))$ be the endowment vector of agent i where $e_i(j)$ denotes the endowment in goods of agent i in state j.

Thus the exchange economy is characterized by the list:

$$\left((\mathbb{R}^l_+)^k \ , \ u_i \ , \ e_i \ ; \ i = 1, \ldots, m \right) .$$

Let $p^l(j)$ be the price of good l to be delivered if state j occurs. The vector $p = [p(1), \ldots, p(k)] \in (\mathbb{R}^l_+)^k$ is called a set of contingent prices.

Given $p \in (\mathbb{R}^l_{++})^k$, the agent determines his budget set, that is to say, the set of plans that are compatible with his set of endowments:

$$B_i(p) \ = \ \{ c_i \in (\mathbb{R}^l_+)^k \mid p \cdot c_i \le p \cdot e_i \} \tag{6.14}$$

where

$$p \cdot c_i \ = \ \sum_{j=1}^{k} p(j) \cdot c_i(j) .$$

Definition 6.4.1. *A contingent Arrow–Debreu equilibrium   is a set of contingent prices $p^* \in (\mathbb{R}^l_{++})^k$ and a set of contingent plans $(c_i^*)_{i=1}^m \in (\mathbb{R}^l_+)^{km}$ such that*

1. *c_i^* maximizes $u_i(c_i)$ under the constraint $c_i \in B_i(p^*)$ for all $i = 1, \ldots, m$.*
2. *The markets clear, i.e., $\sum_{i=1}^m c_i^* = \sum_{i=1}^m e_i$.*

If we make the assumption **U1**, then there exists a contingent Arrow–Debreu equilibrium. This approach has two drawbacks: first of all, it requires a large number of markets to be open (in the previous example there are kl markets). Secondly, the contingent goods are not always for sale. Hence

Arrow's idea, put forward in the article referenced above, of introducing a number k of securities, in order to show the economy can be organized with $k + l$ markets, instead of with kl contingent markets. We describe this model in the case of an economy with two dates (one period).

6.5 The Arrow–Radner Equilibrium Exchange Economy with Financial Markets with Two Dates

As in the previous model, there are two dates. At time 1, the future is uncertain, and there are k possible states of the world at time 2. Agent i has uncertain endowments, and as before, $e_i(j)$ denotes the endowment of agent i in state j. This time however, there are no markets for goods delivered in the future. On the other hand, agents can buy portfolios of securities during the first period. Anticipating the price levels for time 2, they can make consumption plans in terms of the income that they anticipate getting from their exogenous endowments and from their securities.

In this model, we make the assumption that agents have "rational expectations", in other words, the prices that they expected prices do occur. Agents then exchange goods in the one state that does come about, in markets that we call the "spot markets". Let us now fully specify the model.

First we describe the financial part of the economy. There are d securities. Each asset is characterized by the dividend it yields in each state of nature. We say that an asset is "real" if its dividend is expressed in units of the good. We say that it is "nominal" if there is a numéraire in each state of nature, and if the dividend is expressed in monetary units. In the latter case, the matrix V whose i-th column represents the dividend of asset i in the various states, is called the "dividend matrix".

$$V = \begin{array}{c} \overset{\displaystyle i\text{-th asset}}{\downarrow} \\ \begin{bmatrix} v_1^1 \ \dots \ v_1^i \ \dots \ v_1^d \\ \vdots \\ v_j^1 \ \dots \ v_j^i \ \dots \ v_j^d \\ \vdots \\ v_k^1 \ \dots \ v_k^i \ \dots \ v_k^d \end{bmatrix} \end{array} \quad \leftarrow \quad j\text{-th state}$$

From now on, we assume that assets are nominal. The agents construct portfolios for themselves. A portfolio θ is a vector in $\mathbb{R}^d$, whose components can be negative (short selling is allowed). The payoff of this portfolio in state j is $(V\theta)_j$. The securities are traded at time 1 at price $S \in \mathbb{R}_+^d$. We suppose that the agents cannot run into debt, hence that $S \cdot \theta_i \leq 0$ for all i, where θ_i is the portfolio of agent i.

Given expected prices $p = [p(1), \ldots, p(k)] \in (\mathbb{R}^l_+)^k$, agents make consumption plans $c = [c(1), c(2), \ldots, c(k)] \in (\mathbb{R}^l_+)^k$ (where $c(j)$ is the consumption in state j) for time 2.

We say that a pair $(c_i, \theta_i) \in \mathbb{R}^{lk}_+ \times \mathbb{R}^d$ is feasible if it satisfies the following constraints

$$
\begin{cases}
S \cdot \theta_i \leq 0 \\
p(j) \cdot c(j) \leq (V\theta_i)_j + p(j) \cdot e(j) & \text{for all } j = 1, \ldots, k .
\end{cases}
\tag{6.15}
$$

We define the budget set of agent i as the set of consumption plans that he could finance using his exogenous endowments and the income from the securities that he bought at time 1 (without going into debt). That is :

$$
B_i(p, S) = \left\{ c_i \in \mathbb{R}^{lk}_+ \mid \exists\, \theta_i \in \mathbb{R}^d , \; (c_i, \theta_i) \text{ satisfies (6.15)} \right\} .
\tag{6.16}
$$

We assume that the agents have preferences over the set of consumption plans, and that these are represented by utility functions $u_i : (\mathbb{R}^l_+)^k \to \mathbb{R}$ satisfying assumption **U1**.

Definition 6.5.1. *A Radner equilibrium is made up of*

- *a set of prices for the securities* $\overline{S} \in \mathbb{R}^d_+$,
- *expected prices* $\overline{p} \in (\mathbb{R}^l_+)^k$,
- *portfolios of assets* $(\overline{\theta}_1, \ldots, \overline{\theta}_m)$ *and consumption plans* $(\overline{c}_1, \ldots, \overline{c}_m)$ *such that*

  1. *a)* $\overline{c}_i$ *maximizes* $u_i(c_i)$ *under the constraint*

  $$
  c_i \in B_i(\overline{p}, \overline{S}) \qquad \text{for all } i = 1, \ldots, m ,
  $$

  b) $(\overline{c}_i, \overline{\theta}_i)$ *satisfies* (6.15).

  2. *The markets clear, i.e.,*

  a) $\sum_{i=1}^m \overline{c}_i = \sum_{i=1}^m e_i$,

  b) $\sum_{i=1}^m \overline{\theta}_i = 0$.

Remark 6.5.2. If we assume V to be injective, then the equality $\sum_{i=1}^m \overline{\theta}_i = 0$ is satisfied if 1 and 2b) are.

Indeed, suppose that the preferences are increasing and that the constraints are binding at the optimum. It is then the case that $\overline{p}(j) \cdot (\overline{c}_i(j) - e_i(j)) = (V\overline{\theta}_i)_j$, for all (i, j). By summing over i, and using 2a), we thus obtain $V\left(\sum_{i=1}^m \overline{\theta}_i\right) = 0$. As V is injective, $\sum_{i=1}^m \overline{\theta}_i = 0$.

Suppose now that $(\overline{p}, \overline{S}, \overline{c}_i, \overline{\theta}_i ; i = 1, \ldots, m)$ is a Radner equilibrium. A necessary condition for equilibrium is for there to be no arbitrage, i.e., there must not exist any portfolio $\theta \in \mathbb{R}^d$ satisfying $\overline{S} \cdot \theta \leq 0$ and $V\theta > 0$ (otherwise, the wealth of all the agents could become infinite, and there could be no

equilibrium). As we saw in Chap. 1, there then exists $\beta \in \mathbb{R}^k_{++}$ such that $\overline{S} = V^T \beta$.

Henceforth, we will distinguish two separate cases: the one in which rank $V = k$ (called the complete markets case) and the one in which rank $V < k$, the incomplete markets case.

6.6 The Complete Markets Case

We eliminate θ from (6.15) by multiplying the j-th line of (6.15) by β_j, and then summing over all the lines. Using the fact that $\overline{S} = V^T \beta$, we obtain

$$\sum_{j=1}^{k} \beta_j \overline{p}(j) \cdot [c_i(j) - e_i(j)] \leq 0 . \tag{6.17}$$

We set $p^*(j) = \beta_j \overline{p}(j)$, and obtain

$$\sum_{j=1}^{k} p^*(j) \cdot [c_i(j) - e_i(j)] \leq 0 . \tag{6.18}$$

Let us define
$$\overline{B}_i(p^*) = \{c \in \mathbb{R}^{lk}_+ \mid p^* \cdot c \leq p^* \cdot e_i\} .$$
Thus we have shown that $B_i(\overline{p}, \overline{S}) \subseteq \overline{B}_i(p^*)$.

Conversely, let us show that $\overline{B}_i(p^*) \subseteq B_i(\overline{p}, \overline{S})$ if $\overline{S} = V^T \beta$, $\beta \in \mathbb{R}^k_{++}$ and $\overline{p}(j) = \frac{p^*(j)}{\beta_j}$. Let $c_i \in \overline{B}_i(p^*)$. As rank $V = k$, there exists θ_i such that

$$(V\theta_i)_j = \frac{p^*(j)}{\beta_j} \cdot [c_i(j) - e_i(j)] = \overline{p}(j) \cdot [c_i(j) - e_i(j)] \quad \forall j . \tag{6.19}$$

Equation (6.17) entails that

$$\overline{S} \cdot \theta_i = \sum_{j=1}^{k} \beta_j (V\theta_i)_j \leq 0 , \tag{6.20}$$

and (6.19) and (6.20) then imply that $c_i \in B_i(\overline{p}, \overline{S})$.

Hence we can deduce the following result (sometimes called the equivalence theorem):

Theorem 6.6.1. *If $(\overline{p}, \overline{S}, \overline{c}_i, \overline{\theta}_i ; i = 1, \ldots, m)$ is a Radner equilibrium, then there exists $\beta \in \mathbb{R}^k_{++}$ such that $\overline{S} = V^T \beta$ and such that $(p^*, \overline{c}_i ; i = 1, \ldots, m)$ is a contingent Arrow–Debreu equilibrium with $p^*(j) = \overline{p}(j)\beta_j$.*

Conversely, if $(p^*, \overline{c}_i^* ; i = 1, \ldots, m)$ is a contingent Arrow–Debreu equilibrium, then for any $\beta \in \mathbb{R}_{++}^k$, there exists $\overline{\theta}_i$, $i = 1, \ldots, m$ such that $(\overline{p}, V^T \beta, \overline{c}_i, \overline{\theta}_i ; i = 1, \ldots, m)$ is a Radner equilibrium with $\overline{p}(j) = \dfrac{p^*(j)}{\beta_j}$.

Proof. The first implication follows from the equality $B_i(\overline{p}, V^T \beta) = \overline{B}_i(p^*)$. For the converse, let us assume that the k first column vectors of V are linearly independent. Write $V = (V^1, V^2)$ where V^1 is the matrix made up of the first k columns of V, and $\theta_i = (\theta_i^1, \theta_i^2) \in \mathbb{R}^k \times \mathbb{R}^{d-k}$. V^1 is then injective. Let us define $\overline{\theta}_i^1$ for all i, by

$$(V^1 \overline{\theta}_i^1)_j \;=\; \overline{p}(j) \cdot [\overline{c}_i(j) - e_i(j)] \quad \text{and} \quad \overline{\theta}_i^2 = 0 .$$

According to Remark 6.5.2, the equality $\sum_{i=1}^m \overline{c}_i = \sum_{i=1}^m e_i$ implies that $\sum_{i=1}^m \theta_1^i = 0$ (as V^1 is injective). Trivially, we have $\sum_{i=1}^m \overline{\theta}_2^i = 0$.   □

Under the assumptions made previously (the utility functions are continuous, strictly concave and increasing, and agents' endowments are strictly positive in all states), as there exists a contingent Arrow–Debreu equilibrium, we have the following corollary:

Corollary 6.6.2. *Under the assumption* **U1**, *if rank $V = k$ (i.e., if the markets are complete), for all $\beta \in \mathbb{R}_{++}^k$, there exists an equilibrium with financial markets, where $S = V^T \beta$.*

Remark 6.6.3. The proof given above is based on two ideas: that at equilibrium there is no arbitrage, and that markets are complete. These two ideas will recur in continuous time.

The Special Case of a One-Good Economy

Let us assume that there is only a single consumption good in each state of nature, and let us take it as numéraire (the dividend given by an asset in each state is expressed in units of the good). In this case, the spot price is identically equal to 1. Theorem 6.6.1 then becomes :

Theorem 6.6.4. *In the special case of an one-good economy, if $(\overline{S}, \overline{c}_i, \overline{\theta}_i ; i = 1, \ldots, m)$ is a Radner equilibrium in which the consumption good is taken as numéraire in each state of nature, then there exists $\overline{\beta} \in \mathbb{R}_{++}^k$ such that $\overline{S} = V^T \overline{\beta}$ and such that $(\overline{\beta}, \overline{c}_i ; i = 1, \ldots, m)$ is a contingent Arrow–Debreu equilibrium.*

Conversely, if $(p^*, \overline{c}_i ; i = 1, \ldots, m)$ is a contingent Arrow–Debreu equilibrium, then there exists $\overline{\theta}_i$, $i = 1, \ldots, m$ such that $(V^T p^*, \overline{c}_i, \overline{\theta}_i ; i = 1, \ldots, m)$ is a Radner equilibrium in which the consumption good is taken as the numéraire in each state of nature.

In the special case of a one-good economy, we see that there are as many equilibrium prices for assets as there are contingent Arrow–Debreu equilibrium prices in an economy. In general, the equilibrium is not unique. However, note the fundamental result, proved by Debreu [92]: for utility functions that are fixed "generically with respect to endowments", an exchange economy has a finite number of equilibria (we call the situation one of *determinacy*, as opposed to the *indeterminacy* of an infinite number of equilibria). In the special case where the agents' utility functions are additively separable, we can give conditions on agent's coefficients of relative risk aversion and endowments, in such at way that the equilibrium is unique.

Remark 6.6.5. Recall, under the assumption that there is only one consumption good in each state, the probabilistic interpretation of these results. As the market is complete, there is a riskless portfolio (whose payoff is equal to 1 in all states). For the sake of simplicity, we take it to be asset 0, as we will do later, in Chap. 7. Let the asset's price be S^0. We define the interest rate r by $S^0 = \dfrac{1}{1+r}$. Hence we deduce from the formula $S = V^T\beta$ that $\sum_{j=1}^k \beta_j = \frac{1}{1+r}$. Therefore, we can interpret the $\beta_j(1+r)$ as probabilities. It follows from Debreu's result that there are a finite number of interest rates r and probabilities, which are compatible with the equality of supply and demand. In some special cases (if there is a unique contingent Arrow–Debreu equilibrium), the interest rate and these probabilities are completely determined by the equality of supply and demand.

In the complete markets case, Theorems 6.6.1 and 6.6.4 show that the introduction of financial markets does not change agents' consumption at equilibrium. What is the point then, of introducing the financial markets? The aim is to reduce the number of markets or transactions. Indeed, in the one-period model with contingent markets, we need to open kl markets. In the model with financial markets, we need only open $d+l$ markets. When the number of states is sufficiently large, $d+l < kl$. When the number of dates is increased, in the first case the number of markets increases exponentially, whereas in the second, it increases linearly.

In the incomplete markets case, the introduction of financial markets changes agents' consumption at equilibrium.

Under the assumption of a single consumption good in each state of nature, whether or not the markets are complete, asset prices are determined by the equality of supply and demand. If there is more than one consumption good in each state of nature, and if the assets are nominal, then asset prices are subject to the condition of no arbitrage, and there is an indeterminacy in the price of the assets, which is linked to the choice of numéraire.

6.7 The CAPM

We now present the CAPM[4] as a special case of the Arrow-Radner model with two dates and with an infinite number of states at time 2. The CAPM is a necessary condition for equilibrium obtained under restrictive assumptions on agents' preferences. We will see in the next chapter how the introduction of continuous time enables us to relax the very restrictive assumptions of the CAPM. Nonetheless, we will be led to introduce an assumption of market completeness.

We consider an exchange economy with two dates, with a single consumption good, taken as the numéraire, and comprising m agents.

At time 1, there is a stock market comprising d assets (that is to say that the agents own shares of the assets). We assume that the payoff of the j-th asset at time 2 is a random variable d^j with finite variance, and defined on a probability space $(\Omega, \mathcal{F}, P)$. Without loss of generality, we can assume that the assets' payoffs are linearly independent.

Agent i has uncertain endowments for time 2, which are modeled by the random variable e_i defined on $(\Omega, \mathcal{F}, P)$. At time 1, he can modify his future resources by constructing a portfolio $\theta_i = (\theta_i^1, \ldots, \theta_i^d)$, on the condition that he does not run into debt. The resources that will be at his disposal at time 2 are: $c_i = e_i + \sum_{j=1}^{d} \theta_i^j d^j$.

Let C be the finite-dimensional vector space generated by the $(d^j, j = 1, \ldots, d)$. We endow C with a scalar product $\langle c, c' \rangle = E(cc')$. Let $\|c\|_2$ be the associated norm. Note that $\langle 1, c \rangle = E(c)$. We suppose that:

(i) $e_i \in C$ for all i (we can interpret e_i as a payoff on an initial portfolio).

(ii)   The aggregate wealth $e = \sum_{i=1}^{m} e_i$ is a.s. not equal to a constant.

(iii) $d^1 = 1$, that is to say that there exists a riskless asset. In a later remark, we will show how this condition can be relaxed.

In the following, we no longer assume **U1** to hold, but instead we make the following assumption:

U3   the agents have preferences on the elements of C, and these are represented by utility functions $U_i : C \to \mathbb{R}$, $(i = 1, \ldots, m)$, which have the property of "aversion to variance", that is to say that for any pair $(c, c') \in C^2$ satisfying $E(c) = E(c')$, the inequality $\text{Var}(c) < \text{Var}(c')$ implies that $U_i(c) > U_i(c')$.

Remark 6.7.1. It is usually assumed that agents' utility functions depend only on the expectation and on the variance of the random variables, that is, that they are of the form $\tilde{U}_i(E(c), \text{Var } c)$ where $\tilde{U}_i$ is increasing with respect to

---

[4] Capital Asset Pricing Model

its first coordinate and decreasing with respect to its second. Assumption **U5** is less restrictive. To prove the existence of an equilibrium, we would need stronger assumptions.

Given a set of prices $S \in \mathbb{R}^d$ for assets, agent i chooses a portfolio $\theta_i = (\theta_i^1, \ldots, \theta_i^d)$ in such a way as to

$$\text{maximize} \quad U_i\left(e_i + \sum_{j=1}^{d} \theta_i^j d^j\right) \quad \text{under the constraint}$$

$$S \cdot \theta_i \leq 0 .$$

The inequality above means that the agent cannot run into debt.

Given a price set S, the budget set is no longer bounded, as in the previous section. Therefore this problem may not have a solution.

Definition 6.7.2. *We say that $(\overline{S}, \overline{\theta}_i \, ; \, i = 1, \ldots, m)$ is in equilibrium if*

1. *For each i, $\overline{\theta}_i$ maximizes $U_i\left(e_i + \sum_{j=1}^{d} \theta_i^j d^j\right)$ under the constraint*
   $\overline{S} \cdot \theta_i \leq 0$.

2. *The security market clears, that is to say that $\sum_{i=1}^{m} \overline{\theta}_i = 0$.*

We do not discuss the existence of an equilibrium here. The CAPM is a necessary condition for equilibrium. Sufficient conditions are discussed in Nielsen [290, 291], (1989), and Allingham [6], (1991). An elementary proof of existence, under the assumption of the existence of a riskless asset (but without the assumption that C is finite dimensional) is to be found in Dana [83], (1999). In this case, there exists an equilibrium if the agents' utility functions are concave functions of the expectation and variance of the random variables, and are increasing with respect to the expectation coordinate and decreasing with respect to the variance coordinate. On the other hand, in the case where there is no riskless asset, satiation can lead to the non-existence of equilibrium.

Let $(\overline{S}, \overline{\theta}_i \, ; \, i = 1, \ldots, m)$ be an equilibrium. Let us consider the linear functional $\overline{\varphi}$ defined on C by

$$\overline{\varphi}(z) = \overline{S} \cdot \theta \quad \text{for} \quad z = \sum_{j=1}^{d} \theta^j d^j .$$

(i.e., if z is the value of the portfolio at time 2, $\overline{\varphi}(z)$ is its price at time 1).

Note that this linear functional is well-defined, as the d^j are linearly independent. It follows from Riesz's theorem, that there exists $\varphi \in C$ such that $\overline{\varphi}(z) = \langle \varphi, z \rangle$.

Since, by assumption, agent i behaves with aversion to variance, for a given expectation , at equilibrium, he minimizes

$$\text{Var}\left(e_i + \sum_{j=1}^{d} \theta_i^j d^j\right) \qquad \text{under the constraints}$$

$$\overline{S} \cdot \theta_i \leq 0 \quad \text{and} \tag{6.21}$$

$$E\left(e_i + \sum_{j=1}^{d} \theta_i^j d^j\right) = E\left(e_i + \sum_{j=1}^{d} \overline{\theta}_i^j d^j\right) .$$

We set $c_i = e_i + \sum_{j=1}^{d} \theta_i^j d^j$, and $\overline{c}_i = e_i + \sum_{j=1}^{d} \overline{\theta}_i^j d^j$. We have $\langle \varphi, c_i - e_i \rangle = \overline{S} \cdot \theta_i$. Since the expectation is taken to be fixed, the minimization problem comes down to

$$\begin{aligned} \text{minimize} \quad &\|c_i\|_2 \quad \text{under the constraints} \\ \langle \varphi, c_i \rangle \leq \ &\langle \varphi, e_i \rangle = a_0 \\ \langle 1, c_i \rangle = \ &\langle 1, \overline{c}_i \rangle = a_1 \\ &c_i \in c . \end{aligned} \tag{6.22}$$

Therefore, for any i, there are two Lagrange multipliers $\mu_i \geq 0$ and $\lambda_i \in \mathbb{R}$ such that

$$\overline{c}_i = \lambda_i - \mu_i \varphi \qquad \text{a.s. .}$$

Hence we deduce the existence of $\lambda \in \mathbb{R}$ and $\mu \geq 0$ such that

$$e = \sum_{i=1}^{m} \overline{c}_i = \lambda - \mu \varphi \qquad \text{a.s. .} \tag{6.23}$$

As by assumption e is not constant, μ is strictly positive. Therefore for all i, there exist $a_i \geq 0$ and $b_i \in \mathbb{R}$ such that

$$\overline{c}_i = a_i e + b_i \qquad \text{a.s. .} \tag{6.24}$$

We can rewrite (6.23) in the form

$$\varphi = -ae + b \qquad a \geq 0 , \ b \in \mathbb{R} . \tag{6.25}$$

Hence there exists $K \in \mathbb{R}$ such that

$$\overline{\varphi}(z) = -a \, \text{Cov}(e, z) + KE(z) , \qquad z \in C , \tag{6.26}$$

and in particular

$$\overline{S}^j = -a \, \text{Cov}(e, d^j) + KE(d^j) \qquad \text{for all} \quad j = 1, \ldots, d . \tag{6.27}$$

As the price $\overline{S}$ is viable, $\overline{S}^1 = \overline{\varphi}(1)$ is strictly positive.

Let us set $\overline{S}^1 = \dfrac{1}{1+r}$. We can then rewrite formula (6.27) as:

$$\overline{S}^j = -a \, \text{Cov} \, (e, d^j) + \frac{E(d^j)}{1+r} \, . \tag{6.28}$$

Hence we can deduce, for example, that the price of an asset that is positively correlated with aggregate wealth, is lower than the discounted expectation of its returns.

Given a portfolio $\theta \in \mathbb{R}^d$ such that $\overline{S} \cdot \theta \neq 0$, we define its return R_θ by

$$R_\theta = \sum_{j=1}^d \frac{\theta^j d^j}{\overline{S} \cdot \theta} \, . \tag{6.29}$$

Let $M = (M^1, \dots, M^j)$ be such that $e = \sum_{j=1}^d M^j d^j$. We call M the "market portfolio". We have $R_M = \dfrac{e}{\overline{S} \cdot M}$. Let us assume that $\overline{S} \cdot M = \overline{\varphi}(e) > 0$.

We now deduce the traditional beta formula from (6.28). We can rewrite (6.28) as

$$E(R_\theta) = (1+r)[1 + a \, \text{Cov} \, (e, R_\theta)] \, . \tag{6.30}$$

Hence, in particular,

$$E(R_M) = (1+r)[1 + a \, \text{Cov} \, (e, R_M)] \, . \tag{6.31}$$

Hence,

$$\begin{aligned}
E(R_\theta) - (1+r) &= [E(R_M) - (1+r)] \frac{\text{Cov} \, (e, R_\theta)}{\text{Cov} \, (e, R_M)} \\
&= [E(R_M) - (1+r)] \frac{\text{Cov} \, (R_M, R_\theta)}{\text{Var} \, R_M} \, .
\end{aligned} \tag{6.32}$$

We define the "beta" of a portfolio with respect to the market by the formula:

$$\beta_\theta = \frac{\text{Cov} \, (R_\theta, R_M)}{\text{Var} \, R_M} \, . \tag{6.33}$$

Hence we derive what is called the beta formula:

$$E(R_\theta) - (1+r) = \beta_\theta [E(R_M) - (1+r)] \, . \tag{6.34}$$

Now,

$$E(R_M) - (1+r) = \frac{a}{1+r} \, \text{Cov}(e, R_M) > 0 \, .$$

Therefore, if θ is a portfolio such that $\beta_\theta > 0$ then $E(R_\theta) > 1 + r$. We call $E(R_\theta) - (1+r)$ the risk premium of the portfolio. We summarize expressions (6.24) and (6.34) in the following theorems.

Theorem 6.7.3 (The Mutual Fund Theorem). *At equilibrium, the agents' demand can be split into a strictly positive demand for the market portfolio and a demand for the riskless asset.*

Theorem 6.7.4. *The risk premium of a portfolio is a linear function of its beta to aggregate endowment. If the market portfolio has a positive price, and if the return on a portfolio is positively (respectively negatively) correlated with aggregate endowment, the risk premium is positive (respectively negative). If it is not correlated with aggregate endowment, the expectation of its return is equal to the riskless return.*

Remark 6.7.5. When there is no riskless asset, we can carry out the same proof but replacing the function equal to one in every state, by its projection h onto the vector space generated by the d^j. Indeed, at equilibrium, agent i

$$\text{minimizes} \quad \|c_i\|_2 \quad \text{under the constraints}$$
$$\langle \varphi, c_i \rangle \;\leq\; a_0$$
$$E(c_i) \;=\; \langle 1, c_i \rangle \;=\; \langle h, c_i \rangle \;=\; a_1 \;.$$

Expression (6.24) becomes

$$\text{For all } i, \text{ there exist } a_i \geq 0 \text{ and } b_i \text{ such that}$$
$$\bar{c}_i = a_i e + b_i h \;, \tag{6.24bis}$$

and hence we obtain the mutual fund result.

Equation (6.25) becomes

$$\varphi = -ae + bh \quad , \quad a \geq 0 \;, \; b \in \mathbb{R} \;. \tag{6.25bis}$$

Recall expression (6.26), which remains unchanged

$$\overline{\varphi}(z) \;=\; -a \, \mathrm{Cov}(e, z) + K E(z) \;. \tag{6.26}$$

Under the assumption that $0 \;<\; \overline{\varphi}(e)$, we have $K > 0$, for if $K \leq 0$, we would have

$$0 \;<\; \overline{\varphi}(e) \;=\; -a \, \mathrm{Var}\,(e) + K \, E(e) \;<\; 0 \;,$$

and hence a contradiction.

Using the formula

$$K \, E(R_\theta) \;=\; 1 + a \, \mathrm{Cov}(e, R_\theta) \;, \tag{6.35}$$

we can easily show that assets that are not correlated with e have the same expected return $R' = \dfrac{1}{K}$. We obtain a new beta formula:

$$E(R_\theta) - R' \;=\; \beta_\theta [E(R_M) - R'] \;, \tag{6.36}$$

which shows that there is a linear relationship between the expected return on a portfolio and its beta to the market portfolio.

Notes

In this chapter, we have merely touched upon the theory of general equilibrium. The interested reader can refer to the classic works of Debreu [91], (1959) and Arrow and Hahn [14], (1971), or to more recent books by Ekeland [135], (1979), Hildenbrand–Kirman [191], (1989) and the article of Mas-Colell–Winston–Green [268], (1995). For the geometric approach to the theory of equilibrium, we reference the books by Balasko [20], (1988) and the article Mas-Colell [267], (1985). For problems linked to the computation of an equilibrium, we can look for example to articles by Scarf [323, 324], (1967). The book Border [38], (1985) is a classic reference for fixed point methods in general equilibrium theory.

For the theory of contingent markets, one can turn to the papers Arrow [12], (1953) and Debreu [91], (1959).

In this chapter, we have limited ourselves to the case of nominal assets, or of real assets in a world with a single consumption good. This case has been studied by Cass [49], (1984), Duffie [111], (1987), and Werner [364], (1985). When there are several consumption goods, the case described as the real case (where the gains from assets are expressed in terms of the consumption goods) is much more delicate. The first results were obtained by Duffie–Shafer [122, 123], (1985). We only obtain the generic existence of an equilibrium. For the theory of equilibrium with financial markets with either nominal or real assets, the reader can consult Tallon [354], (1995), for an introduction, the book Magill and Quinzii [262], (1994) and Magill and Shafer [264], (1991).

The classic version of the CAPM, in which the market contains a riskless asset, comes from Sharpe [333], (1964) and Lintner [256], (1965). The assumption of the existence of a riskless asset was relaxed by Black [36], (1972). The idea of using a method of projection is due to Chamberlain [52], (1985). Under the assumption that the utility function is Von Neumann–Morgenstern, Chamberlain [51], (1983) characterizes distributions for which the utility function has the "mean–variance" property. The econometric aspects of estimating the CAPM are covered by Huang–Litzenberger [197], (1988), which also provides an abundant bibliography of the subject. Nielsen [290, 291], (1989), Allingham [6], (1991) and Dana [83], (1999) give presentations of the CAPM as an equilibrium, as well as discussions of sufficient conditions for the existence of an equilibrium. One can also consult Leroy and Werner [252].

Factor pricing is discussed in Leroy and Werner [252], (2001). Sect. 6.7 differed from the others in that the agents could choose to have either positive or negative wealth (in general, agents choose to have non-negative consumption). As the set of agents' choice is unbounded, it is harder to prove the existence of an equilibrium. In particular, there is no reason to limit prices to being positive. The articles Page [296], (1996) and Dana et al [86], (1999), cover the problems that this kind of model presents, and gives a abundant bibliography of the subject.

ANNEX 6
The Theorem of the Maximum

Let $X \subseteq \mathbb{R}^l$ and $Y \subseteq \mathbb{R}^m$, let $f : X \times Y \to \mathbb{R}$ be a function such that $f(x, \cdot)$ is continuous for all x, and let $\Gamma : X \to Y$ be a non-empty compact valued correspondence. The aim of this section is to give sufficient conditions for the function defined by

$$h(x) = \max_{y \in \Gamma(x)} f(x, y) \tag{1}$$

to be continuous, and to study the properties of the correspondence defined by

$$G(x) = \{y \in \Gamma(x), \, f(x, y) = h(x)\} . \tag{2}$$

Definition *A correspondence $\Gamma : X \to Y$ is lower semi-continuous (l.s.c.) at x if for all $y \in \Gamma(x)$ and any sequence $x_n \to x$, there exists a sequence (y_n), $y_n \in \Gamma(x_n)$ for all n, such that $y_n \to y$.*

Definition *A non-empty compact valued correspondence $\Gamma : X \to Y$ is upper semi-continuous (u.s.c.) at x, if for any sequence $x_n \to x$ and for any sequence (y_n), $y_n \in \Gamma(x_n)$ for all n, the sequence (y_n) has a limit point $y \in \Gamma(x)$.*

Definition *A correspondence $\Gamma : X \to Y$ is continuous at x if it is upper and lower semi-continuous at x.*

Remark

1. When Γ is a function, Γ is continuous if and only if Γ is upper <u>or</u> lower semi-continuous (as a correspondence) .

2. If X is compact, Γ is upper semi-continuous at any point of X if and only if Γ has a closed graph.

Example

1. $\Gamma(x) = A$ for all x, where A is a non-empty compact set. Then Γ is obviously continuous.

2. $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is defined by $\Gamma(x) = [0, x]$. Then Γ is continuous.

3. $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is defined by $\Gamma(x) = [0, f(x)]$ with f continuous. Then Γ is continuous.

4. $\Gamma : \mathbb{R} \to \mathbb{R}$ is defined by

$$
\begin{array}{lll}
\Gamma(x) & = & \{\{-1\}, \{+1\}\} \qquad & \text{if } x > 0 \\
\Gamma(x) & = & [-1, 1] & \text{if } x = 0 \\
\Gamma(x) & = & 0 & \text{if } x < 0 \,.
\end{array}
$$

Then Γ is u.s.c. but not l.s.c. at zero.

5. Let $\Gamma : \mathbb{R} \to \mathbb{R}$ be defined by $\Gamma(x) = [-1, 1]$ if $x > 0$, $\Gamma(x) = \{0\}$ if $x \le 0$. Then Γ is l.s.c. at zero but is not u.s.c. at zero.

Theorem *Let $f : X \times Y \to \mathbb{R}$ be a continuous function and let $\Gamma : X \to Y$ be a continuous correspondence taking non-empty compact values. Then the function h defined in (1) is continuous, and the correspondence G defined in (2) is non-empty, compact valued, and upper semi-continuous.*

Example $X = \mathbb{R}$, $Y = [-1, +1]$, $f(x, y) = xy^2$, $\Gamma(x) = Y$ for all x. Then:

- $h(x) = \max\{xy^2\,,\; -1 \le y \le 1\} = x^+$,

- $G(x) = \{(-1), (1)\}$ if $x > 0$, $G(0) = [-1, 1]$ and $G(x) = \{0\}$ if $x < 0$,

- G is not lower semi-continuous at zero.

Corollary *If $f : X \times Y \to \mathbb{R}$ is a continuous function such that $f(x, \cdot)$ is strictly concave for all x, and if $\Gamma : X \to Y$ is a non-empty compact and convex valued continuous correspondence then G defined by (2) is a continuous function.*

7

Equilibrium of Financial Markets in Continuous Time. The Complete Markets Case

As we saw in the previous chapter, in discrete time the CAPM assumes very particular specifications for agents' preferences. Originally, continuous-time equilibrium models were introduced to relax these assumptions. The first models considered had one agent, and the methods used to analyze them were essentially those of dynamic programming (see for example Merton [273], Cox et al [70], Breeden [43]). It is only in the last decade that the problem of the existence of equilibrium in continuous time with financial markets has been addressed, and that various properties of the CAPM have been proven. Note that there is no existence result in the incomplete markets case.

In this chapter, we study only the case in which information is generated by a d-dimensional Brownian motion (this assumption can be relaxed). Under the assumption of complete markets, we show that analogous results to those in discrete time hold for equilibria in continuous time, i.e., there is an equivalence theorem for the Radner and Arrow–Debreu equilibria. This is a fundamental result, as it enables us on the one hand to transform the problem of the existence of a Radner equilibrium into that of the existence of an Arrow–Debreu equilibrium, and on the other hand, it shows that asset prices are fully determined at equilibrium. In the case of additively separable utility functions, we prove the existence of Arrow–Debreu equilibria, using the Negishi method. Thence we deduce the existence of Radner equilibria, and provide a characterize for them.

Finally, we show that we can recover the formulae given by Lucas [259], Cox–Ingersoll–Ross [70] and by Breeden [43].

7.1 The Model

As in Chaps. 3 and 4, we work on a finite time interval $[0, T]$ and a probability space $(\Omega, \mathcal{F}, P)$ on which a standard d-dimensional Brownian motion B_t is

given. As usual, information is modeled by the filtration of the Brownian motion, i.e., by $\mathcal{F}_t = \sigma(B_s\,,\ s \le t)$, completed by the addition of the null sets.

We consider an exchange economy with financial markets. In each state of the world, there is a single consumption good, which we take as numéraire.

7.1.1 The Financial Market

There are $d + 1$ real securities (their dividends are expressed in terms of the consumption good).

The first asset is riskless. Its price S^0 evolves according to the equation

$$\mathrm{d}S^0(t)\ =\ S^0(t)\,r(t)\,\mathrm{d}t,\quad S^0(0)\ =\ 1\,.$$

We write $R(t) = \exp[-\int_0^t r(s)\,\mathrm{d}s]$ for the discount factor. In this chapter, we do not assume r to be positive.

The other d assets are characterized by their cumulative dividend process $D = (D^1, \ldots, D^d)$, which we assume to follow an Itô process, and by their price S, which we also assume to follow an Itô process. These processes are then, in particular, continuous. We write $\tilde{S} = (S^0, S) = (S^0, S^1, \ldots, S^d)$ for the stock prices. We define the "discounted cumulative dividend process" D^{d} by $\mathrm{d}D^{\mathsf{d}}(t) = R(t)\mathrm{d}D(t)$ and $D^{\mathsf{d}}(0) = 0$.

The process $G^{\mathsf{d}} = SR + D^{\mathsf{d}}$, which is an Itô process, is called the discounted gains process. (This process was introduced in Sect. 3.4.4). In this chapter we make the following hypothesis:

H1   There exists on $(\Omega, \mathcal{F}_T)$ a probability measure Q, equivalent to P with density $\xi = \frac{\mathrm{d}Q}{\mathrm{d}P}$, such that

$$\mathrm{d}G^{\mathsf{d}}_t\ =\ \sigma^{\mathsf{d}}_G(t)\,\mathrm{d}\tilde{B}_t\,, \tag{7.1}$$

where $\tilde{B}_t$ is a $\mathcal{F}_t$-Q Brownian motion and where the process ξR is uniformly bounded. We assume that $\sigma^{\mathsf{d}}_G(t)$ is $\mathcal{F}_t$-measurable and invertible.

Under (**H1**), any continuous Q-$\mathcal{F}_t$-martingale can be written as a stochastic integral with respect to G^{d}. In what follows, we will use the notation

$$\Theta(G^{\mathsf{d}})\ =\ \left\{\tilde{\theta} = (\theta^0, \theta)\,,\ \tilde{\theta} \text{ predictable and } \left[\int_0^T \|\theta(t)\sigma^{\mathsf{d}}_G(t)\|^2\,\mathrm{d}t\right] < \infty \ a.s.\right\},$$

$$\mathcal{H}(G^{\mathsf{d}})\ =\ \left\{\tilde{\theta} \in \Theta(G^{\mathsf{d}})\ \Big|\ \int_0^t \theta(s)\,\mathrm{d}G^{\mathsf{d}}(s) \text{ is a } Q\text{-}\mathcal{F}_t\text{-martingale}\right\}.$$

Remark 7.1.1. We could have proceeded by noticing that the discounted gains process is an Itô process, and constructing a probability measure Q equivalent to P, and a Brownian motion $\tilde{B}$ such that $dG_t = \sigma_G^\mathsf{d}(t)\,d\tilde{B}_t$, using Girsanov's theorem. We do not take this route however, so as not to introduce additional assumptions and problems of measurability.

Remark 7.1.2. If $\tilde{\theta} \in \mathcal{H}(G^\mathsf{d})$, then $\int_0^t \theta(s)\,dG^\mathsf{d}(s)$ is a Q-$\mathcal{F}_t$-martingale that is zero at time zero. Therefore $E_Q\left(\int_0^t \theta(s)\,dG^\mathsf{d}(s)\right) = 0$, for all $t \in [0,T]$.

Let us customize some of the definitions introduced in Chaps. 3 and 4 to the context of our model.

Let Z be a $\mathcal{F}_T$-measurable random variable. A trading strategy $\tilde{\theta} \in \mathcal{H}(G^\mathsf{d})$ finances Z if

(i) $R(t)(\tilde{\theta}_t \cdot \tilde{S}_t) = \tilde{\theta}_0 \cdot \tilde{S}_0 + \int_0^t \theta(s)\,dG^\mathsf{d}(s)$ for all $t \in [0,T]$,

(ii) $\tilde{\theta}_T \cdot \tilde{S}_T = Z$.

An "arbitrage opportunity" is a strategy $\tilde{\theta} \in \mathcal{H}(G^\mathsf{d})$ with non-positive initial value $\tilde{\theta}_0 \cdot \tilde{S}_0 \le 0$ and a non-negative terminal value Z of strictly positive expectation.

In the model that we are studying, there are no arbitrage opportunities. Indeed, if $\tilde{\theta} \in \mathcal{H}(G^\mathsf{d})$, then $\int_0^t \theta(s)\,dG^\mathsf{d}(s)$ is a Q-martingale that is zero at time zero, and therefore $E_Q\left[\int_0^t \theta(s)\,dG^\mathsf{d}(s)\right] = 0$. Hence $E_Q(R(T)Z) = 0$, and this contradicts the conditions for an arbitrage opportunity.

7.1.2 The Economy

We consider a single consumption good and m agents described by the list: $(L_+^1, (U_i, e_i)\,,\ i = 1, \ldots, m)$ where the space

$$L_+^1 = \left\{ c : \Omega \times [0,T] \to \mathbb{R}_+ \quad \mathcal{F}_t\text{-adapted} \mid E_P\left[\int_0^T c(t)dt\right] < \infty \right\}$$

represents the set of consumption processes available to agents, where $U_i : L_+^1 \to \mathbb{R}_+$ is the utility function of agent i, and where $e_i \in L_+^1$ is his endowment stream.

Though, with regard to the existence of an equilibrium, it would be possible to consider much more general preferences, we restrict ourselves here to the additively separable utility functions,

$$U_i(c) = E_P\left[\int_0^T u_i(t, c(t))\,dt\right].$$

We make the following assumptions concerning the functions u_i:

U1 $u_i(t, \cdot)$ is strictly concave and strictly increasing for all t.

U2 u_i belongs to $C^{0,1}([0, T] \times [0, \infty[)$ and $\frac{\partial u_i}{\partial c} : [0, T] \times \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}_+$ is continuous.

U3 u_i is of class $C^{1,4}$ and $\frac{\partial u_i}{\partial c}(t, 0) = \infty$ for all t.

7.1.3 Admissible Pairs

Agent i holds a portfolio of stocks $\tilde{\theta}_i = (\theta_i^0, \theta_i)$. We suppose that $\tilde{\theta}_i \in \mathcal{H}(G^d)$.

Definition 7.1.3. *Given the discounted gains process of the stocks, the pair $(\tilde{\theta}_i, c_i) \in \mathcal{H}(G^d) \times L_+^1$ is admissible for agent i if it satisfies*

$$R(t)(\tilde{\theta}_i(t) \cdot \tilde{S}_t) = \int_0^t \theta_i(s) dG^d(s) - \int_0^t R(s)[c_i(s) - e_i(s)] ds \tag{7.2}$$
$$Q \text{ a.s. for all } t \in [0, T] ,$$

and
$$\tilde{\theta}_i(T) \cdot \tilde{S}_T = \theta_i^0(T) S^0(T) + \theta_i(T) \cdot S_T \geq 0 \qquad a.s. . \tag{7.3}$$

Equation (7.2) says that the discounted wealth at time t is the sum of the discounted gains or losses produced by the exchange of stocks and of consumption goods. Equation (7.3) supposes there no debt remains at the end of the period. We note firstly that the wealth of an agent is not required to be positive at each instant in time, and secondly that $G^d(t)$, which is a priori a $d + 1$ dimensional process, has a constant first component, so that it will later be considered as a d dimensional process for the purpose of integration.

Remark 7.1.4. If we disregard issues of integrability for a moment, it is natural to define admissible portfolios by

$$\tilde{\theta}_i(t) \cdot \tilde{S}_t = \int_0^t \tilde{\theta}_i(s) d\tilde{G}(s) - \int_0^t (c_i(s) - e_i(s)) ds ,$$

where $\tilde{G} = (S^0, S + D)$, and by (7.3). Let $dZ_t = R(t)(\tilde{\theta}_i(t) \cdot \tilde{S}_t)$. Itô's lemma leads to

$$dZ_t = R(t)\tilde{\theta}_i(t) d\tilde{G}_t - R(t)[c_i(t) - e_i(t)] dt + \tilde{\theta}_i(t) \cdot \tilde{S}_t dR_t .$$

As $R_t dS_t^0 + S_t^0 dR_t = 0$,

$$dZ_t = \theta_i(t) dG_t^d - R(t)[c_i(t) - e_i(t)] dt .$$

Proposition 7.1.5. *We have:*

a) If $(\tilde{\theta}_i, c_i)$ is admissible, then c_i satisfies

$$E_Q\left[\int_0^T R(s)[c_i(s) - e_i(s)]\,\mathrm{d}s\right] \leq 0. \qquad (7.4)$$

b) Conversely, if c_i satisfies (7.4), then there exists $\tilde{\theta}_i \in \mathcal{H}(G^d)$ such that $(\tilde{\theta}_i, c_i)$ is admissible.

Proof. The first implication is obvious, since if $\tilde{\theta}_i \in \mathcal{H}(G^d)$, then $\int_0^t \theta_i(s)\,\mathrm{d}G^a(s)$ is a Q-martingale that is zero at time zero, and so $E_Q\left[\int_0^T \theta_i(s)\,\mathrm{d}G^d(s)\right] = 0$. Therefore we have

$$E_Q\left[\int_0^T R(s)[e_i(s) - c_i(s)]\mathrm{d}s\right] = E_Q\left[\theta_i^0(T)S^0(T) + \theta_i(T)\cdot S_T\right] \geq 0.$$

To prove the converse, we first note that

$$E_Q\left[\int_0^T R(s)[e_i(s) - c_i(s)]\,\mathrm{d}s\right] = E_P\left[\int_0^T \xi(s)R(s)[e_i(s) - c_i(s)]\,\mathrm{d}s\right] < \infty.$$

We introduce the Q-martingale :

$$Y_t = E_Q\left[\int_0^T R(s)[e_i(s) - c_i(s)]\,\mathrm{d}s\right] + E_Q\left[\int_0^T R(s)[c_i(s) - e_i(s)]\mathrm{d}s|\mathcal{F}_t\right]. \qquad (7.5)$$

As $E_Q(Y_T) = 0$, we have $E_Q(Y_t) = 0$, for all t.

From the predictable representation theorem (cf. annex to Chap. 4), there exists $\theta_i = (\theta_i^1, \ldots, \theta_i^d)$ with $\int_0^T \|\theta_i(t)\sigma_G^d(t)\|^2\,\mathrm{d}t < \infty$ *a.s.*, such that

$$Y_t = \int_0^t \theta_i(s)\,\mathrm{d}G^d(s). \qquad (7.6)$$

Let us consider the process $X_i(t)$ defined by

$$R(t)X_i(t) = \int_0^t \theta_i(s)\,\mathrm{d}G^d(s) + \int_0^t R(s)[e_i(s) - c_i(s)]\,\mathrm{d}s.$$

Let $\theta_i^0(t)$ be defined by $\theta_i^0(t)S^0(t) = X_i(t) - \theta_i(t)\cdot S_t$ and $\tilde{\theta}_i = (\theta_i^0, \theta_i)$. By construction, $(\tilde{\theta}_i) \in \mathcal{H}(G^d)$. It follows from (7.5) and (7.6) that

$$R(T)X_i(T) = E_Q\left[\int_0^T R(s)[e_i(s) - c_i(s)]\,\mathrm{d}s\right] \geq 0.$$

$(\tilde{\theta}_i, c_i)$ satisfies (7.2) and (7.3), and is therefore admissible. $\square$

7.1.4 Definition and Existence of a Radner Equilibrium

Definition 7.1.6. *For a given $\tilde{S}$ and D, the pair $(\tilde{\theta}, c) \in \mathcal{H}(G^d) \times L^1_+$ is optimal for agent i if it is admissible and if it maximizes the utility of the agent over the set of admissible strategies.*

Definition 7.1.7. *For a given dividend process D, the list $\{(\tilde{\theta}_i, c_i) \in \mathcal{H}(G^d) \times L^1_+$, $(i = 1, \ldots, m)$; $\tilde{S}\}$ is a Radner equilibrium if:*

(i) the pair $(\tilde{\theta}_i, c_i)$ is optimal for all $i = 1, \ldots, m$,

(ii) markets clear, that is to say that
(a) $\sum_{i=1}^m \tilde{\theta}_i = 0 \qquad P \otimes dt-$ a.s..
(b) $\sum_{i=1}^m c_i = \sum_{i=1}^m e_i \qquad P \otimes dt-$ a.s..

Remark 7.1.8. As in Chap. 6, Remark 6.5.2, we can see that if (i) and (ii)(b) hold, then necessarily (ii)(a) is also satisfied.

Indeed, let $\tilde{\theta}(t) = \sum_{i=1}^m \tilde{\theta}_i(t)$. Define θ by $\tilde{\theta} = (\theta^0, \theta)$. By summing the equalities (7.2) over all the agents, we obtain for all $t \in [0, T]$,

$$
R(t)(\tilde{\theta}(t) \cdot \tilde{S}_t) = \int_0^t \theta(s)\, dG^d(s) + \int_0^t \left[\sum_{i=1}^m R(s)[e_i(s) - c_i(s)] \right] ds
$$

$$
= \int_0^t \theta(s)\, dG^d(s) .
$$

At the final date T, as preferences are monotonous, constraints are binding thus $\tilde{\theta}_T \cdot \tilde{S}_T = 0$, and consequently

$$
0 = \int_0^T \theta(s)\, dG^d(s) .
$$

As the process $0 = \int_0^t \theta(s)\, dG^d(s)$ is a martingale, its increasing process is zero. Hence $\theta(t) \cdot \sigma_G^d(t) = 0$, $Q \otimes dt$ a.e. and so also $P \otimes dt$ a.e.. Since σ_G^d is invertible, $\theta(t) = 0$, $P \otimes dt$ a.s.. Hence $\tilde{\theta}(t) \cdot \tilde{S}_t = 0$ and so $\theta^0(t) = 0$, $P \otimes dt$ a.s..

In what follows:

$$
L^\infty_+ = \Big\{ p : \Omega \times [0, T] \to \mathbb{R}_+ \quad \mathcal{F}_t\text{-adapted} \;\Big|
$$

$$
\exists\, M \text{ such that } p < M \quad P \otimes dt \text{-a.s.} \Big\} .
$$

Let us recall the following definition:

Definition 7.1.9. *The list $[\bar{c}_i, (i = 1, \ldots, m); p] \in (L^1_+)^m \times L^\infty_+$ is a contingent Arrow–Debreu P-equilibrium if*

(i) $\bar{c}_i$ maximizes the utility of agent i under the constraint

$$E_P \left[\int_0^T p(s)(c_i(s) - e_i(s))\,\mathrm{d}s \right] \leq 0 \,,$$

(ii) $\sum_{i=1}^m \bar{c}_i = \sum_{i=1}^m e_i = : e \qquad P \otimes \mathrm{d}t$ a.s..

Using Proposition 7.1.5 and Remark 7.1.8, we now prove an analogous theorem to Theorem 6.6.4.

Theorem 7.1.10. *If $\left((\tilde{\theta}_i, c_i)\,, i = 1, \ldots, m\,; \tilde{S}\, \right)$ is a Radner equilibrium such that* **(H1)** *holds, then $(c_i\,, i = 1, \ldots, m\,; R\xi)$ is a contingent Arrow–Debreu equilibrium.*

Conversely, suppose that $(c_i\,, i = 1, \ldots, m\,; p)$ is a contingent Arrow–Debreu P-equilibrium such that p is an Itô process $\mathrm{d}p_t = \mu_p(t)\,\mathrm{d}t + \sigma_p(t)\,\mathrm{d}B_t$ with strictly positive values and satisfying $p(0) = 1$ and such that $E_P \left(\exp \frac{1}{2} \int_0^T \frac{\|\sigma_p\|^2(t)}{p^2(t)}\,\mathrm{d}t \right) < \infty$. Then there exists a strictly positive P-$\mathcal{F}_t$-martingale ξ and a discount process R, which are unique and such that $p = R\xi$. Let Q be the probability measure on $(\Omega, \mathcal{F}_T)$ that is equivalent to P and of density ξ_T. For any set of prices $\tilde{S}$ such that **(H1)** *is satisfied under Q, there exists $\left(\tilde{\theta}_i \in \mathcal{H}(G^{\mathsf{d}})\,, i = 1, \ldots, m \right)$ such that $((\tilde{\theta}_i, c_i)\,, i = 1, \ldots, m\,; \tilde{S})$ is a Radner equilibrium.*

Proof. The first part of the theorem follows from Proposition 7.1.5.

To prove the second part, we suppose that there exists ξ, a strictly positive P-$\mathcal{F}_t$-martingale and R a discount process, such that $p = R\xi$. As ξ is a positive P-martingale, there exists a unique predictable process q_t that satisfies[1] $\int_0^T q_t^2\,\mathrm{d}t < \infty$ P a.s. and $\mathrm{d}\xi_t = q_t\xi_t\,\mathrm{d}B_t$. We have

$$\begin{aligned}
\mathrm{d}(R\xi)_t &= -r(t)R(t)\xi_t\mathrm{d}t + R(t)\mathrm{d}\xi_t \\
&= -r(t)R(t)\xi_t\,\mathrm{d}t + R(t)q_t\xi_t\mathrm{d}B_t = \mu_p(t)\mathrm{d}t + \sigma_p(t)\mathrm{d}B_t\,.
\end{aligned}$$

By identifying the coefficients of the drift and diffusion terms, we obtain $-r_t R_t \xi_t = \mu_p(t)$ and $R_t q_t \xi_t = \sigma_p(t)$, and hence $r_t = -\frac{\mu_p(t)}{p(t)}$ and $q_t = \frac{\sigma_p(t)}{p(t)}$.

We suppose that p is an Itô process $\mathrm{d}p_t = \mu_p(t)\,\mathrm{d}t + \sigma_p(t)\,\mathrm{d}B_t$ with strictly positive values and such that $E_P \left(\exp \frac{1}{2} \int_0^T \frac{\|\sigma_p\|^2(t)}{p^2(t)}\,\mathrm{d}t \right) < \infty$ and $p(0) = 1$. We define $r = -\frac{\mu_p}{p}$, $R(t) = \exp \left[-\int_0^t r(s)\,\mathrm{d}s \right]$ and $q = \frac{\sigma_p}{p}$. Let us show that $\frac{p}{R}$ is a positive P-martingale. Indeed,

$$\mathrm{d}\left(\frac{p}{R} \right)_t = \frac{1}{R_t}\left(\mu_p(t)\,\mathrm{d}t + p_t q_t\,\mathrm{d}B_t \right) + r_t p_t \frac{\mathrm{d}t}{R(t)} = q_t \left(\frac{p}{R} \right)_t \mathrm{d}B_t\,.$$

---

[1] See for example Revuz–Yor [307] p. 304.

(Cf. annex to Chap. 3.) For any price set $\tilde{S}$ such that (**H1**) is satisfied under Q, since c_i satisfies (7.4) for any i, there exists $(\tilde{\theta}_i \in \mathcal{H}(G^d)$, $i = 1, \ldots, m)$ such that the pair $(\tilde{\theta}_i, c_i)$ is optimal for all $i = 1, \ldots, m$. Finally, according to Remark 7.1.8, it is enough to check that $\sum_{i=1}^m c_i = \sum_{i=1}^m e_i$, which follows from the definition of an Arrow–Debreu equilibrium. $\square$

7.2 Existence of a Contingent Arrow–Debreu Equilibrium

As is often the case in infinite dimensional economies, we use the Negishi method. Under the assumption of the separability of the utility functions, the calculations below are a straightforward transposition of those of Sect. 6.3. We assume here that (**U1**) and (**U2**) hold for all i.

7.2.1 Aggregate Utility

We introduce the following notation. Let $\alpha \in \Delta^{m-1}$ and $c \in \mathbb{R}_{++}$.

$$u(t, c, \alpha) = \max \left\{ \sum_{j=1}^m \alpha_j u_j(t, x_j) \mid \sum_{j=1}^m x_j \le c , \ x_j \ge 0 \text{ for all } j \right\} \quad (7.7)$$

and

$$(C_i(t, c, \alpha))_{i=1}^m = \arg \max \left\{ \sum_{j=1}^m \alpha_j u_j(t, x_j) \mid \right.$$
$$\left. \sum_{j=1}^m x_j \le c , \ x_j \ge 0 \text{ for all } j \right\}. \quad (7.8)$$

First of all, we obtain the following result.

Proposition 7.2.1.

1. *If u_i satisfies (**U1**) for all i, then $u(t, \cdot, \alpha)$ satisfies (**U1**) .*
2. *If u_i satisfies (**U2**) for all i, then $u(t, \cdot, \alpha)$ satisfies (**U2**) .*
3. *If u_i satisfies (**U3**) for all i, then $u(\cdot, \cdot, \alpha)$ is of class $C^{1,3}$.*

Proof. The proof of 1 is left to the reader.

Assuming that $\frac{\partial u_i}{\partial c}(t, 0) = \infty$ for all t and for all i, the proof of 2 follows that of Proposition 6.3.4. Note that in this case, $\frac{\partial u}{\partial c}(t, c, \alpha) = \alpha_i \frac{\partial u_i}{\partial c}[t, C_i(t, c, \alpha)]$

for all i such that $\alpha_i > 0$. The proof can be extended to the case of a general $\frac{\partial u_i}{\partial c}(t, 0)$. However as the extension is a little technical, we will admit the result (the proof can be found in Dana–Pontier [85]).

To show that 3 holds, note that as in Proposition 6.3.4, there exists $\lambda \in \mathbb{R}_+$ such that when $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_p > 0$ and $\alpha_{p+1} = 0 = \cdots = \alpha_m$, we have:

$$
\begin{cases}
\alpha_1 \dfrac{\partial u_1}{\partial c}[t, C_1(t, c, \alpha)] = & \lambda \\[2mm]
\quad\vdots & \vdots \\[2mm]
\alpha_p \dfrac{\partial u_p}{\partial c}[t, C_p(t, c, \alpha)] = & \lambda \\[2mm]
\displaystyle\sum_{i=1}^{m} C_i(t, c, \alpha) = & c .
\end{cases}
\tag{7.9}
$$

As in the proof of Proposition 6.3.4, it follows from the implicit function theorem that $(C_i(\cdot, \cdot, \alpha))_{i=1}^{m}$ is of class $C^{1,3}$. Hence $u(\cdot, \cdot, \alpha)$ is also of class $C^{1,3}$. $\qquad\square$

Remark 7.2.2. Let us write $u_{i,cc}$ for the second derivative of u_i, and $u_{cc}(t, \cdot, \alpha)$ for that of $u(t, \cdot, \alpha)$. We will obtain explicit expressions for them later. As we showed in Proposition 6.3.4,

$$
\frac{\partial u}{\partial c}(t, c, \alpha) = \lambda .
$$

Hence

$$
-u_{cc}(t, c, \alpha) = -\frac{\partial \lambda}{\partial c}
$$

$$
= -\frac{1}{\sum_{i=1}^{m} \frac{1}{\alpha_i} \frac{1}{u_{i,cc}(t, C_i(t,c,\alpha))}} \leq -\alpha_i u_{i,cc}(t, C_i(t, c, \alpha)) \ \forall i
$$

and

$$
\frac{\partial C_i}{\partial c}(t, c, \alpha) = \frac{1}{\alpha_i u_{i,cc}[t, C_i(t, c, \alpha)]} \frac{\partial \lambda}{\partial c} = \frac{u_{cc}(t, c, \alpha)}{\alpha_i u_{i,cc}[t, C_i(t, c, \alpha)]} .
$$

In particular, we note that $\frac{\partial C_i}{\partial c}(t, c, \alpha) > 0$.

7.2.2 Definition and Characterization of Pareto Optima

Let us recall the two following definitions:

Definition 7.2.3. *The consumption vector* $(c_i)_{i=1}^{m} \in (L_+^1)^m$ *is admissible if it satisfies*

$$
\sum_{i=1}^{m} c_i \leq e \qquad P \otimes dt \ a.s. .
$$

Let $\mathcal{A}$ be the set of admissible consumption vectors.

Definition 7.2.4. *An admissible consumption vector $(\bar{c}_i)_{i=1}^m$ is Pareto-optimal if there does not exist a vector $(c_i)_{i=1}^m \in \mathcal{A}$ satisfying:*

$$U_j(c_j) \ge U_j(\bar{c}_j) \quad \text{for all } j \text{ and } U_{j_0}(c_{j_0}) > U_{j_0}(\bar{c}_{j_0}) \quad \text{for at least one } j_0 \,.$$

We can characterize Pareto-optimal allocations as follows:

Proposition 7.2.5. *An admissible consumption vector $(\bar{c}_i)_{i=1}^m$ is Pareto-optimal if and only if there exists $\alpha \in \Delta^{m-1}$ such that*

$$\bar{c}_i(t,\omega) = C_i(t, e(t,\omega), \alpha) \qquad P \otimes \mathrm{d}t \text{ a.s. for all } i = 1, \ldots, m \,.$$

Proof. As in the proof of Proposition 6.3.3, we show that $(\bar{c}_i)_{i=1}^m$ is Pareto-optimal if and only if there exists $\alpha \in \Delta^{m-1}$ such that $(\bar{c}_i)_{i=1}^m$ is an optimal solution to the problem:

$$\text{Maximize} \quad \sum_{i=1}^m \alpha_i \, U_i(x_i) \quad \text{under the constraint} \quad x = (x_i)_{i=1}^m \in \mathcal{A} \,.$$

As the utility functions are additively separable,

$$\max_{x \in \mathcal{A}} \sum_{i=1}^m \alpha_i \, U_i(x_i) = E_P\left[\int_0^T u(t, e(t), \alpha) \, \mathrm{d}t \right],$$

and hence $\bar{c}_i(t,\omega) = C_i(t, e(t,\omega), \alpha) \, P \otimes \mathrm{d}t$ a.s.. $\qquad\square$

As in equation (6.10) of Sect. 6.3, we set

$$p(t, \alpha) = \frac{\partial u}{\partial c}[t, e(t), \alpha] \,. \tag{7.10}$$

We have the following result:

Proposition 7.2.6. *If there exists a constant $k > 0$ such that $k \le e$, $P \otimes \mathrm{d}t$ a.s., then $p(t,\alpha)$ belongs to L_+^∞. For all $\alpha_0 \gg 0$ and for all i, there exists $k_i > 0$ such that $C_i[t, e(t), \alpha_0] \ge k_i$, $P \otimes \mathrm{d}t$ a.s..*

Proof. Let us assume that there exists $k > 0$ such that $k \le e, P \otimes \mathrm{d}t$ a.s.. As $\frac{\partial u}{\partial c}(t, \cdot, \alpha)$ is a decreasing function, we have:

$$p(t, \alpha) \le \frac{\partial u}{\partial c}(t, k, \alpha) \le \max_{s,\beta} \frac{\partial u}{\partial c}(s, k, \beta) \,.$$

Under (**U2**), the maximum in the expression above exists, since according to 2 of Proposition 7.2.1 the function $\frac{\partial u}{\partial c}(\cdot, k, \cdot)$ is continuous, and since the set $[0, T] \times \Delta^{m-1}$ is compact.

Moreover, if $\frac{\partial u_i}{\partial c}(t,0) = \infty$ for all t and for all i, and if the functions u_i are of class $C^{1,2}$, it follows from Proposition 7.2.1 and from Remark 7.2.2 that $C_i(t,\cdot,\alpha)$ is an increasing function. We then have:

$$C_i\left(t,e(t),\alpha_0\right) \geq C_i(t,k,\alpha_0) \geq k_i\,,$$

with

$$k_i = \inf_s C_i(s,k,\alpha_0) > 0\,.$$

We can show that the monotonicity of $C_i(t,\cdot,\alpha)$ still holds when the u_i are of class $C^{1,1}$. Thus the lower bound given above is still valid. $\qquad\square$

Proposition 7.2.7. *If $(\bar{c}_i)_{i=1}^m$ is Pareto-optimal, then there exists $\alpha \in \Delta^{m-1}$ such that $\bar{c}_i(t) = C_i(t,e(t),\alpha)$ for all i and $\bar{c}_i$, is the optimal solution to the problem $\mathcal{P}_i$*

$$\left\{\begin{array}{l} maximize \quad U_i(x_i) \quad under\ the\ constraint \\[2mm] E_P\left[\displaystyle\int_0^T p(t,\alpha)x_i(t)\,\mathrm{d}t\right] \leq E_P\left[\displaystyle\int_0^T p(t,\alpha)\bar{c}_i(t)\,\mathrm{d}t\right]\,. \end{array}\right. \qquad (\mathcal{P}_i)$$

where p is defined as in (7.10).

Proof. As in the proof of Proposition 7.2.1 assertion 2, we make the assumption, which can later be relaxed, that $\frac{\partial u_i}{\partial c}(t,0) = \infty$ for all t and all i. According to Proposition 7.2.5, there exists $\alpha \in \Delta^{m-1}$ such that $\bar{c}_i(t) = C_i[t,e(t),\alpha]$.

In addition,

$$\begin{array}{ll} if \quad \alpha_i > 0 & C_i[t,e(t),\alpha] > 0 \quad and \\ if \quad \alpha_i = 0 & C_i[t,e(t),\alpha] = 0\,. \end{array}$$

First note that the statement of the proposition holds trivially in the case where $\alpha_i = 0$. When $\alpha_i > 0$, as $p(t,\alpha) = \alpha_i \frac{\partial u_i}{\partial c}[t,C_i(t,e(t),\alpha)]$, we have for all $x_i \in L_+^1$,

$$\begin{aligned} u_i[t,\bar{c}_i(t)] - u_i[t,x_i(t)] &\geq \frac{\partial u_i}{\partial c}[t,\bar{c}_i(t)]\,[\bar{c}_i(t) - x_i(t)] \\ &\geq \frac{p(t,\alpha)}{\alpha_i}\,[\bar{c}_i(t) - x_i(t)]\,. \end{aligned}$$

Hence by integrating with respect to (t,ω), we obtain

$$U_i(\bar{c}_i) - U_i(x_i) \geq E_P\left[\int_0^T \frac{p(t,\alpha)}{\alpha_i}\,[\bar{c}_i(t) - x_i(t)]\,\mathrm{d}t\right],$$

and hence the result. $\qquad\square$

7.2.3 Existence and Characterization of a Contingent Arrow–Debreu Equilibrium

Theorem 7.2.8. *Suppose that u_i satisfies (**H1**) and (**H2**) for all i. If there exists $k > 0$ such that $k \leq e \quad P \otimes \mathrm{d}t$ a.s., then there exists a contingent Arrow–Debreu equilibrium of the form $\left(C_i[t, e(t), \alpha_0], \ i = 1, \ldots, m; \ \frac{\partial u}{\partial c}(t, e(t), \alpha_0) \right)$ with $\alpha_0 \gg 0$.*

Proof. As in the proof of Theorem 6.3.5, we introduce the transfer function $\Phi : \Delta^{m-1} \to \mathbb{R}^m$ defined by

$$\Phi_i(\alpha) \ = \ E_P \left[\int_0^T p(t, \alpha) \left(C_i(t, e(t), \alpha) - e_i(t) \right) \mathrm{d}t \right] .$$

As $C_i[t, e(t), \cdot]$ and $p(t, \cdot)$ are continuous functions, and as $C_i[t, e(t), \alpha] \leq e(t)$ for all i and $p(t, \alpha) \in L^\infty$, it follows from the dominated convergence theorem that Φ_i is a continuous function for all i. As $\sum_{i=1}^m C_i[t, e(t), \alpha] = e(t)$, we have $\sum_{i=1}^m \Phi_i = 0$, and if $\alpha_i = 0$, $\Phi_i(\alpha) < 0$. The transfer function $\Phi = (\Phi_i)_{i=1}^m$ therefore satisfies the conditions of Theorem 6.3.5. From this, we deduce the existence of a zero $\alpha_0 \gg 0$ for Φ, and hence the existence of an equilibrium of the form $\left(C_i[t, e(t), \alpha_0], \ i = 1, \ldots, m; \ p(t, \alpha_0) \right)$. $\qquad\square$

Remark 7.2.9. The equilibrium allocations being Pareto-optimal, a consequence of this section is that all the equilibrium allocations take this same form. We can show that the same is true of equilibrium prices. Furthermore, we can show that when we fix k and fix the utility functions generically with respect to endowments, then the transfer function has a finite number of zeros.

7.2.4 Existence of a Radner Equilibrium

Henceforth, we make the following hypotheses:

H2 e is an Itô process of the form $\mathrm{d}e_t \ = \ \mu_e(t) \, \mathrm{d}t + \sigma_e(t) \, \mathrm{d}B_t$.

H3 There exist $k > 0$ such that $k \leq e$ a.s., and M such that $\int_0^T \|\sigma_e(s)\|^2 \, \mathrm{d}s < M$ a.s..

Recall that k_i is a lower bound on the equilibrium consumption of agent i (see Proposition 7.2.6).

U4 There exist i and $A_i > 0$ such that

$$\frac{-u_{i,cc}(t, c)}{u_{i,c}(t, c)} \ \leq \ A_i , \quad t \in [0, T] , \quad c \in [k_i, \infty[\, .$$

Let $(c_i, i = 1, \ldots, m; \; p)$ be a contingent Arrow–Debreu equilibrium. From Theorem 7.1.10, it can be implemented as a Radner equilibrium if p is a strictly positive Itô process $dp_t = \mu_p \, dt + \sigma_p \, dB_t$ satisfying $E_P \left[\exp \; \frac{1}{2} \int_0^T \frac{\|\sigma_p(t)\|^2}{p(t)^2} \, dt \right] < \infty$. We now show that these conditions are in fact satisfied.

Proposition 7.2.10. *Under the assumptions* $(\mathbf{H1 - H3}, \mathbf{U1 - U4})$,

(i) p_t *is a strictly positive Itô process belonging to* L_+^∞,

(ii) $E_P \left[\exp \; \frac{1}{2} \int_0^T \frac{\|\sigma_p(t)\|^2}{p_t^2} \, dt \right] < \infty$.

Proof. From Theorem 7.2.8, there exists $\alpha_0 \in \Delta^{m-1}$ such that $p_t = \frac{\partial u}{\partial c}(t, e(t), \alpha_0)$. According to Proposition 7.2.1, $\frac{\partial u}{\partial c}(\cdot, \cdot, \alpha_0)$ is of class $C^{1,2}$. It follows from Itô's formula that p_t is an Itô process of the form:

$$dp_t = \mu_p(t) \, dt + u_{cc}[t, e(t), \alpha_0] \, \sigma_e(t) \, dB_t . \tag{7.11}$$

As we showed in Remark 7.2.2,

$$-\frac{u_{cc}(t, e(t), \alpha_0)}{u_c(t, e(t), \alpha_0)} \leq \min_i - \frac{u_{i,cc}(t, C_i(t, e(t), \alpha_0))}{u_{i,c}(t, C_i(t, e(t), \alpha_0))} \leq A_i ,$$

hence $E_P \left[\exp \; \frac{1}{2} \int_0^T \frac{\|\sigma_p\|^2}{p^2} \, dt \right] < \infty$. $\square$

Notice that we recover the results of the Cox–Ingersoll–Ross [70] interest rate model with one agent. Indeed, as we showed in the proof of Theorem 7.1.10, we have

$$r_t = -\frac{\mu_p(t)}{p_t} . \tag{7.12}$$

By substituting in for $\mu_p(t)$, and writing u_{ccc} for the third derivative of u, we obtain:

$$\frac{\mu_p(t)}{p(t)} = -\frac{1}{p_t} \left[u_{ct}[t, e(t), \alpha_0] + \mu_e(t) u_{cc}[t, e(t), \alpha_0] \right.$$
$$\left. + \frac{1}{2} \|\sigma_e(t)\|^2 \, u_{ccc}[t, e(t), \alpha_0] \right] . \tag{7.13}$$

There is no reason for the interest rate $-\frac{\mu_p}{p}$ to be positive. Also, we notice that the framework for this section is much more general than the one considered by Cox–Ingersoll–Ross. Indeed, here we consider a model with several agents, and furthermore, we do not assume the processes for asset prices to be Markov.

As a summary, we give the following theorem:

Theorem 7.2.11. *Under the assumptions* $(\mathbf{H1 - H3, U1 - U4})$*, there exists a Radner equilibrium.*

For any Radner equilibrium $\{(\tilde{\theta}_i, c_i), \ (i = 1, \ldots, m) \ ; \ \tilde{S}\}$*, there exists* $\alpha_0 \gg 0$ *such that* $c_i(t) = C_i[t, e(t), \alpha_0]$*,* $i = 1, \ldots, m$*. Moreover,* $\{ \ c_i, \ i = 1, \ldots, m \ \}$ *is Pareto-optimal, and there is a representative agent characterized by an additively separable utility function and by endowments that are equal to the aggregate endowment.*

The interest rate is equal to the inverse of the expectation of the relative variation of its instantaneous marginal utility. The probability measure Q*, which is is equivalent to* P*, has density* $\xi_t = \frac{1}{R_t} \frac{\partial u}{\partial c}(t, e(t), \alpha_0)$ *on* $\mathcal{F}_t$*.*

We conclude these last two sections by giving the formula for asset prices:

$$ S_t \ = \ \frac{1}{R(t)} E_Q \left[\int_t^T \mathrm{d}D^{\mathrm{d}}(s) \,|\, \mathcal{F}_t \right] \ + \ \frac{1}{R(t)} E_Q \left[R(T) S(T) \,|\, \mathcal{F}_t \right] \ . $$

As stated in Remark 7.2.9, if we fix k and if we fix the utility functions generically with respect to endowments, then there are a finite number of equilibrium utility weights, and hence of "equilibrium" measures Q. Prices are not fully determined by the equality between supply and demand. Their final value is arbitrary. As we have seen over the course of this book, and as will be shown in the next section: what matters in the calculations are the risk-neutral measure and the interest rate.

7.3 Applications

We are going to show that we can obtain a continuous version with a finite horizon, of the Lucas [259] one-agent model, and of the Breeden [43] CCAPM (Consumption-based Capital Asset Pricing Model).

In the following, we assume that we have a Radner equilibrium $(\tilde{\theta}_i, c_i, (i = 1, \ldots, m); \tilde{S})$ satisfying the conditions of Theorem 7.2.11. We consider the probability measure Q that is equivalent to P and has density $\xi_t = \frac{1}{R_t} \frac{\partial u}{\partial c}(t, e(t), \alpha_0)$. We have that $R_t \xi_t$ is uniformly bounded.

7.3.1 Arbitrage Price of Real Secondary Assets. Lucas' Formula

We use martingale theory to give an arbitrage price for secondary assets.

Definition 7.3.1. *Let* Y *be a dividend process of the form*

$$ \mathrm{d}Y_t \ = \ \mu_Y(t) \, \mathrm{d}t + \sigma_Y(t) \, \mathrm{d}B_t \ . \tag{7.14} $$

We say that Y *is "attainable" if* $\int_0^t R(s) \, \mathrm{d}Y(s)$ *is* Q*-integrable, or equivalently, if* $\int_0^t p(s) \, \mathrm{d}Y(s)$ *is* P*-integrable. A trading strategy* $\tilde{\theta} \in \mathcal{H}(G^{\mathrm{d}})$ *finances[2]* Y *at*

an initial cost of V_0 if

(i) $\tilde{\theta}_0 \cdot \tilde{S}_0 = V_0$,

(ii) $R(t)(\tilde{\theta}_t \cdot \tilde{S}_t) = \tilde{\theta}_0 \cdot \tilde{S}_0 + \int_0^t \theta(s) \, \mathrm{d}G^{\mathrm{d}}(s) - \int_0^t R(s) \mathrm{d}Y(s)$ for $t \in [0, T]$,

(iii) $\tilde{\theta}_T \cdot \tilde{S}_T = 0$.

The value $\tilde{\theta}_t \cdot \tilde{S}_t =: V_t$ of the trading strategy at time t, is the "arbitrage price" of Y at time t.

Proposition 7.3.2. *Any attainable dividend process Y is financed at an initial cost*

$$V_0 = E_P\left[\int_0^T p(s) \, \mathrm{d}Y(s)\right], \tag{7.15}$$

and its arbitrage price is

$$V_t = \frac{1}{p_t} E_P\left[\int_t^T p(s) \, \mathrm{d}Y(s) \, | \, \mathcal{F}_t\right]. \tag{7.16}$$

Proof. As in the proof of Proposition 7.1.5, we introduce the martingale

$$M_t = E_Q\left[\int_0^T R(s)\mathrm{d}Y(s) \, | \, \mathcal{F}_t\right].$$

It follows from the predictable representation theorem that there exists $\theta = (\theta^1, \ldots, \theta^d)$ with $\int_0^t \|\theta(t)\sigma_G^{\mathrm{d}}(t)\|^2 \, \mathrm{d}t < \infty$ P a.s., such that,

$$M_t = V_0 + \int_0^t \theta(s) \, \mathrm{d}G^{\mathrm{d}}(s).$$

Hence $V_0 = E_Q(M_T) = E_P\left[\int_0^T p(s)\mathrm{d}Y(s)\right]$ is uniquely determined.

Let us define θ^0 by:

$$R(t)S_t^0\theta_t^0 = M_t - \int_0^t R(s)\mathrm{d}Y(s) - R(t)\sum_{i=1}^d \theta_t^i S_t^i.$$

By definition, $\tilde{\theta} = (\theta^0, \theta) \in \mathcal{H}(G^{\mathrm{d}})$ and $\tilde{\theta}$ finances Y at an initial cost V_0 defined by (7.15), and we have

$$V_t = \tilde{\theta}_t \cdot \tilde{S}_t = \frac{1}{R(t)} E_Q\left[\int_t^T R(s)\mathrm{d}Y(s) \, | \, \mathcal{F}_t\right] = \frac{1}{p(t)} E_P\left[\int_t^T p(s)\mathrm{d}Y(s) \, | \, \mathcal{F}_t\right].$$

$\square$

---

[2] See also Sect. 3.4.4.

Remark 7.3.3. Here we assume that the final value of the asset is zero.

Remark 7.3.4. Suppose that the asset also yields a final consumption δ, which is assumed to be a random variable such that $p(T)\delta$ is P-integrable. Then expression (7.16) can be generalized to

$$V_t = \frac{1}{p_t} E_P \left[\int_t^T p(s)\, dY(s) + p(T)\delta \,|\, \mathcal{F}_t \right], \quad t \in [0, T]. \tag{7.17}$$

To reproduce Lucas' valuation formula in continuous time and with a finite horizon, it is enough to assume that there is a single agent, and that $u(t, c) = e^{-\beta t} \varphi(c)$. We then obtain:

$$V_t = \frac{1}{\varphi'(e_t)} E_P \left[\int_t^T e^{-\beta(s-t)} \varphi'[e(s)]\, dY_s \,|\, \mathcal{F}_t \right].$$

Remark 7.3.5. Formula (7.16) holds more generally in the case where Y is a semi-martingale such that $\int_0^t p(s)\, dY(s)$ is well-defined. In particular, the process Y can include jumps. Indeed, suppose we want to calculate the price of a zero coupon bond maturing at time τ. This means that $Y_t = 0$ for $t < \tau$ and $Y_t = 1$ for $t \geq \tau$. The process Y has a jump at time τ. We deduce from (7.16) that the price of the zero coupon bond is zero after time τ, and for $t < \tau$, equals

$$V_t = \frac{1}{p_t} E_P \left(p(\tau) \,|\, \mathcal{F}_t \right).$$

7.3.2 CCAPM (Consumption-based Capital Asset Pricing Model)

We recover Breeden's formula.

Proposition 7.3.6. *For Y an attainable Itô process,*

(i) The process $Z_t = \int_0^t p(s)\, dY(s) + p_t V_t$ is a P-martingale.
(ii) The process V_t is an Itô process.

Proof. Notice that

$$Z_t = \int_0^t p(s)\, dY(s) + p_t V_t = \int_0^t p(s)\, dY(s) + E_P \left[\int_t^T p(s)\, dY(s) \,|\, \mathcal{F}_t \right]$$
$$= E_P \left[\int_0^T p(s)\, dY(s) \,|\, \mathcal{F}_t \right].$$

Hence Z_t is a continuous P-martingale and therefore an Itô process (apply the predictable representation theorem).

Consequently, $p_t V_t = Z_t - \int_0^t p(s)\,dY(s)$ is an Itô process. As, according to Proposition 7.2.6, p_t is a strictly positive Itô process, it follows from Itô's formula that $V_t = \frac{p_t V_t}{p_t}$ is an Itô process that can be written $\mu_V(t)dt + \sigma_V(t)\,dB_t$.

$\square$

Henceforth, we will be looking at the rate of real excess return on stocks.

According to Proposition 7.3.6, for a given asset Y, the process $Z_t = \int_0^t p(s)\,dY(s) + p_t V_t$ is a P-martingale.

By applying Itô's lemma, we get,

$$dZ_t = \Big[p_t \mu_Y(t) + p_t \mu_V(t) + V_t \mu_p(t) + u_{cc}[t, c(t), \alpha_0]$$
$$\sigma_e(t) \cdot \sigma_V(t) \Big]\, dt + \sigma_Z(t)\,dB_t \,, \quad (7.18)$$

where $\sigma_e(t) \cdot \sigma_V(t)$ denotes the matrix product $\sigma_e(t)\sigma_V(t)^T$.

As Z is a martingale, it has zero drift. Therefore, we have a.e.,

$$p_t[\mu_Y(t) + \mu_V(t)] + V_t \mu_p(t) + u_{cc}[t, e(t), \alpha_0]\sigma_e(t) \cdot \sigma_V(t) = 0 \,. \quad (7.19)$$

If we assume that $V_t \neq 0$, and divide through by $p_t V_t$, we obtain, from formula (7.13),

$$\frac{\mu_V(t) + \mu_Y(t)}{V_t} - r_t = -\frac{u_{cc}[t, e(t), \alpha_0]\,\sigma_V(t) \cdot \sigma_e(t)}{\frac{\partial u}{\partial c}[t, e(t), \alpha_0]\,V_t} \,. \quad (7.20)$$

We can associate with each asset Y, the "real rate of return" Γ, which solves the stochastic differential equation

$$d\Gamma_t = \left(\frac{\mu_V(t) + \mu_Y(t)}{V_t} \right) dt + \frac{\sigma_V(t)}{V_t}\,dB_t \,. \quad (7.21)$$

We interpret the term $\frac{\mu_V(t)+\mu_Y(t)}{V_t} = \mu_\Gamma(t)$ as the expectation of the instantaneous real rate of return. First of all, it follows from (7.20) that the interest rate r_t is the instantaneous real rate of return of an asset whose price has zero volatility $\sigma_V(t)$ (a riskless asset).

Let us set

$$\sigma_\Gamma(t) = \frac{\sigma_V(t)}{V_t} \,. \quad (7.22)$$

We can rewrite (7.20) in the form:

$$\overline{\Gamma}_t - r_t = -\frac{u_{cc}[t, e(t), \alpha_0]}{\frac{\partial u}{\partial c}[t, e(t), \alpha_0]}\sigma_\Gamma(t) \cdot \sigma_e(t) \,. \quad (7.23)$$

The term $-\dfrac{u_{cc}[t, e(t), \alpha_0]}{\frac{\partial u}{\partial c}[t, e(t), \alpha_0]}$ can be interpreted as the coefficient of risk aversion of the representative agent, and the term $\sigma_\Gamma(t) \cdot \sigma_e(t)$ as the instantaneous covariance between the real rate of return and the aggregate consumption.

In order to finally obtain a beta formula, we construct the real price process V^*, whose diffusion coefficient can be written $\sigma_{V^*}(t) = k(t)\sigma_e(t)$ where k is a predictable real-valued process. There are different ways of doing this. For example, if we assume that we have an asset whose final dividend δ equals $\frac{1}{p(T)} \int_0^T \sigma_e(s) \, dB(s)$, then according to (7.17) its real arbitrage price is then

$$V_t^* = \frac{1}{p_t} E_P \left[\int_0^T \sigma_e(s) \, dB(s) \mid \mathcal{F}_t \right] = \frac{1}{p_t} \int_0^t \sigma_e(s) \, dB(s) .$$

Using Itô's formula, we then obtain

$$dV_t^* = \mu_V^*(t) \, dt + \sigma_e(t) \left[\frac{1}{p_t} - \frac{\int_0^t \sigma_e(s) \, dB_s}{p_t^2} u_{cc}[t, e(t), \alpha_0] \right] dB_t$$

$$= \mu_V^*(t) \, dt + k(t) \, \sigma_e(t) \, dB_t ,$$

with

$$k_t = \frac{1}{p_t} - \frac{1}{p_t^2} u_{cc}[t, e(t), \alpha_0] \int_0^t \sigma_e(s) \, dB_s .$$

We now set $\sigma_{\Gamma^*}(t) = \frac{\sigma_{V^*}(t)}{V^*(t)}$.

By rewriting (7.23) for this asset, we obtain:

$$\overline{\Gamma}_t^* - r_t = -\frac{u_{cc}[t, e(t), \alpha_0]}{u_c[t, e(t), \alpha_0]} \sigma_{\Gamma^*}(t) \cdot \sigma_e(t) . \tag{7.24}$$

Let us now define the beta of a real asset with respect to aggregate consumption by:

$$\beta_\Gamma(t) = \frac{\sigma_\Gamma(t) \cdot \sigma_{\Gamma^*}(t)}{\sigma_{\Gamma^*}(t) \cdot \sigma_{\Gamma^*}(t)} , \tag{7.25}$$

(assuming that $\sigma_{\Gamma^*}(t) \neq 0$). We then obtain:

$$(\overline{\Gamma}_t - r_t) = \beta_\Gamma(t) (\overline{\Gamma}_t^* - r_t) . \tag{7.26}$$

Hence the excess rate of return on stocks is proportional to the "consumption betas".

Notes

As we said in the introduction, the first continuous time equilibrium models had one agent, and the methods used to analyze them were essentially the methods of dynamic programming. In particular, this is the case for Merton [273], (1973), which introduced the CAPM property in continuous time, Breeden [43], (1979) and finally Cox, Ingersoll and Ross [70], (1985), which deals with determining the interest rate at equilibrium.

The problem of the existence of an equilibrium with financial markets in continuous time is linked to two currents in the literature.

The first one goes back to the end of the sixties (see for example the articles Bewley [30, 31], (1969, 1972)), and deals with the existence of an Arrow–Debreu equilibrium in infinite dimension. Very general spaces were considered and a wide range of analytical methods were used. The literature in this area is plentiful. The interested reader can consult either the article Mas-Colell and Zame [269], (1991), which is based mainly on the Negishi method, or the book Aliprantis, Brown and Burkinshaw [5], (1989). The case of L^p spaces and additively separable utility functions has been studied by Araujo and Monteiro [10, 11], (1989, 1992) and Dana [81, 82], (1993). The issue of uniqueness (or local uniqueness) of an Arrow–Debreu equilibrium is discussed in Dana [84], (2001), and the references therein.

The second current deals with the existence of a Radner equilibrium. Traditionally, and in applications, the utility functions are often assumed to be additively separable. The idea of using martingale theory to implement an Arrow–Debreu equilibrium as a Radner equilibrium was introduced by Duffie and Huang [119], (1985), in a model where agents consumed at time 0 and at the final date. The idea was later generalized by Duffie [110], (1986) to a model allowing for a consumption process. Huang [196], (1987) uses the idea that any Arrow–Debreu equilibrium is Pareto-optimal, to give sufficient conditions for asset prices to be functions of a state variable given by a diffusion process. Karatzas, Lehoczky and Shreve [232], (1990) were the first to take the consumption good as the numéraire. The real prices of the reference risky assets are then Itô processes. Their paper shows that when agents have coefficients of relative risk aversion that are smaller than one, there exists a unique equilibrium measure, under which discounted prices are martingales. Finally, Dana and Pontier [85], (1992) show that very general informational structures can be considered. The case of two agents is considered in Dumas [127], (1989) and Wang [363], (1996).

Duffie and Epstein [116], (1992) extend the results of Sect. 7.3 to "recursive" utility functions in the setting of a representative agent model. The existence of equilibrium allocations with recursive utilities is proved in Duffie et al [118], (1994). The existence of a representative agent with recursive utilities is discussed in Dumas et al [129], (2000). Epstein and Miao [151], (2002) consider a two-agent model, where agents have different sets of multiple priors to solve the equity home bias and the consumption home bias puzzles.

Basak and Cuoco [25], (1998) analyse a model in which agents have heterogeneous investment opportunities. Models containing heteregeneous beliefs or information are developed by Detemple and Murthy [101], (1994), Basak [24], (2000), De Marzo and Skiadas [99, 100], (1998, 1999), Jouini and Napp [219] (2006), Riedel [309], (2002), and Zapatero [377], (1998), Serrat [332] (2001).

Finally, Cuoco [73], (1997), extends Breeden's CCAPM to the case of portfolio constraints.

8

Incomplete Markets

The valuation and hedging of options in incomplete and in imperfect markets is an issue that concerns both academics and practitioners. This brief chapter is an introduction to the topic. A more complete study is beyond the scope of this book. As in previous chapters, we work with continuous processes built on a space $(\Omega, \mathcal{F}, P)$.

8.1 Incomplete Markets

The framework is that of a market with $d + 1$ assets, including one riskless asset, where asset prices are modeled by:

$$
\begin{cases}
\mathrm{d}S_0(t) = r(t)S_0(t)\mathrm{d}t, \qquad S_0(0) = 1 \\[2mm]
\mathrm{d}S_i(t) = S_i(t)[b_i(t)\,\mathrm{d}t + \sum_{j=1}^{n} \sigma_{i,j}(t)\,\mathrm{d}W_j(t)]\,, \qquad i = 1, \ldots, d \qquad (8.1) \\[2mm]
\qquad\quad = S_i(t)[b_i(t)\,\mathrm{d}t + \sigma_i(t)\,\mathrm{d}W_j(t)]\,,
\end{cases}
$$

for $\sigma_i(t)$, a row vector $(\sigma_{i,j},\ j = 1, \ldots, n)$. Here, the dimension of the Brownian motion $W = (W_1, W_2, \ldots, W_n)$ is no longer equal to the number of risky assets. If $d > n$, there is an excess of assets, and the possibility of arbitrage. In what follows, we assume that $n > d$ and that the matrix $\sigma\sigma^T$ is invertible. Recall that a market is complete if any positive, square-integrable r.v. $\zeta \in \mathcal{F}_T$ can be written as the final value of a self-financing strategy. In a complete no-arbitrage market, the equivalent martingale measure is unique.

8.1.1 The Case of Constant Coefficients

We must be careful not to be too hasty in describing a market as incomplete. For example, consider a market consisting of one riskless asset with constant rate of return r, and one risky asset whose price follows the dynamics

$$dS_t = S_t \left(\mu dt + \sigma_1 dB_t^{(1)} + \sigma_2 dB_t^{(t)} \right) \tag{8.2}$$

where σ_i are non-zero constants and where $B^{(i)}$ are independent Brownian motions. With this formulation, there exist an infinity of risk-neutral probability measures R, given by

$$\frac{dR}{dP} = \exp \left(\int_0^t \left[\theta_1(s)dB_s^{(1)} - \frac{1}{2} \left(\theta_1(s) \right)^2 ds + \theta_2(s)dB_s^{(2)} - \frac{1}{2} \left(\theta_2(s) \right)^2 ds \right] \right)$$

with $\sigma_1\theta_1 + \sigma_2\theta_2 + \mu = r$. There exist non-replicable contingent claims: for example, the random variable $B_T^{(2)}$ cannot be written as the final value of a self-financing portfolio. However, this market is complete if we restrict ourselves to contingent claims that are measurable with respect to the price filtration.

To check this is the case, introduce the process $B^{(3)}$ defined by

$$B_t^{(3)} = \frac{1}{\sqrt{\sigma_1^2 + \sigma_2^2}} \left[\sigma_1 B_t^{(1)} + \sigma_2 B_t^{(2)} \right] .$$

This process is a Brownian motion (as it is a martingale and $([B_t^{(3)}]^2 - t, \ t \geq 0)$ is also a martingale). Using this process, and setting $\sigma_3 = \sqrt{\sigma_1^2 + \sigma_2^2}$, the dynamics of the risky asset's price become

$$dS_t = S_t \left(\mu dt + \sigma_3 dB_t^{(3)} \right) . \tag{8.3}$$

We need to specify which are the filtrations being used. The initial filtration was generated by the two Brownian motions $B^{(i)}$. The price filtration is $\tilde{\mathcal{F}}_t = \sigma \left(S_s, \ s \leq t \right) = \sigma \left(B_s^{(3)}, \ s \leq t \right)$. Indeed, from (8.3), we have $\sigma \left(S_s, \ s \leq t \right) \subset \sigma \left(B_s^{(3)}, \ s \leq t \right)$. Moreover, equation (8.3) implies that $dB_t^{(3)} = (dS_t - S_t\mu dt)/S_t\sigma_3$, which shows that the reverse inclusion also holds. Under this new form, the price dynamics are "Black–Scholes", and there exists a unique measure Q, such that $(S_t e^{-rt}, \ t \geq 0)$ is a $\left(\tilde{\mathcal{F}}_t \right)$-$Q$-martingale, satisfying $d \left(S_t e^{-rt} \right) = S_t e^{-rt}\sigma_3 dW_t$, with the process $\left(W_t = B_t^{(3)} - \frac{\mu - r}{\sigma_3}t, \ t \geq 0 \right)$ being a $\left(\tilde{\mathcal{F}}_t \right)$-$Q$-Brownian motion.

Any random variable ζ, which is square integrable and measurable with respect to $\tilde{\mathcal{F}}_T$, is also replicable; this is the statement of the representation theorem under Q. Indeed, the $\left(\tilde{\mathcal{F}}_t \right)$-$Q$-martingale $V_t e^{-rt} = E_Q \left(e^{-rT}\zeta | \mathcal{F}_t \right)$ can be represented as a stochastic integral with respect to W:

$$E_Q \left(e^{-rT}\zeta | \mathcal{F}_t \right) = b + \int_0^t \psi_s dW(s) = b + \int_0^t \pi_s d \left(S_s e^{-rs} \right)$$

where $\pi_s = \psi_s / \left(S_s e^{-rs}\sigma_3 \right)$. From this, it can be deduced that V_t is the value at time t of a self-financing portfolio which replicates ζ, i.e., that $V_t = \alpha_t S_t^0 + \pi_t S_t$

with $\alpha_t = e^{-rt}(V_t - \pi_t S_t)$. So we have checked that a contingent claim that is measurable with respect to the information contained in market prices, is replicable. This no longer holds for a random variable U that is measurable with respect to $\mathcal{F}_T = \sigma\left(B_t^{(1)},\, B_t^{(2)},\, t \leq T\right)$, which is a larger σ-field than $\tilde{\mathcal{F}}_T$. In this case, the representation theorem yields

$$E\left(e^{-rT}U\big|\mathcal{F}_t\right) = u + \int_0^t \psi_1(s)\mathrm{d}B_s^{(1)} + \int_0^t \psi_2(s)\mathrm{d}B_s^{(2)}\ ,$$

which can no longer be written as $u + \int_0^t \pi_s \mathrm{d}\left[e^{-rs}S_s\right]$. However, from a financial point of view, the valuation of such claims holds little interest. The model with two Brownian motions is not satisfactory. It requires too much irrelevant information. All this can be generalized to the case of deterministic coefficients, but not to the case of stochastic coefficients.

8.1.2 No-Arbitrage Markets

Under some regularity assumptions, which we specify later, without however aiming for the most refined possible set of assumptions, the market described in (8.1) does not present any arbitrage opportunities. To prove this is the case, we show that there exists an equivalent martingale measure. Let

$$\theta_t \stackrel{\mathrm{def}}{=} \sigma_t^T[\sigma_t \sigma_t^T]^{-1}[b_1 - r_t\mathbf{1}]\ .$$

Assume that σ and θ are bounded. We use the following notation: $R_t = \exp\left(-\int_0^t r_s \mathrm{d}s\right)$,

$$L_t^0 \stackrel{\mathrm{def}}{=} \exp\left(-\int_0^t \theta_s^T \mathrm{d}W_s - \frac{1}{2}\int_0^t \|\theta_s\|^2 \mathrm{d}s\right)\ ,$$

and $H_t^0 \stackrel{\mathrm{def}}{=} R_t L_t^0$, for the state price process satisfying

$$\mathrm{d}H_t^0 = -H_t^0(r_t\mathrm{d}t + \theta_t^T \mathrm{d}W_t)\ .$$

The process L^0 is a strictly positive P-$(\mathcal{F}_t)$ martingale. Let Q be the probability measure defined on $\mathcal{F}_t$ by $\mathrm{d}Q = L_t^0 \mathrm{d}P$.

Proposition 8.1.1. *Under the measure Q, discounted prices are martingales.*

Proof. A very simple method involves checking that the processes $S_i H^0$ are martingales under P. This follows from Itô's lemma, with the formula for integration by parts leading to:

$$\mathrm{d}\left(S_i H^0\right)(t) = H_t^0 \mathrm{d}S_i(t) + S_i(t)\mathrm{d}H^0(t) - S_i(t)H_t^0 \mathrm{d}\langle\sigma_i W,\ \theta^T W\rangle_t$$

$$= H_t^0 S_i \left[b_i - r - \sigma_i\theta\right]\mathrm{d}t + M_t$$

where M is a martingale. By definition of θ, the drift term is zero.   □

However, it is not possible to replicate every $\mathcal{F}_T$-measurable random variable. Let us explain why, in the case $r = 0$. According to the representation theorem, there exists a family $(\phi_k,\ k \leq n)$ of processes such that $\zeta = x + \sum_{k=0}^{n} \phi_k(t)\, \mathrm{d}W_k(t)$. For ζ to be replicable, we require the existence of $(\pi_i,\ i \leq d)$ such that

$$\sum_{k=0}^{n} \int_0^T \phi_k(t)\mathrm{d}W_k(t) = \sum_{i=1}^{d} \int_0^T \pi_i(t)\mathrm{d}S_i(t)$$

$$= \sum_{k=1}^{n} \int_0^T \sum_{i=1}^{d} [\pi_i(t)\sigma_{i,k}(t)]\ \mathrm{d}W_k(t)\ .$$

By identification of $\mathrm{d}W_k$ terms, this leads to a system of n equations with d unknowns.

8.1.3 The Price Range

In an incomplete market, there exist non-replicable assets. That is, square-integrable $\mathcal{F}_T$-measurable random variables ζ, for which there exists no self-financing portfolio with final value ζ. However, replicable variables exist, as the assets themselves and all random variables of the form $\zeta = x + \int_0^T \pi_s \mathrm{d}S_s^{\mathsf{d}}$ (where S^{d} denotes the discounted prices). As often mentioned previously, if a contingent claim ζ is replicable, its price is $E_Q\left(\mathrm{e}^{-rT}\zeta\right)$. If there exist several martingale measure exists, then the values of $E_Q\left(\mathrm{e}^{-rT}\zeta\right)$ across the different measures Q are all equal. When the claim is not replicable, we refer to any value of $E_Q\left(\mathrm{e}^{-rT}\zeta\right)$ as a viable price. Such a price precludes the possibility of arbitrage, as will be seen later. When the final payoff is not replicable, we can approach the problem in two different ways, just as we did in the first chapter:

- by finding the smallest value x, such that there exists a self-financing portfolio with initial value x and final value greater than ζ,
- by considering the set of "viable" prices, that is the set of values of $E_Q\left(\mathrm{e}^{-rT}\zeta\right)$ given when Q describes the set of risk-neutral probability measures.

We should emphasize the fact that the set of risk-neutral probability measures is convex; if this set contains more than one probability measure, then it contains an infinity of probability measures.

If a claim is replicable, then all the values of $E_Q\left(\mathrm{e}^{-rT}\zeta\right)$ are equal.

Firstly, we look at the set of equivalent martingale measures, and secondly, we study superhedging strategies, before making the link between the two. It

is not possible to give all the proofs involved, as these require mathematical developments that are beyond the scope of this book. Instead, we refer the reader to the papers El Karoui and Quenez [144, 145], Cvitanić and Karatzas [75], Kramkov [241], and Cvitanić, Pham and Touzi [78].

Equivalent Martingale Measures

To determine all the equivalent martingale measures, we look for their Radon–Nykodym densities.

Let $K(\sigma)$ be the kernel of σ. For $\nu \in K(\sigma)$, we notice that $\langle \nu, \theta \rangle = 0$, and define, for a bounded ν, the exponential martingale

$$L_t^\nu \overset{\text{def}}{=} \exp\left(-\int_0^t \left(\theta_s^T + \nu_s^T \right) dW_s - \frac{1}{2} \int_0^t \left(\|\theta_s\|^2 + \|\nu_s\|^2 \right) ds \right),$$

and the process $H_t^\nu = L_t^\nu R_t$ satisfying

$$
\begin{aligned}
dH_t^\nu &= -H_t^\nu \left(r_t\, dt + \left(\theta_t^T + \nu_t^T \right) dW_t \right) \\
H_0^\nu &= 1 .
\end{aligned}
\tag{8.4}
$$

Lemma 8.1.2. *For all $\nu \in K(\sigma)$, the processes $(H^\nu(t)S_i(t),\ t \geq 0)$ are martingales under the measure P.*

Proof. The proof is straightforward. As before, it is enough to apply Itô's formula. □

The result can be stated by saying that the probability measures Q^ν with densities L^ν with respect to P are equivalent martingale measures. Next, we need to check that in this manner, we have obtained all the equivalent martingale measures.

Lemma 8.1.3. *The set of equivalent martingale measures is the set $\mathcal{Q}$ defined by*

$$\mathcal{Q} = \left\{ Q^\nu \,\big|\, dQ^\nu|_{\mathcal{F}_t} = L_t^\nu dP|_{\mathcal{F}_t},\ \nu \in K(\sigma) \right\}$$

with

$$dL^\nu(t) = -L^\nu(t) \left(\theta^T(t) + \nu^T(t) \right) dW_t, \qquad L^\nu(0) = 1 .$$

Proof. If Q is an equivalent martingale measure, the density $(L_t,\ t \geq 0)$ is a strictly positive $\mathcal{F}_t$ martingale, which, thanks to the predictable representation theorem, admits the representation

$$dL_t = L_t \varphi_t^T\, dW(t) .$$

The process $H^\nu S$ is a P-martingale if and only if $\sigma(t)\varphi(t) + (b(t) - r(t)\mathbf{1}) = 0$. Then the various φ can be written as $\theta + \nu$, for $\nu \in K(\sigma)$. □

The range of viable prices at time 0 is given by:

$$\Big] \inf_{\nu} E\left(H^{\nu}(T)\zeta\right), \ \sup_{\nu} E\left(H^{\nu}(T)\zeta\right) \Big[\ .$$

We will write $\bar{p}(\zeta) = \sup_{\nu} E\left(H^{\nu}(T)\zeta\right)$ and $\underline{p}(\zeta) = \inf_{\nu} E\left(H^{\nu}(T)\zeta\right)$.

In general, this price range is very wide. In the case of European option hedging, the upper bound is often equal to the trivial bound (i.e., equal to the value of the underlying). The reader can refer to the work of Eberlein–Jacod [133] and Bellamy–Jeanblanc [26]. It is easy to obtain price ranges of this form for any time t.

If the contingent claim ζ is sold at a price $p(\zeta) \in \ ]\underline{p}(\zeta), \ \bar{p}(\zeta)[$, then an arbitrage cannot be constructed. Indeed, if we sell ζ at price $p(\zeta)$, then we can invest the proceeds in the market by creating a portfolio θ. This portfolio would be an arbitrage opportunity if its final value V_T were greater than ζ. However, under any equivalent martingale measure, the portfolio's discounted value is a martingale. It follows that

$$p(\zeta) = V_0 = E^{\nu}\left(V_T R(T)\right) = E\left(H^{\nu}(T)V_T\right) > E\left(H^{\nu}(T)\zeta\right) \ .$$

We can give an analogous proof to show that a no-arbitrage price must be greater than $\inf_{\nu}\left(H^{\nu}(T)\zeta\right)$. The value $\bar{p}(\zeta) = \sup_{\nu}\left(H^{\nu}(T)\zeta\right)$ is referred to as the selling price of ζ. This choice of terminology will be justified later. El Karoui–Quenez [144, 145] established the following result.

Proposition 8.1.4. *The three following properties are equivalent:*

1. *There exists $\lambda \in K(\sigma)$ such that $E\left(H^{\nu}(T)\zeta\right) = x$.*
2. *The price range is reduced to a single element:*

$$E\left(H^{\nu}(T)\zeta\right) = x, \ \forall \nu \in K(\sigma) \ .$$

3. *There exists an admissible portfolio π such that $X^{x,\pi}(T) = \zeta$.*

8.1.4 Superhedging

We now consider portfolios that superhedge a contingent claim ζ. The seller is willing to sell the contingent claim at a price of x if he can invest x in a portfolio whose final value is greater than the value of the claim, i.e., $V_T^{x,\pi} \geq \zeta$. The selling price is then the smallest amount x that can be used to hedge the claim.

Using the notation of the previous section, it can be shown that $\bar{p}(\zeta) = \inf \left\{x \ : \ \exists \pi, \ V_T^{x,\pi} \geq \zeta\right\}$. Similarly, reasoning from the buyer's point of view, we define the purchase price, and show that $\underline{p}(\zeta) = \sup \left\{x \ : \ \exists \pi, \ V_T^{-x,\pi} \geq -\zeta\right\}$. The essential tool here is the smallest supermartingale (for any given martingale measure equivalent to P) equal to ζ at time T. This supermartingale is equal to $J_0^* = \sup E_{Q^{\nu}}\left(R(T)\zeta\right)$ at time 0, and to $J_t^* = \sup E_{Q^{\nu}}\left(R(T)\zeta|\mathcal{F}_t\right)$

at time t, with the last equality requiring a particular definition of "sup" to avoid difficulties due to null sets (and this leads to the notion of essential sup).

We refer the reader to the papers El Karoui and Quenez [144], and Kramkov [241], as well as to the lecture notes Touzi [355] and Pham [299], for proofs of these results.

8.1.5 The Minimal Probability Measure

One method, which is frequently used to price financial products, involves choosing one specific probability measure, amongst all the different risk-neutral measures available. Different criteria exist for this choice in the literature, and the links between the different probability measures are themselves an area of study.

The Minimal Measure

This probability measure is linked to minimizing quadratic cost. It was introduced by Föllmer and Sondermann [163], and was exhaustively studied by Schweizer [330]. The chosen criterion is the minimization of $E\left(\zeta - X_T^{\pi,x}\right)^2$. The associated dual to the problem is $\min_\nu E\left[\left(\mathrm{e}^{-rT} L_T^\nu\right)^2\right]$, where the parameter ν varies in such a way that the set of risk-neutral probability measures is described.

The Optimal Variance Martingale Measure

The idea here is to choose Q close to the historic probability measure, selecting the equivalent martingale measure that minimizes $E_P\left[\left(R(T)\frac{\mathrm{d}Q}{\mathrm{d}P}\right)^2\right]$. Further details are to be found in Schweizer [331], and in Delbaen and Schachermayer [96, 97].

The Minimal Entropy Martingale Measure

Miyahara [280] on the other hand chooses the equivalent martingale measure that is closest to the historic measure, in the sense of a distance that is linked to "entropy", and which minimizes $E_P\left[\frac{\mathrm{d}Q}{\mathrm{d}P}\ln\frac{\mathrm{d}Q}{\mathrm{d}P}\right]$. Fritelli [166] has devoted numerous articles to this topic, and the dual approach has been studied by Delbaen et al [95]

Risk

This new approach consists in accepting that the hedge will not be perfect, and that only 95% of the risk will be hedged. So we look for a portfolio, with a minimal initial value x, such that $P\left(X_T^{x,\pi} \geq \xi\right) = 0.95$. (Föllmer and Leukert [161]; Cvitanić and Karatzas, [76]).

8.1.6 Valuation Using Utility Functions

Let U be a utility function that is strictly increasing, strictly concave, and satisfies Inada's conditions. An investor puts his initial wealth x into the markets for the risky assets and for the riskless asset, and seeks to maximize $E\left(U\left(X_T^{x,\pi}\right)\right)$, where the expectation is taken under the historic measure. Davis [87] determines the fair price of a contingent claim ζ using an argument based on the marginal rate of substitution. We suppose that the claim ζ is sold at a price of $p(\zeta)$. If the agent invests an amount δ in the contingent claim ζ, and if he keeps this position in place until expiry, his final wealth will be given by $X_T^{x-\delta,\pi} + \frac{\delta}{p(\zeta)}\zeta$. His optimization problem consists in achieving

$$W(\delta, x, p) = \sup_{\pi} E\left(U\left(X_T^{x-\delta,\pi} + \frac{\delta}{p(\zeta)}\zeta\right)\right).$$

Definition 8.1.5. *Suppose that the equation*

$$\frac{\partial W}{\partial \delta}(0, p, x) = 0$$

has a unique solution p^*. Then the fair price for the contingent claim ζ is p^*.

An "infinitesimal" investment in the contingent claim has a neutral effect on the agent's optimization problem.

Theorem 8.1.6. *Let $V(x) = \sup_{\pi} E\left(U\left(X_T^{\pi,x}\right)\right) = E\left(U\left(X_T^{\pi^*,x}\right)\right)$.*
We assume that V is differentiable and that $V'(x) > 0$. Then p^* is given by

$$p^* = \frac{E\left(U'\left(X_T^{\pi^*,x}\right)\zeta\right)}{V'(x)}.$$

In the complete market case, or when the contingent claim is replicable, the value of p^* is the same as the one obtained by the usual approach: the fair price is then the expectation of $R(T)\zeta$ under a risk-neutral measure. Indeed, when ζ is replicable, $\tilde{p} = E_Q\left(R(T)\zeta\right)$ is the initial value of the replicating portfolio: the initial wealth $\tilde{p}$ can be used to construct a self-financing portfolio made up of the market assets $(\theta_t, \ 0 \leq t \leq T)$, and having initial value $V_0(\theta) = \tilde{p}$, and final value $V_T(\theta) = \zeta$. If the contingent claim ζ is for sale at price p, an investor can, for any choice of δ, buy an amount δ/p of the contingent claim at price p, and invest his remaining wealth $x - \delta$ in a portfolio of market assets. Initially, he seeks the optimal strategy for a given fixed δ, and next he optimizes his choice of δ. As the contingent claim is replicable, the optimal solution is to choose $\delta = 0$.

Let us implement the first part of the optimization scheme, with δ fixed. The δ/p shares in the contingent claim have a final value $\delta\zeta/p$, and the portfolio $\delta\theta/p$ has initial value $V_0\left(\delta\theta/p\right) = \tilde{p}\delta/p$, and final value $V_T\left(\delta\theta/p\right) = \delta\zeta/p$.

It is easy to see that the optimal way of investing the wealth $x - \delta$ is to sell the portfolio $\delta\theta/p$ and invest the wealth $x - \delta + V_0\left(\delta\theta/p\right) = x - \delta + \tilde{p}\delta/p$, in an optimal manner, in the complete market. The optimal solution to our initial problem is then to choose $\delta = 0$. The marginal rate of substitution is $\frac{\mathrm{d}}{\mathrm{d}\delta}V\left(\frac{\delta p}{\tilde{p}} + x - \delta\right)\Big|_{\delta=0} = \left(\frac{\tilde{p}}{p-1}\right)V'(x)$, and this is zero only if $p = \tilde{p}$.

Davis' price is a viable price in that the fair price is within the price range determined previously, and thus it does not give rise to any arbitrage opportunities. We do not reproduce the proof here (see for example Pham [299]).

Hodges and Neuberger [194] suggest the following method. For a given utility function, the "reservation price" is the value of p such that

$$\max_{\pi} E\left(U\left(X_T^{x,\pi}\right)\right) = \max_{\pi} E\left(U\left(X_T^{x-p,\pi} + \xi\right)\right) .$$

The reader can refer to El Karoui–Rouge [147].

8.1.7 Transaction Costs

Valuation in the case where transactions incur costs is an interesting topic, as it is very close to what happens in the real world. The results however are often disappointing. In the case of proportional transaction costs and of a European option, Shreve et al. [341] show that the minimal superhedge is the trivial hedge, which involves buying the underlying.

8.2 Stochastic Volatility

Models with stochastic volatility are such that

$$\mathrm{d}S_t = S_t\left(\mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t\right)$$

where the volatility follows a stochastic process. Numerous studies have been carried out in the case where the volatility dynamics are specified using a noise process that is different from that of the underlying:

$$\mathrm{d}\sigma_t = b(t, \sigma_t)\mathrm{d}t + \Sigma(t, \sigma_t)\mathrm{d}B_t .$$

Models in which $\mathrm{d}S_t = S_t\left(\mu_t \mathrm{d}t + \sigma\left(t, S_t\right)\mathrm{d}W_t\right)$, where $\sigma : \mathbb{R}^+ \times \mathbb{R}$ is a deterministic function are not considered to be of stochastic volatility. In such models, there exists a sole risk-neutral measure, given by $\frac{\mathrm{d}Q}{\mathrm{d}P}\Big|\mathcal{F}_t = L_t$ with $L_t = \exp\left[\int_0^t \theta_s \mathrm{d}W_s - \frac{1}{2}\int_0^t \theta_s^2 \mathrm{d}s\right]$ and $\theta_s = \frac{\mu_s - r}{\sigma(s, S_s)}$. Nonetheless, it is not straightforward to provide explicit valuation calculations in these models.

8.2.1 The Robustness of the Black–Scholes Formula

As stated earlier, the price of a European call in the Black–Scholes model is increasing with respect to volatility. This property can be extended to the stochastic volatility case as follows. Let $dS_t = S_t(r dt + \sigma_t dB_t)$ be the price process for the risky asset under a risk-neutral measure Q. We use the notation $\mathcal{BS}(t, x, \sigma)$ for the Black–Scholes function defined by

$$\mathcal{BS}(t, x, \sigma) = E\left(e^{-r(T-t)}(X_T - K)^+ \,|X_t = x\right),$$

where X is the geometric Brownian motion with constant drift r and constant volatility σ. The function gives the price of a European option in the Black–Scholes model with volatility σ. Now we examine the error a financial agent makes if he does not know the exact value of $(\sigma_t, \ t \geq 0)$, but knows that this volatility remains between two bounds, which are either constant or deterministic. He evaluates the price of the option using the Black–Scholes function, and hedges using the delta obtained by applying the Black–Scholes formula.

Theorem 8.2.1. *If $\sigma_1 \leq \sigma_t \leq \sigma_2$ for all t and for almost all ω, where the σ_i's are constants,*

$$\mathcal{BS}(t, S_t, \sigma_1) \leq E_Q\left(e^{-r(T-t)}(S_T - K)^+ \,|\mathcal{F}_t\right) \leq \mathcal{BS}(t, S_t, \sigma_2).$$

Proof. The proof is very simple and uses the fundamental properties of the Black–Scholes function, in the shape of the its PDE and its convexity. We apply Itô's formula to the discounted Black–Scholes function $e^{-rt}\mathcal{BS}(t, S_t, \sigma_2)$ evaluated at the "real" market levels. We obtain:

$$e^{-rt}\mathcal{BS}(t, S_t, \sigma_2) = \mathcal{BS}(0, S_0, \sigma_2) + \int_0^t e^{-rs}\frac{\partial \mathcal{BS}}{\partial t}(s, S_s, \sigma_2)\, ds$$

$$+ \int_0^t e^{-rs}\frac{\partial \mathcal{BS}}{\partial x}(s, S_s, \sigma_2)\, dS_s + \frac{1}{2}\int_0^t e^{-rs}\frac{\partial^2 \mathcal{BS}}{\partial x^2}(s, S_s, \sigma_2)S_s^2\sigma_s^2\, ds$$

$$- \int_0^t r e^{-rt}\mathcal{BS}(s, S_s, \sigma_2)\, ds$$

$$= \mathcal{BS}(0, S_0, \sigma_2) + \int_0^t e^{-rs}\frac{\partial \mathcal{BS}}{\partial x}(s, S_s, \sigma_2)\sigma_s S_s\, ds$$

$$+ \int_0^t e^{-rs}\left[\frac{\partial \mathcal{BS}}{\partial t} + \frac{\partial \mathcal{BS}}{\partial x}r S_s + \frac{1}{2}\frac{\partial^2 \mathcal{BS}}{\partial x^2}S_s^2\sigma_s^2 - r e^{-rs}\mathcal{BS}\right](s, S_s, \sigma_2)\, ds.$$

The Black–Scholes function $\mathcal{BS}(t, x, \sigma_2)$ satisfies

$$\frac{\partial \mathcal{BS}}{\partial t} + xr\frac{\partial \mathcal{BS}}{\partial x} + x^2\sigma_2^2\frac{1}{2}\frac{\partial^2 \mathcal{BS}}{\partial x^2} - r\mathcal{BS} = 0.$$

It follows that

$$e^{-rt}\mathcal{BS}\left(t, S_t, \sigma_2\right) = \mathcal{BS}\left(0, S_0, \sigma_2\right)$$
$$+ \frac{1}{2}\int_0^t e^{-rs}\frac{\partial^2 \mathcal{BS}}{\partial x^2}\left(s, S_s, \sigma_2\right)S_s^2\left(\sigma_s^2 - \sigma_2^2\right)\mathrm{d}s + M_t$$

where M is a martingale under the risk-neutral measure. Now it is enough to use the same formula at time T, and take expectations under the risk-neutral measure, to obtain

$$E_Q\left[e^{-rT}\mathcal{BS}\left(T, S_T, \sigma_2\right)\right] = E_Q\left(e^{-rT}\left(S_T - K\right)^+\right)$$

$$= \mathcal{BS}\left(0, S_0, \sigma_2\right) + \frac{1}{2}\int_0^T e^{-rs}\frac{\partial^2 \mathcal{BS}}{\partial x^2}\left(s, S_s, \sigma_2\right)S_s^2\left(\sigma_s^2 - \sigma_2^2\right)\mathrm{d}s$$
$$\leq \mathcal{BS}\left(0, S_0, \sigma_2\right) .$$

$\square$

To hedge himself, the agent calculates the hedging "delta" using the Black–Scholes methodology, that is using the function $\frac{\partial \mathcal{BS}}{\partial x}\left(t, x, \sigma_2\right)$ evaluated at the observed levels, and constructs a portfolio containing $\Delta = \frac{\partial \mathcal{BS}}{\partial x}\left(t, S_t, \sigma_2\right)$ shares. The value of this self-financing portfolio is $\Pi_\Delta(t)$ and its price dynamics are given by

$$\mathrm{d}\Pi_\Delta(t) = r\Pi_\Delta(t)\mathrm{d}t + \Delta\left[\mathrm{d}S_t - rS_t\mathrm{d}t\right] .$$

The value of this portfolio is always greater than the estimated value of the contingent claim ζ. Indeed, if we write $e(t) = \Pi_\Delta(t) - \mathcal{BS}\left(t, S_t, \sigma_2\right)$, Itô's formula shows that

$$e(t) = \frac{1}{2}e^{rt}\int_0^t e^{-ru}\left[\sigma_2^2 - \sigma_u^2\right]S_u^2\frac{\partial \mathcal{BS}}{\partial x^2}\left(u, S_u, \sigma_2\right)\mathrm{d}u .$$

So the agent is superhedged. A counter-example in the case of a stochastic σ_2 is provided by El Karoui and Jeanblanc-Picqué [138].

This study is carried out in a more general setting in Avellaneda, Levy and Parás [17] and in Gozzi and Vargioglu [176].

8.3 Wealth Optimization

One approach to the problem of portfolio optimization in the case of incomplete markets, is to complete the market by adding fictitious assets that the agent is not allowed to use in his portfolio.

Let $\rho(t)$ be a $(n - d) \times n$ matrix whose row vectors are orthogonal and such that $\sigma(t)\rho^T(t) = 0$. We complete the market with the assets

$$\mathrm{d}S_i(t) = S_i(t) \left[\beta_i(t)\mathrm{d}t + \sum_{j=1}^{n} \rho_{ij}(t)\mathrm{d}W_j(t) \right] , \qquad i = d + 1, \ldots, n .$$

The matrix ρ is taken to be fixed, and the vector β is adjusted in such a way that the optimal trading strategy in the completed market is given by a portfolio that does invest in any of the fictitious assets. The idea is an intuitive one: if β is large, the investor would prefer to have a long position in the fictitious assets in his portfolio, and if β is small he would prefer to have a short position. It must be possible to adjust the β coefficient in such a way that the agent does not have a position in the fictitious assets, in his portfolio. Karatzas [230] provides further details, as well as a rigorous proof.

Another approach is to complete the market and work with constrained portfolios. Duality techniques are then very effective, and can also be applied to other types of constraints. The reader can consult Cvitanić's lecture notes, published in [32], and Karatzas [230].

Notes

Pricing in an incomplete market remains a challenge for practitioners. One can refer to Bingham and Kiesel [33], (1998), and Björk [34], (1998), for a general presentation, to Miyahara [280], (1997), and Frittelli [166], (2000), for the minimal entropy measure, and to Cvitanić [32, 74], (2001), for a utility approach.

9

Exotic Options

The valuation and hedging of the ever increasing number of exotic options, is a topic that interests many practitioners seeking to answer their customers' need to hedge risk (in particular in the foreign exchange markets). This last chapter is devoted to the mathematical problems related to these products.

Exotic options, or "path-dependent" options, are options whose payoff depends on the behaviour of the price of the underlying between time 0 and the maturity (here assumed to be fixed), rather than merely on the final price of the underlying. We only deal with the case of European options here. American options (with which the buyer can exercise his right at any time between 0 and the maturity) are covered in Lamberton–Lapeyre [250], as well as in Elliott and Kopp [149]. The paper Myneni [287] is also good reference, though it is less accessible. One can also consult Jarrow and Rudd [213], Bensoussan [28], Karatzas [229]. Path-dependent options of the American type are traded, though their pricing formulae are not known explicitly.

Path-dependent options are traded mainly in the foreign exchange markets. Numerous authors have studied them in discrete time. One can look to Willmott et al. [371], Chesney et al. [57], Musiela–Rutkowski [285], Zhang [379], Pliska [301] and to the references therein. These works also tackle the continuous time case, to which we devote the final part of this chapter.

As in the Black–Scholes framework, let us assume that the price of the underlying follows, under the risk-neutral measure Q, the dynamics

$$\mathrm{d}S_t = S_t(r\mathrm{d}t + \sigma\mathrm{d}B_t)$$

with B a Brownian motion. We suppose that $\sigma > 0$. This is not a very restrictive assumption: if it did not hold, we would only need to change B into $-B$, which is also a Brownian motion.

Our first section is devoted to the probability distributions associated with the Brownian motion and its supremum. The results will serve later on in the

chapter. In the following section, we study the valuation problem for barrier options, and subsequent sections touch upon other exotic options.

For all that follows, we take a space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and a Brownian motion $(B_t, t \geq 0)$ starting at zero and constructed on this space.

We write $\mathcal{N}(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-\frac{u^2}{2}} \, du$ for the standard cumulative normal distribution.

9.1 The Hitting Time and Supremum for Brownian Motion

We study the distribution of the pair of random variables (B_t, M_t), where M is the process for the Brownian motion's maximum, i.e., $M_t \overset{\text{def}}{=} \sup_{s \leq t} B_s$. From this distribution, we deduce the distribution of the hitting time: the first time that the Brownian motion B hits a given level.

9.1.1 Distribution of the Pair (B_t, M_t)

Theorem 9.1.1. *Let B be a Brownian motion starting at 0 and let $M_t = \sup (B_s, 0 \leq s \leq t)$. Then:*

$$
\begin{cases}
for \quad y \geq 0, x \leq y \qquad P(B_t \leq x, M_t \leq y) = \mathcal{N}\left(\dfrac{x}{\sqrt{t}}\right) - \mathcal{N}\left(\dfrac{x - 2y}{\sqrt{t}}\right), \\[3mm]
for \quad y \geq 0, x \geq y \qquad P(B_t \leq x, M_t \leq y) = \mathcal{N}\left(\dfrac{y}{\sqrt{t}}\right) - \mathcal{N}\left(\dfrac{-y}{\sqrt{t}}\right), \\[3mm]
for \quad y \leq 0 \qquad\qquad\quad P(B_t \leq x, M_t \leq y) = 0 \, .
\end{cases}
$$
$$\tag{9.1}$$

Proof. Calculating the law of the pair (M_t, B_t) depends on which area of the plane we are in.

- *The Reflection Principle*
  Let us show that

$$
for \quad 0 \leq y, \, x \leq y, \quad P(B_t \leq x \,, M_t \geq y) = P(B_t \geq 2y - x) \, . \tag{9.2}
$$

  Let $T_y = \inf\{t \,:\, B_t \geq y\}$ be the hitting time for level y. It is a stopping time and, trivially, $(T_y \leq t) = (M_t \geq y)$; also, for $y \geq 0$, by continuity of the Brownian motion's paths, we have $T_y = \inf\{t \,:\, B_t = y\}$ and $B_{T_y} = y$. Hence

$$
P(B_t \leq x \,, M_t \geq y) = P(B_t \leq x \,, T_y \leq t) = P(B_t - B_{T_y} \leq x - y \,, T_y \leq t) \, .
$$

We introduce conditioning with respect to T_y, and to simplify the notation we write $P(A|T_y) = E_P(\mathbb{1}_A|T_y)$. Using the strong Markov property, the previous expression becomes

$$E(\mathbb{1}_{T_y \leq t}\, P(B_t - B_{T_y} \leq x - y\,|T_y)) = E(\mathbb{1}_{T_y \leq t}\, \Phi(T_y))$$

with $\Phi(u) = P(\tilde{B}_{t-u} \leq x - y)$ where $\tilde{B} = (\tilde{B}_t = B_{t+T_y} - B_{T_y},\, t \geq 0)$ is a Brownian motion that is independent of $(B_t, t \leq T_y)$, and has the same distribution as $-\tilde{B}$. It follows that $\Phi(u) = P(\tilde{B}_{t-u} \geq y - x)$. Going back to our previous computations, we get

$$\begin{aligned}
E(\mathbb{1}_{T_y \leq t}\, \Phi(T_y)) &= E(\mathbb{1}_{T_y \leq t} P(B_t - B_{T_y} \geq y - x\,|T_y)) \\
&= P(B_t \geq 2y - x,\, T_y \leq t)\,,
\end{aligned}$$

where this last term equals $P(B_t \geq 2y - x)$, since $2y - x \geq x$.

- In the area of the plane $0 \leq y \leq x$, as $M_t \geq B_t$, we have

$$P(B_t \leq x, M_t \leq y) = P(B_t \leq y, M_t \leq y) = P(M_t \leq y)\,,$$

  where the last term can be evaluated using the previous calculations.

- Finally, for $y \leq 0$, $P(B_t \leq x, M_t \leq y) = 0$ as $M_t \geq M_0 = 0$.

$\square$

9.1.2 Distribution of Sup and of the Hitting Time

Proposition 9.1.2. *The random variable M_t has the same distribution as $|B_t|$.*

Proof. Indeed, for $x \geq 0$:

$$P(M_t \geq x) = P(M_t \geq x, B_t \geq x) + P(M_t \geq x, B_t \leq x)\,.$$

Using (9.2) (with $x = y$), we obtain:

$$\begin{aligned}
P(M_t \geq x) &= P(B_t \geq x) + P(B_t \geq x) \\
&= P(B_t \geq x) + P(B_t \leq -x) = P(|B_t| \geq x)\,.
\end{aligned}$$

$\square$

We obtain the distribution of $T_x = \inf\{s : B_s \geq x\}$ by noting that

$$P(T_x \leq t) = P(M_t \geq x)\,.$$

Hence, the density of T_x is, for $x > 0$,

$$P\left(T_x \in dy\right) \frac{x}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right) \mathbb{1}_{t\geq 0} \, dt \ . \tag{9.3}$$

It is worth noting that the process M does not have the same distribution as the process $|B|$. The former is increasing, the latter is not. The reader can refer to Revuz–Yor [307] for more information.

9.1.3 Distribution of Inf

The distribution of inf is obtained using the same principles as above, or by noticing that

$$m_t \stackrel{\text{def}}{=} \inf_{s\leq t} B_s = -\sup_{s\leq t}(-B_s) \, ,$$

where $-B$ is a Brownian motion. Hence

$$\text{for} \quad y \leq 0, x \geq y \qquad P(B_t \geq x, m_t \geq y) \;=\; \mathcal{N}\left(\frac{-x}{\sqrt{t}}\right) - \mathcal{N}\left(\frac{2y-x}{\sqrt{t}}\right)$$

$$\text{for} \quad y \leq 0, x \leq y \qquad P(B_t \geq x, m_t \geq y) \;=\; \mathcal{N}\left(\frac{-y}{\sqrt{t}}\right) - \mathcal{N}\left(\frac{y}{\sqrt{t}}\right)$$

$$\text{for} \quad y \geq 0 \qquad\qquad P(B_t \geq x, m_t \geq y) \;=\; 0 \, .$$

$$\tag{9.4}$$

In particular, $P(m_t \geq y) = \mathcal{N}\left(\frac{-y}{\sqrt{t}}\right) - \mathcal{N}\left(\frac{y}{\sqrt{t}}\right)$.

9.1.4 Laplace Tranforms

We know that for any λ, the process $\left(\exp\left(\lambda B_t - \frac{\lambda^2}{2}t\right), t \geq 0\right)$ is a martingale. Let $y \geq 0, \lambda \geq 0$ and let T_y be the hitting time for level y. The martingale

$$\exp\left(\lambda B_{t\wedge T_y} - \frac{\lambda^2}{2}(t \wedge T_y)\right)$$

is bounded by $e^{\lambda y}$, so is uniformly integrable, and the optimal stopping theorem (see annex) yields

$$E\left[\exp\left(\lambda B_{T_y} - \frac{\lambda^2}{2}T_y\right)\right] = 1$$

i.e.,

$$E\left[\exp\left(-\frac{\lambda^2}{2}T_y\right)\right] = \exp(-y\lambda) \, .$$

9.1.5 Hitting Time for a Double Barrier

Let $a < 0 < b$, let T_a and T_b be the corresponding hitting times, and let $T^* = T_a \wedge T_b$ be the hitting time for the double barrier. We continue to use the notation M for the Brownian motion's maximum, and m for its minimum.

Proposition 9.1.3. *The Laplace transform of T^* is*

$$E\left[\exp\left(-\frac{\lambda^2}{2}T^*\right)\right] = \frac{\cosh[\lambda(a+b)/2]}{\cosh[\lambda(b-a)/2]}.$$

The joint distribution of (M_t, m_t, B_t) is given by

$$P\left(a \le m_t < M_t \le b, B_t \in E\right) = \int_E k(x)\,\mathrm{d}x$$

where, for $E \subset [a, b]$,

$$k(x) = \frac{1}{\sqrt{2\pi t}} \sum_{k=-\infty}^{\infty} \left[\exp\left(-\frac{1}{2t}(x + 2k(b-a))^2\right)\right.$$
$$\left. - \exp\left(-\frac{1}{2t}(x - 2b + 2k(b-a))^2\right)\right]. \quad (9.5)$$

Proof. The Laplace transform of T^* can be obtained by applying the optimal stopping theorem. Indeed,

$$\exp\left[-\lambda\left(\frac{a+b}{2}\right)\right] = E\left[\exp\left(\lambda\left(B_{T^*} - \frac{a+b}{2}\right) - \frac{\lambda^2 T^*}{2}\right)\right]$$
$$= \exp\left(\lambda\frac{b-a}{2}\right) E\left[\exp\left(-\frac{\lambda^2 T^*}{2}\right) \mathbb{1}_{T^*=T_b}\right]$$
$$+ \exp\left(\lambda\frac{a-b}{2}\right) E\left[\exp\left(-\frac{\lambda^2 T^*}{2}\right) \mathbb{1}_{T^*=T_a}\right]$$

and, using $-B$,

$$\exp\left[-\lambda\left(\frac{a+b}{2}\right)\right] = E\left[\exp\left(\lambda\left(-B_{T^*} - \frac{a+b}{2}\right) - \frac{\lambda^2 T^*}{2}\right)\right]$$
$$= \exp\left(\lambda\frac{-3b-a}{2}\right) E\left[\exp\left(-\frac{\lambda^2 T^*}{2}\right) \mathbb{1}_{T^*=T_b}\right]$$
$$+ \exp\left(\lambda\frac{-b-3a}{2}\right) E\left[\exp\left(-\frac{\lambda^2 T^*}{2}\right) \mathbb{1}_{T^*=T_a}\right].$$

By solving this system of two equations with two unknowns, we find:

$$\begin{cases} E\left[\exp\left(-\dfrac{\lambda^2 T^*}{2}\right) \mathbb{1}_{T^*=T_b}\right] = \dfrac{\sinh(-\lambda a)}{\sinh(\lambda(b-a))} \\[4mm] E\left[\exp\left(-\dfrac{\lambda^2 T^*}{2}\right) \mathbb{1}_{T^*=T_a}\right] = \dfrac{\sinh(\lambda b)}{\sinh(\lambda(b-a))}. \end{cases}$$

The first assertion follows by noting that

$$E\left[\exp\left(-\frac{\lambda^2}{2}T^*\right)\right]$$
$$= E\left[\exp\left(-\frac{\lambda^2\,T^*}{2}\right)1\!\!1_{T^*=T_b}\right] + E\left[\exp\left(-\frac{\lambda^2\,T^*}{2}\right)1\!\!1_{T^*=T_a}\right].$$

The density can then be obtained by inverting the Laplace transform. A direct proof of (9.5) appears as an exercise in Revuz–Yor [307]. $\qquad\square$

9.2 Drifted Brownian Motion

9.2.1 The Laplace Transform of a Hitting Time

Let X be the process defined by $X_t = \mu t + B_t$ and let $T_a^\mu = \inf\{t \geq 0; X_t = a\}$. Then we have

$$E\left(\exp\left(-\frac{\lambda^2}{2}T_a^\mu\right)\right) = \exp\left(\mu a - |a|\sqrt{\mu^2+\lambda^2}\right).$$

To see this is the case, we use Girsanov's theorem. Let Q be the probability measure defined by $dQ|_{\mathcal{F}_t} = \exp\left(-\mu B_t - \frac{1}{2}\mu^2 t\right)dP|_{\mathcal{F}_t}$, which can also be written

$$dP|_{\mathcal{F}_t} = \exp\left(\mu X_t - \frac{1}{2}\mu^2 t\right)dQ|_{\mathcal{F}_t}.$$

Under Q, the process X is a Brownian motion, and T_a is a.s. finite. Furthermore,

$$E_P\left(\exp\left(-\frac{\lambda^2}{2}T_a\right)\right) = E_Q\left[\exp\left(\mu X_{T_a} - \frac{1}{2}\mu^2 T_a\right)\exp\left(-\frac{\lambda^2}{2}T_a\right)\right],$$

where on the LHS as on the RHS, $T_a = \inf\{t \geq 0; X_t = a\}$. The RHS equals

$$E_Q\left[\exp\left(\mu a - \frac{1}{2}(\mu^2+\lambda^2)T_a\right)\right] = e^{\mu a}E_Q\left[\exp\left(-\frac{1}{2}(\mu^2+\lambda^2)T_a\right)\right],$$

which we know how to calculate since under Q, T_a is the hitting time of level a for a Brownian motion.

If X is defined by $X_t = \mu t + \sigma B_t$, it is enough to note that $\inf\{t \mid X_t = a\} = \inf\left\{t \mid \frac{\mu}{\sigma}t + B_t = \frac{a}{\sigma}\right\}$, in order to obtain the transform of the hitting time for level 0.

9.2.2 Distribution of the Pair (Maximum, Minimum)

Let $X_t = \mu t + \sigma B_t$ with $\sigma > 0$, and let $M_t^X = \sup(X_s, s \leq t)$, and $m_t^X = \inf(X_s, s \leq t)$. Girsanov's theorem enables us to transform $\sigma^{-1}X$ into a Brownian motion. Let Q be defined by $dP = L\, dQ$, where $L_t = \exp\left[-\dfrac{\mu}{\sigma}B_t - \dfrac{1}{2}\dfrac{\mu^2}{\sigma^2}t\right]$. Then $dP = L^{-1}dQ$ with $L_t^{-1} = \exp\left[\dfrac{\mu}{\sigma}B_t + \dfrac{1}{2}\dfrac{\mu^2}{\sigma^2}t\right] = \exp\left[\dfrac{\mu}{\sigma}W_t - \dfrac{1}{2}\dfrac{\mu^2}{\sigma^2}t\right]$ where $W_t = B_t + \dfrac{\mu}{\sigma}t$ is a Q-Brownian motion, and

$$P(X_t \leq x, M_t^X \leq y) = E_Q\left[\exp\left[\frac{\mu}{\sigma}W_t - \frac{\mu^2}{2\sigma^2}t\right] \mathbb{1}_{\left\{W_t \leq \frac{x}{\sigma}, M_t^W \leq \frac{y}{\sigma}\right\}}\right].$$

Straightforward but lengthy calculations lead to

- for $y \geq 0$, $y \geq x$,

$$P(X_t \leq x, M_t^X \geq y) = e^{\frac{2\mu y}{\sigma^2}} P(X_t \geq 2y - x + 2\mu t)$$

$$P(X_t \leq x, M_t^X \leq y) = \mathcal{N}\left(\frac{x - \mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu y}{\sigma^2}}\mathcal{N}\left(\frac{x - 2y - \mu t}{\sigma\sqrt{t}}\right)$$

- and for $y \leq 0$, $y \leq x$,

$$P(X_t \geq x, m_t^X \geq y) = \mathcal{N}\left(\frac{-x + \mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu y}{\sigma^2}}\mathcal{N}\left(\frac{-x + 2y + \mu t}{\sigma\sqrt{t}}\right).$$

In particular, we can deduce from above the distribution of the maximum as well as that of the minimum:

$$
\begin{aligned}
P(M_t^X \leq y) &= \mathcal{N}\left(\frac{y - \mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu y}{\sigma^2}}\mathcal{N}\left(\frac{-y - \mu t}{\sigma\sqrt{t}}\right), && y \geq 0, \\
P(m_t^X \geq y) &= \mathcal{N}\left(\frac{-y + \mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu y}{\sigma^2}}\mathcal{N}\left(\frac{y + \mu t}{\sigma\sqrt{t}}\right), && y \leq 0.
\end{aligned}
\tag{9.6}
$$

Let $y > 0$ and let $T_y = \inf\{t \geq 0 \,|\, X_t \geq y\}$. The density of T_y can be computed by using the equality $\{T_y \geq t\} = \{M_t^X \leq y\}$. In particular, by letting $t \to \infty$ in the expression $P(T_y \geq t) = \mathcal{N}\left(\frac{y-\mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu y}{\sigma^2}}\mathcal{N}\left(\frac{-y-\mu t}{\sigma\sqrt{t}}\right)$, we find that if $\mu \leq 0$ and $y > 0$, then $P(T_y = \infty) = 1 - e^{\frac{2\mu y}{\sigma^2}}$, which is zero only when $\mu = 0$.

9.2.3 Evaluation of $E(e^{-rT_y}\mathbb{1}_{T_y < a})$

Let $X_t = \mu t + B_t$. We write

$$E\left(e^{-\mu T_y}\mathbb{1}_{T_y < a}\right) = e^{-y\sqrt{2\mu}}E\left(e^{B_{T_y}\sqrt{2\mu} - \mu T_y}\mathbb{1}_{T_y < a}\right).$$

Applying Girsanov's theorem with $\tilde{Q}$ defined by $d\tilde{Q} = L_t\, dP$ for $L_t = e^{B_t\sqrt{2\mu} - \mu t}$, yields $E_P\left(e^{B_{T_y}\sqrt{2\mu} - \mu T_y}\,\mathbb{1}_{T_y < a}\right) = E_{\tilde{Q}}\left(\mathbb{1}_{T_y < a}\right)$, with $\tilde{B}_t = B_t - \sqrt{2\mu}t$, a Brownian motion under $\tilde{Q}$. The stopping time T_y can be written as

$$T_y = \inf\{t\,:\, B_t = y\} = \inf\{t\,,\, \tilde{B}_t + \sqrt{2\mu}t = y\}\,.$$

Calculating $E_Q(\mathbb{1}_{T_y < a})$ depends on the density of the hitting time for a Brownian motion with drift. Making use of formula (9.6), we obtain

$$E(e^{-\mu T_y}\mathbb{1}_{T_y < a}) = e^{-y\sqrt{2\mu}}\mathcal{N}\left(-\frac{y}{\sqrt{a}} + \sqrt{2a\mu}\right)$$

$$+\, e^{y\sqrt{2\mu}}\mathcal{N}\left(-\frac{y}{\sqrt{a}} - \sqrt{2a\mu}\right)\,. \quad (9.7)$$

9.3 Barrier Options

For this kind of option, a level (the barrier) L is fixed in the same way that the strike price of an option is fixed. Both calls (options to buy) and puts (options to sell) can be associated with a barrier. We give explicit workings only for call options with strike K; in the case of a put the calculations are very similar.

9.3.1 Down-and-Out Options

The buyer of the option looses his right of exercise if the price of the underlying $(S_t, t \geq 0)$ falls beneath the level L before maturity (we assume that $L < S_0$). Otherwise, the option holder receives a payoff $\phi(S_T)$, with $\phi(x) = (x - K)^+$ for a call option (respectively $(K - x)^+$ for a put), where K is the fixed strike price.

The price of a down-and-out call is:

$$\mathrm{DOC}(S_0, K, L) \stackrel{\mathrm{def}}{=} E_Q(e^{-rT}(S_T - K)^+\mathbb{1}_{T_L > T})$$

where T_L is the stopping time at which the underlying price crosses the barrier for the first time:

$$T_L \stackrel{\mathrm{def}}{=} \inf\{t\,|\, S_t \leq L\} = \inf\{t\,|\, S_t = L\}\,,$$

with the last equality being a consequence of the assumption $L < S_0$ and of path continuity.

We can also consider the case where the option holder receives a compensation F if the barrier is crossed. This compensation is agreed upon when the contract is signed, and is received either when the barrier is crossed or at maturity. The compensation is bounded above by $F\, E_Q(e^{-rT_L}\mathbb{1}_{T_L \leq T})$ for the first mode of payment, and by $F\, E_Q(e^{-rT}\mathbb{1}_{T_L \leq T})$ for the second. These two expressions only involve the distribution of T_L, which is known; we will provide explicit calculations later.

9.3.2 Down-and-In Options

The buyer of a down-and-in call receives the payoff only in the case where the level L is reached before time T. The price of such an option is:

$$\mathrm{DIC}(S_0, K, L) \stackrel{\mathrm{def}}{=} E_Q \left(\mathrm{e}^{-rT} (S_T - K)^+ \mathbb{1}_{T_L < T} \right) .$$

It is then obvious that:

$$\mathrm{DOC}(S, K, L) + \mathrm{DIC}(S, K, L) = C(S, K)$$

where $C(S, K)$ is the price of a call with strike K on an underlying that has initial value S. The relation above enables us to limit ourselves to the study of down-and-in options.

9.3.3 Up-and-Out and Up-and-In Options

An up-and-out option becomes worthless if the level $H > S_0$ is reached before time T, whereas the up-and-in option is activated when the underlying crosses the barrier on its way up.

The price of an up-and-out option is

$$E_Q(\mathrm{e}^{-rT}(S_T - K)^+ \mathbb{1}_{T < T_H}) ,$$

and that of an up-and-in option is

$$E_Q(\mathrm{e}^{-rT}(S_T - K)^+ \mathbb{1}_{T > T_H}) ,$$

where

$$T_H = \inf\{t \,|\, S_t \geq H\} = \inf\{t \,|\, S_t = H\} .$$

9.3.4 Intermediate Calculations

- *Evaluating the Integrals*

    We start by evaluating the integrals that will appear in subsequent calculations. Let $a < b$, m and c be real numbers. Elementary calculations using a change of variable yield

    $$I(a, b; m, c) \stackrel{\mathrm{def}}{=} \frac{1}{\sqrt{2\pi T}} \int_a^b \exp\left(mx - \frac{1}{2T}(x - c)^2 \right) \mathrm{d}x$$
    $$= (\mathcal{N}(\beta) - \mathcal{N}(\alpha)) \exp\left(\frac{m^2 T}{2} + mc \right)$$

    with

$$\alpha = \frac{a - (c + mT)}{\sqrt{T}} \, , \qquad \beta = \frac{b - (c + mT)}{\sqrt{T}} \, .$$

Another integral that we will need later is, for $a > 0$, $x > 0$ and $y > 0$,

$$J(a, x, y) \stackrel{\text{def}}{=} \int_0^a e^{-xt} \frac{1}{\sqrt{t^3}} \, e^{-y/t} \, dt \, ,$$

which can be written as $\dfrac{1}{\sqrt{y}} \tilde{J}\left(\dfrac{a}{y}, \, xy\right)$ with

$$\tilde{J}(a, \, x) \stackrel{\text{def}}{=} \int_0^a e^{-xt} \frac{1}{\sqrt{t^3}} \, e^{-1/t} \, dt \, .$$

We can evaluate the integral above by introducing the change of variable $v = \sqrt{\dfrac{2}{t}}$, and then $v - \dfrac{\sqrt{x}}{v} = u$, so that $v = \dfrac{1}{2}\left(u + \sqrt{u^2 + 4\sqrt{x}}\right)$, and finally $t = \sqrt{u^2 + 4\sqrt{x}}$. A more elegant approach involves using probabilistic representations of the integrals above. Firstly, using the density of T_z (see formula (9.3)), we find that

$$E\left(e^{-rT_z} \mathbb{1}_{T_z \leq a}\right) = \frac{z}{\sqrt{2\pi}} J\left(a, r, \frac{z^2}{2}\right) = \frac{1}{\sqrt{\pi}} \tilde{J}\left(\frac{2a}{z^2}, \frac{rz^2}{2}\right) \, .$$

It follows from (9.7) that

$$\tilde{J}(a, x) = \sqrt{\pi}\left[e^{-2\sqrt{x}} \mathcal{N}(\alpha) + e^{2\sqrt{x}} \mathcal{N}(\beta)\right]$$

with

$$\alpha = -\sqrt{\frac{2}{a}} + \sqrt{xa} \, ; \quad \beta = -\sqrt{\frac{2}{a}} - \sqrt{xa} \, .$$

This brings us back to the Laplace transform of the hitting time for level $z > 0$:

$$E\left(e^{-rT_z}\right) = \frac{1}{\sqrt{2\pi}} z \, J\left(\infty, r, \frac{z^2}{2}\right) = e^{-z\sqrt{2r}} \, .$$

The integral

$$K(x, z) \stackrel{\text{def}}{=} \int_0^\infty \frac{e^{-xt}}{\sqrt{2\pi t}} \, e^{-z^2/2t} dt$$

is obtained by differentiating $J\left(\infty, x, \dfrac{z^2}{2}\right)$ with respect to x:

$$K(x, z) = \frac{1}{\sqrt{2x}} \exp\left(-|z|\sqrt{2x}\right) \, .$$

- *Change of Measure*

  We express the various terms we need to evaluate to price barrier options as functions of the Brownian motion B, and then of the process $(W_t = B_t + mt, t \leq T)$ where $m = \dfrac{1}{\sigma} \left(r - \dfrac{\sigma^2}{2} \right)$.

  The price of the underlying asset can be written as

  $$S_t = S_0 \exp \left(\sigma B_t + rt - \frac{1}{2}\sigma^2 t \right) = S_0 \exp(\sigma W_t) \,,$$

  and the hitting time T_L can be expressed as a function of S or as a function of W by

  $$T_L = \inf \{ t \,:\, S_t \leq L \} = \inf \left\{ t \,:\, \sigma W_t \leq \ln \frac{L}{S_0} \right\} = \inf \{ t \,:\, W_t \leq \ell \} \overset{\text{def}}{=} T_\ell$$

  where we have set $\ell = \dfrac{1}{\sigma} \ln \dfrac{L}{S_0}$. Similarly, setting $h = \dfrac{1}{\sigma} \ln \dfrac{H}{S_0}$, we have

  $$T_H = \inf \left\{ t \,:\, \sigma W_t \geq \ln \frac{H}{S_0} \right\} = \inf \{ t \,:\, W_t \geq h \} \overset{\text{def}}{=} T_h \,.$$

  An expression of the form $E_Q(\Phi(S_T)\mathbb{1}_{T_L < T})$ can be evaluated either by using Girsanov's theorem and transforming the process W into a Brownian motion, or by using the density distribution of the pair (W, M) when W is a Brownian with drift, and M is its maximum. Here we elaborate on the method that uses Girsanov's Theorem.

  Let R be the probability measure defined by

  $$\frac{\mathrm{d}R}{\mathrm{d}Q} = \exp \left(-mB_T - \frac{1}{2}m^2 T \right) \,,$$

  i.e., by

  $$\frac{\mathrm{d}Q}{\mathrm{d}R} = \exp \left(mW_T - \frac{1}{2}m^2 T \right) \,.$$

  It follows that

  $$E_Q \left(\Phi(S_T)\mathbb{1}_{T_L < T} \right) = E_R \left(\Phi(S_0 \mathrm{e}^{\sigma W_T}) \exp \left(mW_T - \frac{m^2 T}{2} \right) \mathbb{1}_{T_\ell < T} \right)$$

  where W is a R-Brownian motion. Therefore, we need to evaluate $R(W_T \in \mathrm{d}x, T_\ell < T)$, which can be done using the reflection principle.

- *The Resolvent Operator*

  Let γ be a given function, and let B be a Brownian motion starting at x. To evaluate the Laplace transform of the function $t \to \gamma(B_t)$, i.e.,

  $$\Gamma(\lambda) = E\left(\int_0^\infty \exp\left[-\frac{\lambda^2 t}{2}\right] \gamma(B_t)\, dt\right) ,$$

  let us use our preliminary calculations and the density of B_t given by $\dfrac{1}{\sqrt{2\pi t}}\exp-\dfrac{(y-x)^2}{2t} dy$:

  $$\Gamma(\lambda) = \int_{-\infty}^\infty dy\, \gamma(y) \int_0^\infty dt \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \exp\left(-\frac{\lambda^2 t}{2}\right)$$
  $$= \frac{1}{\lambda}\int_{-\infty}^\infty dy\, \gamma(y) \exp(-\lambda|x-y|) .$$

  The last equality uses the function K from the previous section. The function Γ is known in the literature as the *resolvent operator* of Brownian motion.

9.3.5 The Value of the Compensation

- If the compensation is paid when the barrier is touched, we need to calculate $E_Q(\mathbb{1}_{T_L<T}e^{-rT_L}) = \Phi(L,T)$. To do this, we express it under the probability measure R:

  $$\Phi(L,T) = E_R\left[\exp\left(mW_{T_\ell} - \frac{m^2}{2}T_\ell\right) e^{-rT_\ell}\, \mathbb{1}_{T_\ell<T}\right]$$
  $$= e^{m\ell}E_R\left[\mathbb{1}_{T_\ell<T} \exp-\frac{2r+m^2}{2}T_\ell\right] = \frac{e^{m\ell}}{\sqrt{\pi}}\tilde{J}\left(\frac{2T}{\ell^2}, \frac{2r+m^2}{2}\frac{\ell^2}{2}\right) .$$

  Using the earlier integral calculations, and noting that $2r+m^2 = \left(\dfrac{r}{\sigma} + \sigma2\right)^2$, we obtain:

  $$\Phi(L,T) = e^{-\ell\sigma}\mathcal{N}(\gamma_1) + e^{2\ell r/\sigma}\mathcal{N}(\gamma_2)$$

  with

  $$\gamma_1 = \frac{1}{\sigma\sqrt{T}}\left(rT - \ell\sigma + \frac{\sigma^2 T}{2}\right) , \quad \gamma_2 = -\frac{1}{\sigma\sqrt{T}}\left(rT + \ell\sigma + \frac{\sigma^2 T}{2}\right) .$$

- If the compensation is paid at maturity, it is enough to calculate $E_Q(\mathbb{1}_{T_L<T})$, which has already been done above. Specifically:

$$E_Q(\mathbb{1}_{T_L < T}) = \frac{e^{m\ell}}{\sqrt{\pi}} \, \tilde{J}\left(\frac{2T}{\ell^2}, \frac{m^2\ell^2}{4}\right) = \mathcal{N}(\gamma_3) + e^{2m\ell}\mathcal{N}(\gamma_4)$$

with

$$\gamma_3 = \frac{1}{\sigma\sqrt{T}}\left(rT - \ell\sigma - \frac{\sigma^2 T}{2}\right) \;,\; \gamma_4 = -\frac{1}{\sigma\sqrt{T}}\left(rT + \ell\sigma + \frac{\sigma^2 T}{2}\right) \;.$$

9.3.6 Valuation of a DIC Option

Since $S_0 > L$, the value of $\ln\dfrac{L}{S_0}$ is negative. Setting $S_0 = x$, the price of the option $\mathrm{DIC}(S_0, L, K)$ can be written

$$
\begin{aligned}
e^{rT}\mathrm{DIC}(x, L, K) &= E_Q((S_T - K)^+ \mathbb{1}_{T_L < T}) \\
&= E_R\left((xe^{\sigma W_T} - K)^+ \exp\left(mW_T - \frac{m^2 T}{2}\right)\mathbb{1}_{T_\ell < T}\right) \;.
\end{aligned}
$$

The payoff can be decomposed into:

$$
\begin{aligned}
(xe^{\sigma W_T} - K)^+ &= (xe^{\sigma W_T} - K)\mathbb{1}_{\{xe^{\sigma W_T} - K \geq 0\}} \\
&= xe^{\sigma W_T}\mathbb{1}_{W_T \geq k} - K\mathbb{1}_{W_T \geq k}
\end{aligned}
$$

where $k = \dfrac{1}{\sigma}\ln\dfrac{K}{x}$. Substituting this expression into that for the DIC, we obtain

$$
\begin{aligned}
e^{rT}\mathrm{DIC}(L, K) &= \exp\left(-\frac{m^2 T}{2}\right)\left[x\, E_R\left(e^{(\sigma + m)W_T}\mathbb{1}_{W_T \geq k}\mathbb{1}_{\{T_\ell < T\}}\right)\right. \\
&\qquad \left. - K E_R\left(e^{mW_T}\mathbb{1}_{\{W_T \geq k\}}\mathbb{1}_{T_\ell < T}\right)\right] \\
&= \exp\left(-\frac{m^2 T}{2}\right)\left[x\Psi(\sigma + m) - K\Psi(m)\right]
\end{aligned}
$$

where $\Psi(y) = E_R\left(e^{yW_T}\mathbb{1}_{W_T \geq k}\mathbb{1}_{T_\ell < T}\right)$. To evaluate the last expression above, we use the reflection principle for Brownian motion and the elementary fact that

$$\{T_\ell < T\} = \{m_T \overset{\text{def}}{=} \inf_{s \leq T} W_s \leq \ell\} \;.$$

- Hence, in the case $k \leq \ell$ (that is $K \leq L$) we deduce that

$$\Psi(y) = \int_{-\infty}^{\infty} e^{yu} \, \mathbb{1}_{u \geq k} R(W_T \in du, T_\ell < T)$$

$$= \frac{1}{\sqrt{2\pi T}} \left[\int_k^\ell \exp\left(yx - \frac{1}{2T}x^2\right) dx \right.$$

$$\left. + \int_\ell^{\infty} \exp\left(yx - \frac{1}{2T}(2\ell - x)^2\right) dx \right]$$

$$= \exp\left[\frac{Ty^2}{2}\right] [\mathcal{N}(y_1) - \mathcal{N}(y_2)] + \exp\left[\frac{Ty^2}{2} + 2y\ell\right] \mathcal{N}(y_3) \,,$$

where the last equality makes use of some basic calculations, and where we have set

$$y_1 = \frac{1}{\sqrt{T}} [\ell - yT] \,, \quad y_2 = \frac{1}{\sqrt{T}} [k - yT] \,, \quad y_3 = \frac{1}{\sqrt{T}} [y + \ell] \,.$$

Once the computations are done, we find that for $K \leq L$

$$\text{DIC}(L, K) = S_0 \left[\mathcal{N}(z_1) - \mathcal{N}(z_2) + \left(\frac{L}{S_0}\right)^{\frac{2r}{\sigma^2} + 1} \mathcal{N}(z_3) \right]$$

$$- Ke^{-rT} \left[\mathcal{N}(z_4) - \mathcal{N}(z_5) + \left(\frac{L}{S_0}\right)^{\frac{2r}{\sigma^2} - 1} \mathcal{N}(z_6) \right]$$

where

$$z_1 = \frac{1}{\sigma\sqrt{T}} \left[(r + \frac{1}{2}\sigma^2)T + \ln(x/K)\right] \qquad z_4 = z_1 - \sigma\sqrt{T}$$

$$z_2 = \frac{1}{\sigma\sqrt{T}} \left[(r + \frac{1}{2}\sigma^2)T + \ln(x/L)\right] \qquad z_5 = z_2 - \sigma\sqrt{T}$$

$$z_3 = \frac{1}{\sigma\sqrt{T}} \left[(r + \frac{1}{2}\sigma^2)T - \ln(x/L)\right] \qquad z_6 = z_3 - \sigma\sqrt{T} \,.$$

- In the case $K \geq L$, we find that

$$\text{DIC}(L, K) = S_0 \left(\frac{L}{S_0}\right)^{\frac{2r}{\sigma^2} + 1} \mathcal{N}(z_7) - Ke^{-rT} \left(\frac{L}{S_0}\right)^{\frac{2r}{\sigma^2} - 1} \mathcal{N}(z_8)$$

where

$$z_7 = \frac{1}{\sigma\sqrt{T}} \left[\ln(L^2/S_0 K) + (r + \frac{1}{2}\sigma^2)T\right]$$

$$z_8 = z_7 - \sigma\sqrt{T} \,.$$

The formula can also be written as

$$\mathrm{DIC}(S_0, L, K) = \left(\frac{L}{S_0}\right)^{1 + \frac{2r}{\sigma^2}} \left[S_0 \mathcal{N}(z_7) - \frac{K S_0^2}{L^2} \mathcal{N}(z_8)\right] ,$$

or, if we use the notation $C(x, K)$ for the price of a European call with strike K on an underlying with initial value x, as

$$\mathrm{DIC}(S_0, L, K) = \left(\frac{L}{S_0}\right)^{1 + \frac{2r}{\sigma^2}} C\left(S_0, \frac{K S_0^2}{L^2}\right) .$$

This expression enables us to give a strategy for hedging barrier options with European calls.

9.3.7 Up-and-In Options

Their value is given by $\mathrm{UIC}(S_0, K, L) = \mathrm{e}^{-rT} E_Q[(S_T - K)^+ \mathbb{1}_{T_H < T}]$ with $T_H = \inf\{t : S_t \geq H\}$, in the case $S_0 < H$.

• Before we go into the calculations, we note that if $K \geq H$, then it follows from the equality $\{S_T > K\} \cap \{T_H < T\} = \{S_T > K\}$ that the option price is the Black–Scholes price.

• In the case $K \leq H$, we proceed as we did previously, and establish that

$$E\left[(S_T - K)^+ \mathbb{1}_{T_H < T}\right] = \exp\left(-\frac{m^2 T}{2}\right) [S_0 \Phi(m + \sigma) - K \Phi(m)]$$

where, setting $h = \frac{1}{\sigma} \ln \frac{H}{x}$,

$$\Phi(y) = \int \mathrm{e}^{yu} \mathbb{1}_{u \geq k} R\left(W_T \in \mathrm{d}u, T_h < T\right)$$

$$= \frac{1}{\sqrt{2\pi T}} \left[\int_k^h \exp\left(yx - \frac{1}{2T}(2h - x)^2\right) \mathrm{d}x \right.$$

$$\left. + \int_h^\infty \exp\left(yx - \frac{1}{2T}x^2\right) \mathrm{d}x\right]$$

$$= \exp\left[\frac{Ty^2}{2}\right] [\mathcal{N}(a_1) - \mathcal{N}(a_2)] + \exp\left[\frac{Ty^2}{2} + 2yh\right] \mathcal{N}(a_3) ,$$

and where the numbers a_i depend on the various parameters involved. Once we are done, we find that for $K \leq H$,

$$\text{UIC}(S_0, L, K) = S_0 \left[\left(\frac{H}{S_0} \right)^{\frac{2r}{\sigma^2} + 1} (\mathcal{N}(b_1) - \mathcal{N}(b_2)) + \mathcal{N}(b_3) \right]$$

$$- K\mathrm{e}^{-rT} \left[\left(\frac{H}{S_0} \right)^{\frac{2r}{\sigma^2} - 1} (\mathcal{N}(b_4) - \mathcal{N}(b_5)) + \mathcal{N}(b_6) \right]$$

with

$$b_1 = \frac{1}{\sigma\sqrt{T}} \left[(r + \sigma^2/2)T - \ln(Kx/H^2) \right] \qquad b_4 = b_1 - \sigma\sqrt{T}$$

$$b_2 = \frac{1}{\sigma\sqrt{T}} \left[(r + \sigma^2/2)T + \ln(H/x) \right] \qquad b_5 = b_2 - \sigma\sqrt{T}$$

$$b_3 = \frac{1}{\sigma\sqrt{T}} \left[(r + \sigma^2/2)T - \ln(H/x) \right] \qquad b_6 = b_3 - \sigma\sqrt{T} .$$

The case in which coefficients depend on time in a deterministic way is much more complicated, and no explicit formula is known. The problem involves studying the distribution of (B_T, T^*) with $T^* = \inf\{t, B_t \geq g(t)\}$ where g is a deterministic function. The calculation is straightforward if g is an affine function, but few other cases yield explicit solutions. Roberts and Shortland [311] transform the problem by conditioning on the final value of the Brownian motion, and writing

$$E_Q \left(\Phi(S_T) \mathbb{1}_{T_L \leq T} \right) = E_Q \left(E_Q \left(\Phi(S_T) \mathbb{1}_{T_L \leq T} | B_T \right) \right) .$$

This comes down to considering a Brownian bridge, and yields numerical procedures for obtaining an approximation.

9.3.8 P. Carr 's Symmetry

The formula obtained in the case of a DIC option when $K > L$ is particularly simple, and can in fact be obtained by another approach, which involves less mathematics, and uses the symmetry formula of Peter Carr [47]. The formula (see the following proposition) can be explained in a very intuitive way in the case of the foreign exchange markets, when the exchange rate has the dynamics

$$\mathrm{d}X_t = X_t[(r^d - r^f)\mathrm{d}t + \sigma\mathrm{d}W_t]$$

where r^d is the rate in the domestic country, and r^f is the rate in the foreign country.

Let us consider a call on the foreign currency, with maturity T and an exercise price of K euro. The option guarantees the purchase of \$1, at a maximum rate of K. Its price in domestic currency terms at time t is written

$\mathrm{Call}^d(t, X_t, K, T, r^d, r^f)$.

For the foreign investor, the same option guarantees the sale of K euro at time T at a maximum price of \$1. Thus we have K foreign puts on the euro-dollar rate $Y_t = X_t^{-1}$, with strike K^{-1}, and a price at time t of $K\mathrm{Put}^f(t, X_t^{-1}, K^{-1}, T, r^f, r^d)$.

The absence of arbitrage implies that there is a symmetry between these two markets (domestic and foreign):

$$\mathrm{Call}^d(t, X_t, K, T, r^d, r^f) = KX_t\mathrm{Put}^f(t, X_t^{-1}, K^{-1}, T, r^f, r^d) \qquad (9.8)$$

In the special case where $r^d = r^f$, the homogeneity of the price of a put

$$a\mathrm{Put}(t, x, K, T) = \mathrm{Put}(t, ax, aK, T) ,$$

leads to the following symmetry formula. (We have used the notation $\mathrm{Put}(t, x, y, T)$ for the price of a put with strike y: the strike is given as the third variable, and the price of the underlying as the second.)

Proposition 9.3.1. *If the underlying asset has the dynamics* $\mathrm{d}S_t = S_t\sigma\mathrm{d}W_t$ *under the risk-neutral measure, then we have the relationship*

$$\mathrm{Call}(t, S_t, K, T) = KS_t\mathrm{Put}(t, S_t^{-1}, K^{-1}, T) = \mathrm{Put}(t, K, S_t, T) ,$$

which is called the symmetry formula.

Proof. We give a proof that uses Girsanov's Theorem. The price of a call is

$$\mathrm{Call}(t, S_t, K, T) = E_Q((S_T - K)^+|\mathcal{F}_t) .$$

By taking out $S_T K$ as a factor within the expectation, we get

$$\mathrm{Call}(t, S_t, K, T) = KE_Q(S_T(K^{-1} - S_T^{-1})^+|\mathcal{F}_t) .$$

Under Q, we can write $S_t = xM_t$ where M is the martingale $\mathrm{d}M_t = M_t\sigma\mathrm{d}W_t$, $M_0 = 1$; moreover M^{-1} has the dynamics $\mathrm{d}M_t^{-1} = -\sigma M_t^{-1}(\mathrm{d}B_t - \sigma\mathrm{d}t)$. Next, $E_Q(S_T(K^{-1} - S_T^{-1})^+|\mathcal{F}_t)$ can be evaluated by changing measure (and numéraire), and introducing the density

$$\left.\frac{\mathrm{d}R}{\mathrm{d}Q}\right|_{\mathcal{F}_t} = M_t = \exp\left(\sigma B_t - \frac{\sigma^2}{2}t\right) .$$

Under the new measure R, the process $\tilde{B}_t = B_t - \sigma t$ is a Brownian motion, and the dynamics of $Z = M^{-1}$ are given by $\mathrm{d}Z_t = -\sigma Z_t\mathrm{d}\tilde{B}_t = \sigma Z_t\mathrm{d}\overline{W}_t$, where $\overline{W}$ is an R-Brownian motion. Using Bayes' formula, $M_t E_R(A|\mathcal{F}_t) = E_Q(AM_T|\mathcal{F}_t)$, we obtain

$$\mathrm{Call}(t, S_t, K, T) = KxM_t\, E_R\left(\left(K^{-1} - \frac{Z_T}{x}\right)^{+} |\mathcal{F}_t\right) .$$

The term $E_R\left((K^{-1} - \frac{Z_T}{x})^{+}|\mathcal{F}_t\right)$ corresponds to the price of a put with strike $1/K$, on an asset Z following the process $\mathrm{d}Z_t = -\sigma Z_t \mathrm{d}\overline{W}_t$ and with initial value $1/x$. The last equality in the statement of the proposition follows from the homogeneity of the put. □

In the case where $K > L$, and for an underlying that is a martingale under the risk-neutral measure,

$$E((S_T - K)^{+}\mathbb{1}_{T_L < T}) = E(\mathbb{1}_{T_L < T} E((S_T - K)^{+})|\mathcal{F}_{T_L}) .$$

The term $E((S_T - K)^{+}|\mathcal{F}_{T_L})$ corresponds to the value of a European call with strike K, initial value of the underlying L, and maturity $T - T_L$. Using the symmetry formula, it is equal to $\mathrm{Put}(T_L, K, L, T)$. From the homogeneity of the put,

$$E\left[\mathbb{1}_{T_L < T} E((S_T - K)^{+}|\mathcal{F}_{T_L})\right] = E\left[\mathbb{1}_{T_L < T} \frac{K}{L}\,\mathrm{Put}\left(T_L, L, \frac{L^2}{K}, T\right)\right] .$$

At maturity, the put with strike price L^2/K is worth $\left(\frac{L^2}{K} - S_T\right)^{+}$, and only has strictly positive value if the underlying is below $\frac{L^2}{K}$, which in turn is smaller than L, as $K > L$. If the payoff is non-zero, it must be the case that the barrier L has been reached between times 0 and T, and it follows that

$$\frac{K}{L} E\left[\mathbb{1}_{T_L < T}\mathrm{Put}\left(T_L, L, \frac{L^2}{K}, T\right)\right] = \frac{K}{L} E\left[\left(\frac{L^2}{K} - S_T\right)^{+}\right]$$

$$= \frac{K}{L}\,\mathrm{Put}\left(x, \frac{L^2}{K}\right)$$

It remains to employ Carr's symmetry formula once again to obtain

$$\mathrm{DIC}(x, K, L) = \mathrm{Call}\left(L, \frac{Kx}{L}\right) .$$

By differentiating this expression with respect to K, we obtain

$$Q(S_T > K, \min S_t \le L) = \frac{x}{L} Q\left(S_T > \frac{Kx^2}{L^2}\right) .$$

Indeed, $\mathrm{Call}\left(L, \frac{Kx}{L}\right) = E\left(\left(S_T \frac{L}{x} - \frac{Kx}{L}\right)^{+}\right)$, and we know that

$$\frac{\mathrm{d}}{\mathrm{d}K} E_Q\left((S_T - K)^{+}\right) = -K Q(S_T > K) .$$

The General Case

In the Black–Scholes case, we revert to the martingale case by using the following result.

Proposition 9.3.2. *Let X have the dynamics*

$$\mathrm{d}X_t = X_t(r\mathrm{dt} + \sigma \mathrm{d}B_t)$$

and let $\gamma^{-1} \stackrel{def}{=} 1 - \dfrac{2r}{\sigma^2}$. Then, $X_t = X_0(M_t)^\gamma$ where M is such that $\mathrm{d}M_t = M_t\gamma^{-1}\sigma\mathrm{d}B_t$.

Proof. The proof is a direct application of Itô's formula, from which we deduce:

$$Q(S_T > K, \min_{t \leq T} S_t \leq L) = Q\left(M_T > \left(\frac{K}{S_0}\right)^{1/\gamma}, \min_{t \leq T} M_t \leq \left(\frac{L}{S_0}\right)^{1/\gamma}\right)$$

$$= \left(\frac{x}{L}\right)^{1/\gamma} Q\left(M_T > \left(\frac{Kx}{L^2}\right)^{1/\gamma}\right)$$

$$= \left(\frac{x}{L}\right)^{1/\gamma} Q\left(S_T > \frac{Kx^2}{L^2}\right).$$

By integrating the last equality with respect to K, we obtain:

$$\int_K^\infty \mathrm{d}k\, Q\left(S_T > k, \min_{t \leq T} S_t \leq L\right) = E\left((S_T - K)^+ \mathbb{1}_{T_H \leq T}\right)$$

$$= \left(\frac{x}{L}\right)^{1/\gamma} E_Q\left(\left(S_T\frac{L^2}{x^2} - K\right)^+\right).$$

Hence we obtain the same formula as before. □

9.4 Double Barriers

The reader can also refer to the papers Kunimoto and Ikeda [248] and Geman and Yor [173].

The payoff of a double barrier option is $(S_T - K)^+$ if the underlying remains within the range $[L, H]$ between time 0 and the maturity of the option. The option's price is therefore $E_Q(\mathrm{e}^{-rT}(S_T - K)^+ \mathbb{1}_{T^* > T})$ where $T^* \stackrel{def}{=} T_H \wedge T_L$. Instead of evaluating $E_Q(\mathrm{e}^{-rT}(S_T - K)^+ \mathbb{1}_{T^* > T})$, we calculate $E_Q(\mathrm{e}^{-rT}(S_T - K)^+ \mathbb{1}_{T^* < T})$, as the sum of these two terms is $E_Q(\mathrm{e}^{-rT}(S_T - K)^+)$, which is known from the Black–Scholes formula. We can use the same change of measure as before, and put this expression in the form

$$\mathrm{e}^{-\left(r+\frac{m^2}{2}\right)T} E_R\left((x\mathrm{e}^{\sigma W_T} - K)^+ \mathrm{e}^{mW_T}\mathbb{1}_{T^*<T}\right) \ .$$

Evaluating this expression requires explicit knowledge of the distribution of the pair (W_T, T^*). The distribution is not a simple one, as it is given by the double series in (9.5).

We can use a different approach (Geman–Yor). We can evaluate the Laplace transform of $\Sigma(t) = E_R\left[\mathrm{e}^{mW_t}(x\mathrm{e}^{\sigma W_t} - K)^+ \mathbb{1}_{T^*<t}\right]$. Thus, using the Markov property, we evaluate

$$\Phi(\lambda) = \int_0^\infty \exp\left[-\frac{\lambda^2 t}{2}\right] \Sigma(t)\,\mathrm{d}t = E_R\left(\int_{T^*}^\infty \exp\left[-\frac{\lambda^2 t}{2}\right] \varphi(W_t)\,\mathrm{d}t\right)$$

$$= E_R\left(\exp\left[-\frac{\lambda^2 T^*}{2}\right]\int_0^\infty \exp\left[-\frac{\lambda^2 t}{2}\right] \varphi(\tilde{W}_t + W_{T^*})\,\mathrm{d}t\right)$$

where $\varphi(y) = \mathrm{e}^{my}(x\mathrm{e}^{\sigma y} - K)^+$ and where $\tilde{W}$ is a Brownian motion that is independent of W. We can simplify the expression for the expectation in the last line, by separating it into two and using hitting times:

$$\Phi(\lambda) = E_R\left[\exp\left[-\frac{\lambda^2 T^*}{2}\right]\mathbb{1}_{T^*=T_h}\right]\Psi(h) + E_R\left[\exp\left[-\frac{\lambda^2 T^*}{2}\right]\mathbb{1}_{T^*=T_\ell}\right]\Psi(\ell)$$

where

$$\Psi(y) = E_y\left[\int_0^\infty \exp\left[-\frac{\lambda^2 t}{2}\right]\varphi(\widetilde{W}_t)\,\mathrm{d}t\right]\ ,$$

and where the notation E_y specifies that the Brownian motion $\widetilde{W}$ starts from y. The values of the expressions of the form $E_R\left[\exp\left[-\frac{\lambda^2 T^*}{2}\right]\mathbb{1}_{T^*=T_h}\right]$ were given previously, in the proof of Proposition 9.1.3.

- Let $K \in [L, H]$. Using the results on the Brownian motion's resolvent operator, and for values of λ such that $m + \sigma - \lambda < 0$, we get

$$\Psi(h) = \frac{\mathrm{e}^{-\lambda h}}{\lambda}\Big[x(\Psi_h(\sigma + m + \lambda)$$
$$- \Psi_k(\sigma + m + \lambda)) - K[\Psi_h(m + \lambda) - \Psi_k(m + \lambda)]\Big]$$
$$+ \frac{\mathrm{e}^{\lambda h}}{\lambda}\left[K\Psi_h(m - \lambda) - x\Psi_h(\sigma + m - \lambda)\right]$$

$$\Psi(\ell) = \frac{\mathrm{e}^{\lambda \ell}}{\lambda}\left[K\Psi_k(m - \lambda) - x\Psi_k(\sigma + m - \lambda)\right]$$

with $\Psi_a(u) = \frac{1}{u}\mathrm{e}^{ua}$.

- For $K < L$, taking $z = \ell$ or $z = k$, we get

$$\Psi(z) = \frac{\mathrm{e}^{-\lambda h}}{\lambda} \Big[x(\Psi_h(\sigma + m + \lambda) - \Psi_k(\sigma + m + \lambda))$$
$$-K[\Psi_h(m + \lambda) - \Psi_k(m + \lambda)] \Big]$$
$$+ \frac{\mathrm{e}^{\lambda h}}{\lambda} [K\Psi_h(m - \lambda) - x\Psi_h(\sigma + m - \lambda)] .$$

It then remains to invert the Laplace Transform.

9.5 Lookback Options

A *standard lookback* option pays out $S_T - m_0^T$ where $m_0^T = \inf\{S_t, t \in [0, T]\}$. The amount $S_T - m_0^T$ is positive. The price of this option can be expressed as

$$\mathrm{Look} = \mathrm{e}^{-rT} E_Q(S_T - m_0^T)$$

and its price at an interim time t is Look_t where

$$\mathrm{Look}_t \mathrm{e}^{-rt} = \mathrm{e}^{-rT} E_Q \left(S_T - m_0^T | \mathcal{F}_t \right) .$$

We note that $m_0^T = m_0^t \wedge m_t^T$, with $m_s^t = \inf\{S_u, u \in [s, t]\}$, which allows us to write

$$\mathrm{Look}_t \mathrm{e}^{-rt} = E_Q(\mathrm{e}^{-rT} S_T | \mathcal{F}_t) - \mathrm{e}^{-rT} E_Q(m_0^t \wedge m_t^T | \mathcal{F}_t) .$$

- The first term equals $\mathrm{e}^{-rt} S_t$, as discounted prices are Q-martingales.

- To evaluate the second term, we decompose the expectation into two parts.

$$E_Q \left(m_0^t \wedge m_t^T | \mathcal{F}_t \right) = E_Q \left(m_0^t \mathbb{1}_{m_0^t < m_t^T} | \mathcal{F}_t \right) + E \left(m_t^T \mathbb{1}_{m_t^T < m_0^t} | \mathcal{F}_t \right) .$$

  – The first term above can be calculated using the fact that for $u > t$ the price process can be written

$$S_u = S_t \exp \left(r(u - t) + \sigma(B_u - B_t) - \frac{\sigma^2}{2}(u - t) \right)$$
$$= S_t \exp \left(r(u - t) + \sigma \tilde{B}_{u-t} - \frac{\sigma^2}{2}(u - t) \right) \stackrel{\text{def}}{=} S_t Z_{t-u}$$

  where $\tilde{B}$ is independent of $\mathcal{F}_t$, and where $Z_u \stackrel{\text{def}}{=} \exp \left(ru + \sigma \tilde{B}_u - \frac{\sigma^2}{2} u \right)$. Now $E_Q(m_0^t \mathbb{1}_{m_0^t < m_t^T} | \mathcal{F}_t) = m_0^t \Phi(m_0^t, S_t)$ where $\Phi(m, x) = Q(m < x m_{T-t})$ with $m_{T-t} = \inf\{Z_u, u \in [0, T - t]\}$. An explicit expression can be obtained for Φ either by using the results on the distribution of

the minimum for a Brownian motion with drift, or by using the results proved in the section on barrier options.

$$\Phi(m,x) = \mathcal{N}\left(\frac{\ln\dfrac{x}{m} + (r - \sigma^2/2)(T - t)}{\sigma\sqrt{T-t}}\right)$$

$$- \exp\left[-(1 - \frac{2r}{\sigma^2})\ln\frac{m}{x}\right]\mathcal{N}\left(\frac{\ln\dfrac{m}{x} + (r - \sigma^2/2)(T - t)}{\sigma\sqrt{T-t}}\right),$$

which by setting

$$d = \frac{\ln\dfrac{x}{m} + (r + \sigma^2/2)(T - t)}{\sigma\sqrt{T-t}},$$

becomes

$$\Phi(m,x) = \mathcal{N}(d - \sigma\sqrt{T-t}) - \left(\frac{x}{m}\right)^{1 - \frac{2r}{\sigma^2}}\mathcal{N}(-d + \frac{2r}{\sigma}\sqrt{T-t}).$$

– It remains to evaluate the second term

$$E(m_t^T \mathbb{1}_{m_t^T < m_0^t}|\mathcal{F}_t)$$

which can be written as $\Psi(m_0^t, S_t)$ where

$$\Psi(m,x) = E_Q(xm_{T-t}\mathbb{1}_{xm_{T-t}<m}).$$

Once again, the relevant computations were given in the section on barrier options.

Proposition 9.5.1. *The price of a lookback option is*

$$\text{Look}_t = S_t\mathcal{N}(d_t) - \mathrm{e}^{-r(T-t)}m_0^t\mathcal{N}\left(d_t - \sigma\sqrt{T-t}\right)$$

$$+ \mathrm{e}^{-r(T-t)}\frac{S_t\sigma^2}{2r}\left[\left(\frac{S_t}{m_0^t}\right)^{-\frac{2r}{\sigma^2}}\mathcal{N}(-d_t + \frac{2r\sqrt{T-t}}{\sigma}) - \mathrm{e}^{rT}\mathcal{N}(-d_t)\right]$$

where
$$d_t = \frac{1}{\sigma\sqrt{T-t}}\ln\left(\frac{S_t}{m_0^t} + r(T-t) + \frac{1}{2}\sigma^2(T-t)\right).$$

Further lookback options are to be found in Conze and Viswanathan [65, 66].

9.6 Other Options

We give examples of a few more options that are traded in the markets. We do not go into the details of the calculations, which are often tedious. Nor do we deal with the crucial issue of hedging, which is even more complex than that of valuation.

9.6.1 Options Linked to the Hitting Time of a Barrier

Digital Options

- *Asset-or-nothing* options are linked to an "exercise price" K. The final payoff is equal to the price of the underlying if the underlying is "in the money" at maturity, i.e., equal to $S_T \mathbb{1}_{S_T \geq K}$, and to 0 otherwise. The "exercise price" plays the role of a barrier. The value of such an option is $\mathrm{e}^{-rT} E_Q(S_T \mathbb{1}_{S_T \geq K})$, and is easily obtained (it is the first term in the Black–Scholes price).

- *Digital* options (or binary options or cash-or-nothing options) are associated with a barrier. The payoff of an up-and-out digital option is 1 if the underlying does not go through the barrier before maturity, and 0 otherwise. The price of this option is $\mathrm{e}^{-rT} E_Q(\mathbb{1}_{T_L > T}) = \mathrm{e}^{-rT} Q(T_L > T)$. The distribution of T_L under Q is the distribution of the hitting time for the barrier $\ell = \dfrac{1}{\sigma} \ln \dfrac{L}{S_0}$ by the Brownian motion with drift $\dfrac{r}{\sigma}t + B_t$. The relevant workings were carried out when we evaluated the compensation from barrier options.

- Asset-or-nothing options can also have an up-and-in feature linked to a barrier. Their price is then given by $\mathrm{e}^{-rT} E_Q(S_T \mathbb{1}_{S_T \geq K} \mathbb{1}_{T_L} > T)$. These options are used to build hedging portfolios for barrier options.

Forward-Start Barrier Options

For this type of option, the barrier is only put into place at time $t < T$ where T is the maturity. The payoff is $(S_T - K)^+ \mathbb{1}_{T_H^t \geq T}$ where $T_H^t = \inf\{u \geq t : S_u \geq H\}$. Meanwhile, *early-ending* options have a barrier that is only active between the emission of the contract, and a time t.

Boost Options

A boost option is associated with two barriers: an upper barrier H and a lower barrier L. The payoff of the option is proportional to the amount of time spent between the barriers, before the first exit from the range.

The price of a boost option can be calculated using the Laplace Transform of $T^* = T_{H,L} \wedge T$, where $T_{H,L}$ is the first exit time from the range, and where $T_{H,L} \overset{\text{def}}{=} T_H \wedge T_L$. This Laplace transform is defined by $\Psi(\lambda) = E_Q(e^{-\lambda T^*})$. Knowledge of the transform will give us the solution we are after, as differentiating the transform with respect to λ leads to $-\Psi'(\lambda) = E_Q(T^* e^{-\lambda T^*})$.

To evaluate Ψ, we can split $E_Q(e^{-\lambda T^*})$ into two parts:

$$E_Q(e^{-\lambda T^*}) = E_Q(e^{-\lambda T_{H,L}} \mathbb{1}_{T_{H,L} < T}) + e^{-\lambda T} Q(T_{H,L} > T) .$$

We know the distribution of $T_{H,L}$: it is given by a double series.

- If payments are made before maturity, the price of a boost option is, up to a coefficient of proportionality, $E_Q(e^{-rT} T^*) = -e^{-rT} \Psi'(0)$.
- If payments are made "at hit", the price is $E_Q(e^{-rT^*} T^*) = -\Psi'(r)$.

9.6.2 Options Linked to Occupation Times

Cumulative Options

These options become worthless if the underlying spends more than a certain amount of time (specified in the contract) above a barrier. If L is the barrier and D is the maximum duration of time allowed above the barrier, the payoff is given by $(S_T - K)^+ \mathbb{1}_{A_T \leq D}$ where $A_T = \int_0^T \mathbb{1}_{S_t \geq L} dt$ is a measure of the amount of time that the underlying spends above the barrier L. The problem is then solved by evaluating the distribution of the pair (A_T, S_T). This requires calculations that we cannot reproduce here. The reader can refer to Chesney et al. [55, 56] and Hugonnier [199].

Cumulative–Boost Options

Unilateral cumulative boost options have a payoff that is proportional to the amount of time spent above the barrier, that is, proportional to $\int_0^T \mathbb{1}_{S_t \geq L} dt$. Their value is then

$$\text{BC} = E_Q \left(e^{-rT} \int_0^T \mathbb{1}_{S_t \geq L} dt \right) ,$$

which can be easily evaluated under the form

$$\text{BC} = e^{-rT} \int_0^T Q(S_t \geq L) dt$$

where, as for the Black–Scholes formula, the term $Q(S_t \geq L)$ can be expressed using the cumulative normal distribution.

Step Options

These options were introduced by Linetzky [255] in 1997, and have the payoff $e^{-\nu A_T^L}(S_T - K)^+$ where $A_T^L = \int_0^T \mathbb{1}_{S_t \leq L} dt$ and where ν is a positive coefficient that is agreed upon, along with the other parameters K and L, at the signing of the contract. This payoff is smaller that the payoff of a European option with identical strike, with equality occurring when the underlying remains under the barrier throughout the life of the option. Linetzky prices these options by obtaining the distribution of the pair (S_T, A_T^L). We can also transform the expression for the price by using the Brownian motion W and the distribution of the pair $\left(W_T, \int_0^T \mathbb{1}_{W_t \leq \ell} \, dt \right)$, which appears in Borodin–Salminen [39].

Quantile Options

These were introduced by Akahori [3] in 1995, and have the payoff $(A_T^\alpha - K)^+$ where

$$A_T^\alpha = \inf\left\{ x \in \mathbb{R} : \int_0^T \mathbb{1}_{S_t \leq x} dt \geq \alpha T \right\} \, .$$

Parisian Options

These options bear similarities to the cumulative options, but here time does not cumulate. The option becomes worthless if the underlying remains over a certain level L for an interval of time of length D. This option is much harder to price. The first step is to write the payoff in a mathematical form. To do this, we introduce the last instant before time t when the underlying reaches level L, written $g_t = \sup\{s \leq t | S_s = L\}$, and the stopping time H, at which the option disappears, given by

$$H = \inf\{t | (t - g_t) \geq D, S_t \leq L\} \, .$$

In the expression above, $t - g_t \geq D$ means that between times g_t and t, the underlying does not take the value L, and the inequality $S_t \leq L$ specifies that at the instant t (i.e., between times g_t and t) the value of the underlying is below the level L. The value of the option is then

$$E_Q\left(e^{-rT}(S_T - K)^+ \mathbb{1}_{H \geq T} \right) \, .$$

To evaluate this expression, we need to know the distribution of the pair (S_T, H), which is not easy to obtain. We calculate the Laplace transform of the price with respect to time (Chesney et al. [55, 56] and Yor et al. [376]), i.e.,

$$\int_0^\infty dt \, e^{-\lambda t} E_Q(e^{-rt}(S_t - K)^+ \mathbb{1}_{H \geq t}) \, .$$

In fact, it is easier to calculate

$$\int_0^\infty \mathrm{d}t \, \mathrm{e}^{-\lambda t} \, E_Q(\mathrm{e}^{-rt}(S_t - K)^+ \mathbb{1}_{H \le t}))$$

$$= \int_0^\infty \mathrm{d}t \, \mathrm{e}^{-\lambda t} \, E_Q(\mathrm{e}^{-rt}(S_t - K)^+)) - \int_0^\infty \mathrm{d}t \, \mathrm{e}^{-\lambda t} \, E_Q(\mathrm{e}^{-rt}(S_t - K)^+ \mathbb{1}_{H \ge t})) \, .$$

The first term of the line above can be given explicitly, using the Black–Scholes formula. This technique allows us to apply the Markov property at time H.

9.7 Other Products

9.7.1 Asian Options or Average Rate Options

These options have a fixed maturity T at which the final payoff is $\left(\dfrac{1}{T} \displaystyle\int_0^T S_u \, \mathrm{d}u - K \right)^+$. Exact pricing formulae are known, but to present them, we would need material that is beyond the scope of this book. Instead, we refer the reader Geman–Yor [171, 172], which make intensive use of the Bessel process. The idea is to evaluate the Laplace transform with respect to time of the price, rather than the price itself. That is:

$$\int_0^\infty \mathrm{e}^{-\lambda t} E_Q \left[\frac{1}{t} \int_0^t S_u \, \mathrm{d}u - K \right]^+ \mathrm{d}t \, .$$

It is possible to calculate this, using the result, owed to Lamperti[1], that expresses S_t as a function of a Bessel process evaluated at a time other than t (the time-change formula).

   Another approach, developed by Stanton [344], Rogers and Shi [313] and Alziary, Decamps and Koehl [7], involves writing down the valuation PDE. The value of an Asian option is given by $C_t^{As} = X(t)\mathcal{A}(t, Y_t)$ where

$$Y_t \stackrel{\mathrm{def}}{=} \frac{1}{X(t)} \left(\frac{1}{T} \int_0^t X(u) \, \mathrm{d}u - K \right) \, ,$$

and where $\mathcal{A}$ solves

$$\frac{\partial \mathcal{A}}{\partial t} + \left(\frac{1}{T} - ry \right) \frac{\partial \mathcal{A}}{\partial y} + \frac{1}{2} \sigma^2 y^2 \frac{\partial^2 \mathcal{A}}{\partial y^2} = 0$$

with the boundary condition $\mathcal{A}(T, y) = y^+$.

   The markets trade simpler products, which are based on the arithmetic average $\left(\text{payoff } \dfrac{1}{n} \displaystyle\sum_{j=1}^n S_{\frac{jT}{n}} \right)$ or on the geometric average $\left(\text{payoff } \Pi_{j=1}^n S_{\frac{jT}{n}}^{1/n} \right)$.

---

[1] See Revuz–Yor [307].

9.7.2 Products Depending on an Interim Date

For the following products, we consider the maturity T and a fixed date t_1 such that $t_1 < T$.

Compound Options

They are also called options on options. We distinguish the underlying option and the dependent option. At time 0, the buyer purchases an option with maturity t_1 and strike K_1, on an option with maturity $T > t_1$ and strike K. The dependent option can be priced by noting that its payoff at time t_1 is, for a call on a call, $(C(t_1, K, T) - K_1)^+$.

Chooser Options

The buyer of the option can decide at a fixed time $t_1 < T$ on the nature of the product that he has bought: is it a call or a put? By using put–call parity, it is easy to show that the payoff of this option, in the case where the put and call have identical strikes and maturities, is given by

$$\max(C(t_1), P(t_1)) = C(t_1) + (Ke^{-r(T-t_1)} - S_{t_1})^+$$

and it is then easy to valuate the option.

Cliquet Options

Their payoff is $\max(S_T - K, S_{t_1} - K, 0)$. It would be possible to have cliquet options involving several interim dates.

Bermudan Options

These are mid-way between European and American options, hence their name. The holder of such an option can exercise his right before maturity, but only at certain predetermined dates.

9.7.3 Still More Products

In this section, we content ourselves with giving some product definitions. Our list is by no means exhaustive, and new products will no doubt have appeared by the time this book is on the shelves.

Quanto Options

Quanto options involve two countries and the exchange rate between their currencies. These options have been studied by Chérif [54]. Their valuation is based on the principle that a foreign asset becomes a domestic asset once its price is transcribed into the domestic currency.

Take for example a call on a foreign asset with a strike given in the foreign currency. The payoff is $(S_T^f - K^f)^+$ where we use f to denote prices in the foreign country. This payoff is converted into the domestic currency, by using the exchange rate X, and can then be valued using the domestic risk-neutral measure, which is indexed by d. This leads us to evaluate $E^d(X_T(S_T^f - K^f)^+)$.

Another approach involves valuating the product in the foreign currency, using the foreign risk-neutral measure, and then transcribing its value into the domestic currency using the rate of exchange: we get $X_t E^f((S_T^f - K^f)^+)$. The two approaches are identical according to the assumption of no-arbitrage. Thus the risk-neutral measures of the two countries are linked.

Russian Options

These are American-style options. If they are exercised at time τ their payoff is $Z_\tau \stackrel{\text{def}}{=} K \vee \max_{t \le \tau} S_t$. The point is to determine τ^* such that it optimizes $E(Z_\tau e^{r(\tau-t)})$.

Rainbow Options

They are based on two underlyings. Their payoff is $\max(S_1(T), S_2(T), K)$.

Notes

The reader can consult the following articles: Rubinstein and Reiner [319], Bowie and Carr [42], Rich [308], Heynen and Kat[190], Carr and Chou [48].

Recent studies have been carried out when the underlying process is a general Lévy process. In this case, the market is incomplete and the valuation is done under a particular risk-neutral measure. The reader can refer to Shiryaev [336] and to the collective book [295], which also give wide choice of references.

We also mention here some recent books on exotic options and derivative products. The two Deutsche Bank volumes [46, 295] present the practitioner's view point. Jarrow and Turnbull, [214], (1996), Kat [235], (2001), Haug [180], (1998), and Hull [200], (2000), contain an extensive study of various options. Lipton [257], (2001), on the other hand, is devoted to the financial engineer's approach. The mathematics of derivatives can be found in Hunt and Kennedy [203], (2000), and Kallianpur and Karandikar [228], (1999).

ANNEX 9

1 The Laplace Transformation

Let E be the set of real-valued functions f defined on $I\!R$ such that:

(i) $f(x) = 0$, $\forall x < 0$,

(ii) there exists $a > 0$ such that $\int_0^a |f(x)| \mathrm{d}x < \infty$,

(iii) $\exists \lambda_0$, $|f(x)| \mathrm{e}^{-\lambda_0 x} \to 0$ when $x \to \infty$.

If $f \in E$, the integral $\mathcal{L}(f)(\lambda) = \int_0^\infty f(x) \mathrm{e}^{-\lambda x} \mathrm{d}x$ exists for all $\lambda > \lambda_0$. We introduce the notation

$$\sigma(f) = \inf\{\lambda \in I\!R : \lim_{x \to \infty} f(x) \mathrm{e}^{-\lambda x} = 0\}.$$

Examples

- Let Y be the Heaviside function given by $Y(t) = 0, t < 0$ and $Y(t) = 1, t \geq 0$. Then $\mathcal{L}(Y)(\lambda) = \dfrac{1}{\lambda}$.
- Let $f \in E$, $a \in I\!R$ and let $_a f$ be defined by $_a f(t) = f(t) \mathrm{e}^{at}$. Then $\mathcal{L}(_a f)(\lambda) = \mathcal{L}(f)(\lambda - a)$.
- Let $f \in E, a > 0$ and let f_a be defined by $f_a(t) = f(t-a)$. Then $\mathcal{L}(f_a)(\lambda) = \mathrm{e}^{-a\lambda} \mathcal{L}(f)(\lambda)$.
- Let $f(t) = Y(t) t^n \mathrm{e}^{at}$. Then $\mathcal{L}(f)(\lambda) = \dfrac{n!}{(\lambda - a)^{n+1}}$.
- An important example in probability is the hitting time of a Brownian motion.
  Let $a > 0$ and $f(t) = \dfrac{a}{\sqrt{2\pi t^3}} \exp{-\dfrac{a^2}{2t}}$. Then $\mathcal{L}(f)(\lambda) = \exp{-(a\sqrt{2\lambda})}$.

Properties

- Let $f \in E$. Then $\lim_{\lambda \to \infty} \mathcal{L}(f)(\lambda) = 0$.
- If $f_n(t) = (-1)^n t^n f(t)$, then $\mathcal{L}(f_n)(\lambda) = [\mathcal{L}(f)]^{(n)}(\lambda)$.
- If f is continuous to the right at 0 and if f' exists, is continuous on $]0, \infty[$ and belongs to E, then $\mathcal{L}(f')(\lambda) = \lambda \mathcal{L}(f)(\lambda) - f(0)$.
- If $f, g \in E$ and $h(x) = \int_0^x f(x - t) g(t) \mathrm{d}t$, then $\mathcal{L}(h) = \mathcal{L}(f)\mathcal{L}(g)$.
- The Laplace transformation is injective.

2 The Optional Stopping Theorem

Let $(M_t, t \geq 0)$ be a martingale and let T be a finite stopping time. The process $(M^T = (M_{t \wedge T}; t \geq 0)$ is a martingale that is called the martingale stopped at T.

If M is a uniformly integrable martingale and S, T are two stopping times with $S \leq T$, then

$$M_S = E(M_T | \mathcal{F}_S) = E(M_\infty | \mathcal{F}_S), \text{ a.s.}$$

In particular, if M is a bounded martingale, then

$$M_S = E(M_T | \mathcal{F}_S) = E(M_\infty | \mathcal{F}_S), \text{ a.s.}$$

for any pair of stopping times $S \leq T$.

This theorem often serves as a basic tool to determine quantities defined up to a first hitting times and laws of hitting times. However, in many cases, the u.i. hypothesis has to be checked carefully. For example, if W is a Brownian motion and T_a the first hitting time of a, then $E(W_{T_a}) = a$, while a blind application of Doob's theorem would lead to the equality between $E(W_{T_a})$ and $W_0 = 0$. The process $(W_{t \wedge T_a}, t \geq 0)$ is not uniformly integrable.

A

Brownian Motion

A.1 Historical Background

The botanist Robert Brown, in 1828, observed the irregular movements of particles of pollen suspended in water. In 1877, Delsaux explained the ceaseless changes of direction in the particles' paths by the collisions between the particles of pollen and the water molecules. A motion of this type was described as being a "random motion".

In 1900, Bachelier [18], with a view to studying price movements on the Paris exchange, exhibited the "Markovian" nature of Brownian motion: the position of a particle at time $t + s$ depends on its position at time t, and does not depend on its position before time t. It is worth emphasizing that Bachelier was a forerunner in the field, and that the theory of Brownian motion was developed for the financial markets before it was developed for physics.

In 1905, Einstein [134] determined the transition density function of Brownian motion by means of the heat equation, and so linked Brownian motion to partial differential equations of the parabolic type. That same year, Smoluchowski described Brownian motion as a limit of random walks.

The first rigorous mathematical study of Brownian motion was carried out by N. Wiener [366] (1923), who also gave a proof of the existence of Brownian motion. P. Lévy [254], (1948) worked on the finer properties of Brownian paths, without any knowledge of concepts such as filtration or stopping time. Since then, Brownian motion has continued to fascinate probabilists, for the study of its paths just as much as for that of stochastic integration theory. See for example the books Knight [238] and Yor [373, 374, 375].

A.2 Intuition

The easiest Brownian motion to imagine is probably Brownian motion in the plane: at each instant in time, the particle randomly chooses a direction,

and then makes a "step" in that direction. However, for an approach that is both intuitive and rigorous, we must study the real Brownian motion: at each instant that is a multiple of Δt, the particle "randomly" chooses to move left or right to a distance Δx from its starting point. To model this "randomness", we turn to a sequence of independent identically distributed random variables $(Y_i,\ i \geq 1)$ such that $P(Y_i = \Delta x) = P(Y_i = -\Delta x) = \dfrac{1}{2}$.

At time t, the particle will have made $\left[\dfrac{t}{\Delta t}\right]$ moves (where $[a]$ denotes the integer part of a). The particle's position will be $V_t = Y_1 + Y_2 + \cdots + Y_{[\frac{t}{\Delta t}]}$.

All this takes place on a very small scale: we would like to let both Δt and Δx tend to zero in an appropriate way. Note that $EV_t^2 \simeq (\Delta x)^2 \dfrac{t}{\Delta t}$. In order for this quantity to have a limit, we must impose that $\dfrac{(\Delta x)^2}{\Delta t}$ have a limit. The increment Δt will be "very small" and Δx will be "small", so that $(\Delta x)^2$ will also be "very small". The most straightforward choice is $\Delta x = \sqrt{\Delta t}$ and $\Delta t = \dfrac{1}{n}$.

Let us now give a precise formulation of this approach.

A.3 Random Walk

On a probability space $(\Omega, \mathcal{F}, P)$, let

$$P(X_i = 1)\ =\ P(X_i = -1)\ =\ \frac{1}{2}\quad,\quad i \in \mathbb{N}^*,$$

be a family of independent identically distributed random variables (the X_i are said to be independent Bernoulli variables). To this family, we associate the sequence $(S_n,\ n \geq 0)$ defined by

$$S_0\ =\ 0$$
$$S_n\ =\ \sum_{i=1}^{n} X_i\ . \tag{A.1}$$

We have $E(S_n) = 0$, $\mathrm{Var}\,(S_n) = n$. We say that the sequence S_n is a *random walk*. We can interpret it as a game of tossing a coin: a player tosses a coin, he wins one euro if it comes up tails, and loses one euro if it comes up heads. He has no initial wealth ($S_0 = 0$). His wealth at **time** n (after n games) is S_n. We represent the series of the results obtained over N successive games as a graph (see Fig. A.1).

We note that the sequence $(S_m - S_n, m \geq n)$ is independent of $(S_0, S_1, \ldots, S_n)$ and that $S_m - S_n$ has the same probability law as S_{m-n} (the binomial distribution depends only on $m - n$).

We now proceed with a two-fold normalization. Let N be fixed.

Fig. A.1. A random walk
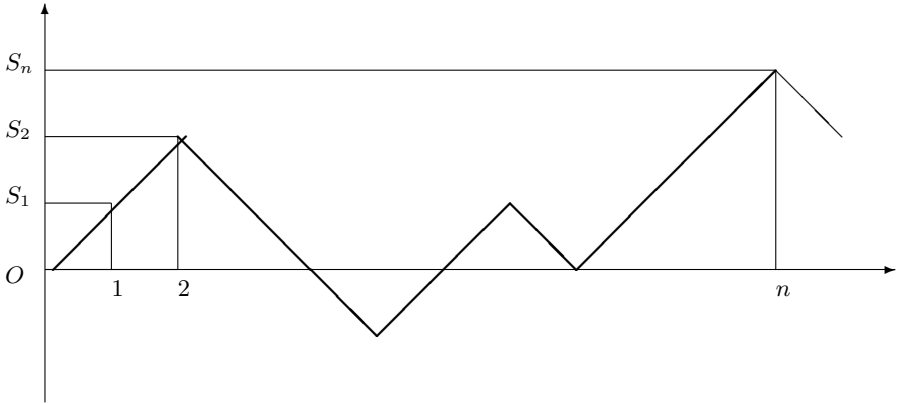
- We transform the time interval $[0, N]$ into the interval $[0, 1]$,
- and we change the scale of values taken by S_n.

More precisely, we define a family of random variables indexed by real numbers of the form $\frac{k}{N}$, $k \in \mathbb{N}$:

$$U_{\frac{k}{N}} \;=\; \frac{1}{\sqrt{N}}\, S_k \;. \tag{A.2}$$

We move from $U_{\frac{k}{N}}$ to $U_{\frac{k+1}{N}}$ in a "very small" interval of time equal to $\frac{1}{N}$, by making a step of a "small length" $\frac{1}{\sqrt{N}}$ (towards the left or towards the right). We have

$$E\big(U_{\frac{k}{N}}\big) \;=\; 0 \quad \text{and} \quad \mathrm{Var}\big(U_{\frac{k}{N}}\big) = \frac{k}{N} \;.$$

The independence and stationarity properties of the random walk still hold, i.e.,

- if $k \geq k'$, $U_{\frac{k}{N}} - U_{\frac{k'}{N}}$ is independent of $U_{\frac{p}{N}}$ for $p \leq k'$;
- if $k \geq k'$, $U_{\frac{k}{N}} - U_{\frac{k'}{N}}$ has the same probability law as $U_{\frac{k-k'}{N}}$.

We define a continuous-time process, that is, a family of random variables $(U_t, t \geq 0)$ starting from $U_{\frac{k}{N}}$, by requiring the function $t \to U_t$ to be affine between times $\frac{k}{N}$ and $\frac{k+1}{N}$. To do this, for N fixed, we note that for all $t \in \mathbb{R}_+$, there exists a unique $k(t) \in \mathbb{N}$ such that $\frac{k(t)}{N} \leq t < \frac{k(t)+1}{N}$, and we set

$$U_t^N = U_{\frac{k}{N}} + N \left(t - \frac{k}{N}\right) \left(U_{\frac{k+1}{N}} - U_{\frac{k}{N}}\right)$$

where $k = k(t)$.

(The process $(U_t, t \geq 0)$ does not have independent increments. However, if $t \geq t'$ and $\frac{k'+1}{N} > t' \geq \frac{k'}{N}$, we have that

$$U_t^N - U_{t'}^N \quad \text{is independent of} \quad U_{\frac{p}{N}}^N, \; p \leq k' \;).$$

Let us return for a brief moment to writing U as a function of the random walk S.

For $t = 1$ we have $U_1^N = \frac{1}{\sqrt{N}} S_N$. The central limit theorem then implies that U_1^N converges in distribution to a standard normal random variable.

Exercise A.3.1.

1. Show that U_t^N converges in distribution to a normal random variable with mean 0 and variance t as $N \to \infty$. Notice how $0 \leq t - \frac{k(t)}{N} \leq \frac{1}{N}$, and how

$$\left|U_{\frac{k+1}{N}}^N - U_{\frac{k}{N}}^N\right| \leq \frac{1}{\sqrt{N}} \qquad \text{with} \quad k = k(t).$$

2. Show that $(U_{t_1}^N, U_{t_2}^N, \ldots, U_{t_n}^N)$ with $t_1 < t_2 < \cdots < t_n$ converges in distribution to a vector $\left(Z_{t_1}, Z_{t_2}, \ldots, Z_{t_n}\right)$, such that $Z_{t_i} - Z_{t_{i-1}}$ is independent of $\left(Z_{t_1}, Z_{t_2}, \ldots, Z_{t_n}\right)$ and such that $(Z_{t_i} - Z_{t_{i-1}})$ has a normal distribution with mean 0 and variance $(t_i - t_{i-1})$ (use the central limit theorem for vectors).

It can be shown that U^N converges[1] to a process B, which has continuous paths (i.e., for almost all ω, the mapping $t \to B_t(\omega)$ is continuous), and which satisfies

(i) $B_0 = 0$.

---

[1] In the sense of convergence in distribution. This is stronger than the convergence in distribution of finite families. See Karatzas and Shreve [233]. It is also possible to construct a probability space on which all the random walks S_N are defined, and on which the normalized sums U^N converge a.s. to a Brownian motion (Knight [238]).

(ii) $B_{t+s} - B_t$ has the normal distribution $\mathcal{N}(0, s)$.

(iii) $B_{t+s} - B_t$ is independent of $B_{t_i} - B_{t_{i+1}}$, for $t_0 < \cdots < t_n = t$.

Remark A.3.2. We can show that Brownian motion is the only process satisfying (i), (iii) and

(ii)$'$ The distribution of $B_{t+s} - B_t$ depends only on s.

We introduce the notation $\Delta B(t) = B(t + \Delta t) - B(t)$ where $B(t) = B_t$ and $\Delta t > 0$. The Brownian motion then satisfies:

- $E[\Delta B(t)] = 0$ $\quad\quad$ $\mathrm{Var}[\Delta B(t)] = \Delta t$ $\quad\quad$ (using (ii))
- $E_t[\Delta B(t)] = 0$ $\quad\quad$ $E_t[(\Delta B(t))^2] = \Delta t$ $\quad\quad$ (using (ii) and (iii))

where E_t is the conditional expectation with respect to $\mathcal{F}_t = \sigma(B_s, \, s \leq t)$. The equality $E_t(\Delta B(t)) = 0$ can be interpreted as follows: if the position of the Brownian motion at time t is known, then the average move between times t and $t + \Delta t$ is zero. This property is a result of the independence and of the Gaussian nature of Brownian motion.

A.4 The Stochastic Integral

Brownian motion represents the path of a particle that incessantly changes direction. The graph of such a path has many sharp peaks and troughs, and is not differentiable at these points (the left and right derivatives are not equal). We can prove the following result:

Theorem A.4.1. *For almost all ω, the function $t \mapsto B_t(\omega)$ is a.s. nowhere differentiable (i.e., the set of t for which $B_t(\omega)$ is differentiable has Lebesgue measure zero).*

The lack of differentiability of Brownian paths forbids **the interpretation of the symbol** $\mathrm{d}B_t$ as B'_t, and makes it impossible to define $\int \theta(t) \, \mathrm{d}B_t$ using the usual methods (such as writing $\mathrm{d}B_t = B'(t)\mathrm{d}t$).

As Brownian motion has unbounded variation, the Stieljes integration theory cannot be applied. However, we can draw on the ideas of Riemann integration theory, as long as we carefully check each step along the way. The aim is to define a new integral, in such a way that it is additive with respect to θ, and satisfies

$$\int_{[a,b]} \mathrm{d}B_t = B(b) - B(a) \, .$$

Hence the idea of defining the integral for a step function[2] θ (i.e., such that $\theta(t) = \theta(t_i)$, $t \in]t_i, t_{i+1}]$, $t_0 = 0 < t_1 \cdots < t_p = T$), as:

$$\int_0^T \theta(s) \, \mathrm{d}B_s \; = \; \sum_{i=0}^{p-1} \theta(t_i)[B(t_{i+1}) - B(t_i)] \; .$$

When θ is a process, we impose conditions of measurability, which are slightly stronger than assuming the process to be adapted to the Brownian motion's filtration. For technical reasons, we also impose integrability conditions on the process θ, in order for $\sum \theta(t_i)(B(t_{i+1}) - B(t_i))$ to converge when the time-step tends to 0 (we approximate the process θ with a step process).

As seen in Chaps. 2 to 4, we are led to study Itô processes, i.e., processes X of the form

$$X_t \; = \; x + \int_0^t \mu(s) \, \mathrm{d}s + \int_0^t \sigma(s) \, \mathrm{d}B_s \; . \tag{A.3}$$

It is important to understand that the notation

$$\mathrm{d}X_t \; = \; \mu(t) \, \mathrm{d}t + \sigma(t) \, \mathrm{d}B_t \; , \tag{A.4}$$

is only a symbolic notation, with which we can develop a stochastic calculus. The exact meaning of (A.4) is given by writing X in the form of (A.3).

Still working symbolically, and interpreting $\mathrm{d}B_t$ (and $\mathrm{d}X_t$) as small increments ΔB_t of B (or ΔX_t), we obtain

$$E(\mathrm{d}X_t) = \mu(t)\mathrm{d}t, \quad \mathrm{Var}(\mathrm{d}X_t) = \sigma^2(t)\mathrm{d}t$$

and similarly,

$$\begin{aligned} E_t(\mathrm{d}X_t) \; &= \; \mu(t) \, \mathrm{d}t \\ \mathrm{Var}_t \, (\mathrm{d}X_t) \; &= \; E_t[\mathrm{d}X_t - E_t(\mathrm{d}X_t)]^2 \; = \; \sigma^2(t) \, \mathrm{d}t \; . \end{aligned}$$

(Exact calculations would lead to $E_t(\Delta X_t) \; = \; \int_t^{t+\Delta t} \mu(s) \, \mathrm{d}s$ and $E_t\left(\Delta X_t - \int_t^{t+\Delta t} \mu(s)\mathrm{d}s\right)^2 = \int_t^{t+\Delta t} \sigma^2(s)\mathrm{d}s$.)

It is worth emphasizing that whilst $\mathrm{d}t$ and $\mathrm{d}B_t$ are both "small", their sizes are of different orders. Indeed, we have "$E(\mathrm{d}B_t) = 0$" and "$E(\mathrm{d}B_t)^2 = \mathrm{d}t$". This symbolic representation of (A.4) has an advantage: it suggests that if we apply Taylor's expansion to a function of X_t, and if we wish to keep the $\mathrm{d}t$ terms, we will need to include terms from the expansion of "$(\mathrm{d}B_t)^2$", which will have a role to play.

---

[2] Taking functions that are left-continuous. This is a small technical difficulty that we do not dwell upon here.

A.5 Itô's Formula

Using this very intuitive approach to Itô processes, we can persuade ourselves (and persuade the reader) that Itô's lemma is "quite natural". Let

$$\mathrm{d}X_t \;=\; \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t\;, \tag{A.5}$$

be an Itô process, and let f be a function of class C^2. We can apply Taylor's expansion to f :

$$f(X_{t+\Delta t}) - f(X_t) \;=\; (X_{t+\Delta t} - X_t)\,f'(X_t) + \frac{1}{2}(X_{t+\Delta t} - X_t)^2 f''(X_t)$$
$$+\, o(X_{t+\Delta t} - X_t)^2\;.$$

Setting $\Delta X_t = X_{t+\Delta t} - X_t$ and identifying ΔX_t with $\mathrm{d}X_t$ as we did before, we obtain from the expression for ΔX_t given in (A.5)

$$\Delta f(X_t) \;=\; \mu(t, X_t)f'(X_t)\Delta t + \sigma(t, X_t)f'(X_t)\Delta B_t$$
$$+ \frac{1}{2}\{\mu^2(t, X_t)(\Delta t)^2 + \sigma^2(t, X_t)(\Delta B_t)^2$$
$$+ 2\mu(t, X_t)\sigma(t, X_t)\Delta t\,\Delta B_t\}f''(X_t) + o(\Delta X_t)^2\;.$$

We saw above that the $(\Delta B_t)^2$ term is "of the same order as" Δt. Therefore, we must keep it in this form. However the $(\Delta t)^2$ and $(\Delta t)\,(\Delta B_t)$ terms are $o(\Delta t)$. It is appropriate to keep only terms of order lesser than or equal to that of Δt. We obtain

$$\Delta f(X_t) \;=\; \mu(t, X_t)f'(X_t)\Delta t + \frac{1}{2}\sigma^2(t, X_t)f''(X_t)\,\Delta t$$
$$+ \sigma(t, X_t)f'(X_t)\,\Delta B_t\;.$$

Table A.1. Multiplication Table 1

	$\mathrm{d}t$	$\mathrm{d}B$
$\mathrm{d}t$	0	0
$\mathrm{d}B$	0	$\mathrm{d}t$

We remark that a rigorous proof of Itô's lemma rests on the same idea.

Furthermore, we note that this intuitive approach to Itô's formula makes it easy to write down: we take Taylor's expansion of order 2, and use the "multiplication table" in Table A.1.

A similar technique can be used to move up the case of a multi-dimensional Brownian motion. If B^1 and B^2 are two independent Brownian motions, $\Delta B_t^1 \; \Delta B_t^2$ has zero expectation, so we neglect these terms in the Taylor expansion. This leads to the multiplication table in Table A.2.

Table A.2. Multiplication Table 2

	$\mathrm{d}t$	$\mathrm{d}B_t^1$	$\mathrm{d}B_t^2$
$\mathrm{d}t$	0	0	0
$\mathrm{d}B_t^1$	0	$\mathrm{d}t$	0
$\mathrm{d}B_t^2$	0	0	$\mathrm{d}t$

Example A.5.1.

$$\mathrm{d}X_t^1 \;=\; \mu_1 \,\mathrm{d}t + \sigma_1 \,\mathrm{d}B_t^1$$
$$\mathrm{d}X_t^2 \;=\; \mu_2 \,\mathrm{d}t + \sigma_2 \,\mathrm{d}B_t^2$$
$$\mathrm{d}(X_t^1 X_t^2) \;=\; X_t^1 \,\mathrm{d}X_t^2 + X_t^2 \,\mathrm{d}X_t^1 + \mathrm{d}X_t^1 \,\mathrm{d}X_t^2$$
$$\;=\; X_t^1 \,\mathrm{d}X_t^2 + X_t^2 \,\mathrm{d}X_t^1 + \sigma_1 \sigma_2 \,\mathrm{d}t \;.$$

B

Numerical Methods

We present here several methods for approximating solutions to partial differential equations (PDEs) of the parabolic type that are analogous to those appearing in the Black–Scholes model. We have chosen to give our exposition of these methods in a simple case, assuming the coefficients to be constant, for example. In this case, we know an explicit solution to the Black–Scholes equation, and numerical methods are of little interest. However, we hope to show which are the difficulties that arise, and to make it easier for the reader to access specialist works on the subject such as Cessenat et al. [50], Kloeden and Platen [237], Dupuis and Kushner [131], and Rogers and Talay [314].

To simplify the exposition, we assume that the market includes one riskless bond whose price is given by

$$\mathrm{d}S_t^0 \;=\; S_t^0\, r(t, S_t)\, \mathrm{d}t \;,$$

and a stock whose price satisfies

$$\mathrm{d}S_t \;=\; b(t, S_t)\, \mathrm{d}t + \sigma(t, S_t)\, \mathrm{d}B_t \;,$$

where B is a real-valued Brownian motion.

We showed in Chap. 3 (Sect. 3.4) that to calculate the value of the contingent product $g(S_T)$, we need to solve the following PDE:

$$\begin{cases} \mathcal{L}C(t, x) - r(t, x)C(t, x) \;=\; 0 \\ C(T, x) \;=\; g(x) \end{cases} \tag{B.1}$$

with

$$\mathcal{L}C(t, x) = xr(t, x)\frac{\partial C}{\partial x}(t, x) + \frac{1}{2}\sigma^2(t, x)\frac{\partial^2 C}{\partial x^2}(t, x) + \frac{\partial C}{\partial t}(t, x) \;.$$

This solution can be written as

$$C(t,x) \;=\; E\left(\mathrm{e}^{-\phi(T)}g(W_T^{x,t})\right) \tag{B.2}$$

with

$$\phi(s) \;=\; \int_t^s r(u, W_u^{x,t})\,\mathrm{d}u \ .$$

In this formulation, $W_u^{x,t}$ denotes the solution to the following stochastic differential equation

$$\begin{cases} \mathrm{d}W_u^{x,t} \;=\; \mu(u, W_u^{x,t})\mathrm{d}u + \sigma(u, W_u^{x,t})\mathrm{d}B_u \ , \\ W_t^{x,t} \;=\; x \ , \end{cases}$$

where we have set $\mu(t,x) = x\,r(t,x)$.

We present two methods for approximating C, the first one using (B.1) and techniques for approximating solutions to parabolic equations, and the second using (B.2) and simulating the process W.

B.1 Finite Difference

We are going to use the fact that C is the unique solution to the partial differential equation (B.1) satisfying conditions of regularity.

Let us give an example of regularity conditions in a particular case. Let us assume that $r(t,x) = r$ and $\sigma(t,x) = x\sigma$. In this case, we solve (B.1) on $[0,T]\times]0,\infty[$. We then have the following result (Karatzas et al. [233]): if Δ is continuous on $[0,T]\times]0,\infty[$, Hölder continuous in x uniformly with respect to (t,x) on a compact set, if g is continuous, and if Δ and g satisfy

$$\begin{cases} |g(x)| \;\leq\; K(1 + x^\alpha + x^{-\alpha}) \\ \max_{0\leq t\leq T} |\Delta(t,x)| \;\leq\; K(1 + x^\alpha + x^{-\alpha}) \qquad 0 < x < \infty \ , \end{cases} \tag{B.3}$$

then equation B.1 has a unique solution in the set of $C^{1,2}([0,T]\times]0,\infty[)$ functions satisfying (B.3).

Note that in this particular case, in order to solve

$$\frac{\partial C}{\partial t} + rx\frac{\partial C}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 C}{\partial x^2} - rC \;=\; \Delta$$

we first make a change of variable, setting $H(t,x) = C(t,\mathrm{e}^x)$. We are thus led to solve $\dfrac{\partial H}{\partial t} + \left(r - \dfrac{1}{2}\sigma^2\right)\dfrac{\partial H}{\partial x} + \dfrac{1}{2}\sigma^2 \dfrac{\partial^2 H}{\partial x^2} - rH \;=\; \Delta$, which has constant coefficients.

Let us suppose therefore that (B.1) **has a unique solution**

The first difficulty is that the domain on which we are studying (B.1) is unbounded. Therefore, let us first solve the problem on $[0, T] \times [-K, +K]$.

To obtain the uniqueness of the solution, boundary conditions are needed. Let us either impose *Dirichlet* conditions: we take $a \in \mathbb{R}$ and impose

$$C(t, K) = C(t, -K) = a \qquad t \in [0, T] , \tag{B.4}$$

or *Neumann* conditions:

$$\frac{\partial C}{\partial x}(t, K) = \frac{\partial C}{\partial x}(t, -K) = b \qquad t \in [0, T] , \tag{B.4bis}$$

with $b \in \mathbb{R}$.

In the case of constant coefficients, $C(t, x)$ can be expressed as a function of the normal distribution. In this case, it is easy to show that

$$C(t, x) \xrightarrow[x \to \infty]{} \infty \qquad t \in [0, T]$$

$$\frac{\partial C}{\partial x}(t, x) \xrightarrow[x \to \infty]{} 1 \qquad t \in [0, T] .$$

In this case, the Neumann conditions with $b = 1$ are best suited to the problem.

B.1.1 Method

We continue our exposition of the method in the cases of Dirichlet and Neumann conditions.

We define a grid on the domain $[0, T] \times [-K, +K]$, with steps of size $h = \dfrac{2}{N+1}$ for the space variable x and of size ε for the time variable t.

We use the notation

$$t_n = n\varepsilon \quad 0 \leq n \leq M \qquad \text{with} \quad M\varepsilon = T ,$$

$$x_i = -K + i\frac{2K}{N+1} \qquad \text{with} \quad 0 \leq i \leq N+1 ;$$

(both the step sizes h and ε will tend to 0.)

The *finite difference* method is a means of obtaining an approximation to the solution, by using the nodes (t_n, x_i) on the grid. Let $C(t, x)$ be the solution to (B.1). We are looking for a family M of vectors $(C^n(i), 1 \leq i \leq N)_{n<M}$ such that $C^n(i)$ is close to $C(t_n, x_i)$ for $i = 1, \ldots, N$ and $n = 0, \ldots, M-1$ (from our choice of the boundary condition in time, we know that $C(t_M, x_i) = g(x_i) := C^M(i)$). If we are working with Dirichlet conditions, we impose

$$C(t_n, x_0) = C(t_n, x_{N+1}) = a \qquad n < M$$

i.e.,

$$C^n(0) \;=\; C^n(N+1) \;=\; a \qquad n < M \;.$$

If we are working with Neumann conditions with $b = 0$, we take

$$C(t_n, x_{N+1}) \;=\; C(t_n, x_N) \qquad \text{and} \qquad C(t_n, x_0) \;=\; C(t_n, x_1)$$

i.e.,

$$C^n(N+1) \;=\; C^n(N) \qquad \text{and} \qquad C^n(0) = C^n(1)$$

(if $b \ne 0$ we take for example $C^n(N+1) = C^n(N) + bh$).

Next, we approximate

$$\frac{\partial C}{\partial t}(t_n, x_i) \qquad\qquad \text{by} \qquad \frac{C^{n+1}(i) - C^n(i)}{\varepsilon} \qquad\qquad \text{(scheme 1)}$$

$$\qquad\qquad\qquad \text{or by} \qquad \frac{C^n(i) - C^{n-1}(i)}{\varepsilon} \qquad\qquad \text{(scheme 2)}$$

$$\frac{\partial C}{\partial x}(t_n, x_i) \qquad\qquad \text{by} \qquad \frac{C^n(i+1) - C^n(i-1)}{2h}$$

$$\frac{\partial^2 C}{\partial x^2}(t_n, x_i) \qquad\qquad \text{by} \qquad \frac{C^n(i+1) - 2C^n(i) + C^n(i-1)}{h^2} \;.$$

Exercise B.1.1. Why have we chosen these approximations?

B.1.2 The Implicit Scheme Case

By substituting the expressions above into the partial differential equation, we obtain in the case of scheme 1 (called the *implicit scheme*)

$$\frac{C^{n+1}(i) - C^n(i)}{\varepsilon} \;=\; -\, \frac{\sigma^2(n,i)}{2} \, \frac{C^n(i+1) - 2C^n(i) + C^n(i-1)}{h^2}$$
$$-\, \mu(n,i) \, \frac{C^n(i+1) - C^n(i-1)}{2h} + r(n,i) C^n(i)$$

where $\sigma(n,i) = \sigma(t_n, x_i)$, $\mu(n,i) = \mu(t_n, x_i)$ and $r(n,i) = r(t_n, x_i)$. Hence C^n can be computed as a function of C^{n+1} (remember that it is C^M rather than C^0 that is known at the outset).

Let us carry through our analysis in the case where r, σ and μ depend only on the space variable x. We can write the previous equation in the matrix form:

$$\frac{1}{\varepsilon}(C^{n+1} - C^n) \;=\; AC^n \qquad \text{where} \qquad C^n = C^n(i) \,,$$

and where the matrix A is a tridiagonal matrix.

In the case of Dirichlet conditions, $C^n(0)$ and $C^n(N+1)$ are known. It remains to determine $(C^n(i) \,,\, 1 \le i \le N)$, where C^n is a vector of $\mathbb{R}^N$. In the case $a = 0$, the matrix A has the form

$$A = \begin{bmatrix} a_1 & b_1 & 0 & 0 \\ c_2 & a_2 & b_2 & 0 \\ 0 & c_3 & a_3 & b_3 \\ & 0 & & \ddots & 0 \\ & & & c_N & a_N \end{bmatrix}$$

with

$$a_i = \frac{\sigma^2(i)}{h^2} + r(i)$$

$$b_i = -\frac{\sigma^2(i)}{2h^2} - \frac{\mu(i)}{2h}$$

$$c_i = -\frac{\sigma^2(i)}{2h^2} + \frac{\mu(i)}{2h} .$$

Exercise B.1.2. Modify A in order to study the case $a \neq 0$.

We obtain C^n as a function of C^{n+1} by solving the system $\frac{1}{\varepsilon}(C^{n+1}-C^n) = AC^n$.

In the case of Neumann conditions, we determine $C(n,i)$ for $1 \leq i \leq N$ from the equalities $C(n,0) = C(n,1)$ and $C(n,N+1) = C(n,N)$. Matrix A is written as

$$A = \begin{bmatrix} \alpha_1 & b_1 & 0 \\ c_2 & a_2 & b_2 \\ & c_3 & a_3 & b_3 \\ & & \cdots & & \cdots & \cdots \\ & & & & c_N & \alpha_N \end{bmatrix}$$

(only the first and last lines have changed) with

$$\alpha_1 = \frac{\sigma^2(1)}{2h^2} + \frac{\mu(1)}{2h} + r(1) \qquad \alpha_N = \frac{\sigma^2(N)}{2h^2} - \frac{\mu(N)}{2h} + r(N) .$$

Scheme 2 seems more straightforward. We can obtain C^{n-1} as a function of C^n using $\frac{1}{\varepsilon}(C^n - C^{n-1}) = AC^n$, so it is no longer necessary to solve a system. This scheme is called *explicit*, but it is not as efficient as scheme 1, for reasons of stability (see for example Ciarlet [60]). Let us return to scheme 1.

B.1.3 Solving the System

Solving the system $\frac{1}{\varepsilon}(C^{n+1} - C^n) = AC^n$ calls on methods for solving the equation

$$(I + \varepsilon A)C^n = C^{n+1}$$

where $I + \varepsilon A$ is a tridiagonal matrix. We can then proceed using the pivot method, which consists in writing $I + \varepsilon A$ as a product of two matrices $L\,U$,

where L is upper triangular and U is lower triangular, and in solving $LU(X) = B$ in two steps:

- solve $LY = B$
- solve $UX = Y$.

We can also solve $(I + \varepsilon A)X = Y$ by iterative methods. These are based on the idea that $I + \varepsilon A$ can be written as $C - D$ where C and D are two matrices, with C being invertible (there are a number of possible decompositions).

We then need to solve $CX = Y + DX$. We construct a sequence of vectors $(U^n, n \geq 1)$ defined by recurrence for any fixed U^0, with

$$U^{n+1} \qquad \text{such that} \qquad CU^{n+1} = Y + DU^n \ .$$

We show that (if the spectral radius of $C^{-1}D$ is smaller than 1) the sequence U^n converges to X, the solution to $(I + \varepsilon A)X = Y$.

B.1.4 Other Schemes

We can use other schemes than schemes 1 and 2. Let us assume that $\mu = r = 0$ and that σ is constant, and describe some of the other possibilities. Scheme 1 is then written

$$\frac{C^{n+1} - C^n}{\varepsilon} = \frac{\sigma^2}{2} A_h C^n$$

where A_h is the operator $[A_h C]_i = -\frac{1}{h^2}\{C(i+1) - 2C(i) + C(i-1)\}$, whose matrix we already know.

We could use the *Crank-Nicholson* scheme:

$$\frac{C^{n+1} - C^n}{\varepsilon} = \frac{\sigma^2}{2} A_h(\theta C^{n+1} + (1 - \theta)C^n)$$

(In the case $\theta = 1$, we are back to the implicit case of scheme 1, and the case $\theta = 0$ returns us to the explicit case).

The choice between the different schemes based on error estimation.

B.2 Extrapolation Methods

Let us assume that $\mu(t, x) = \mu x$ and $\sigma(t, x) = \sigma x$. We have seen how a change of variable can bring us back to an equation with constant coefficients. Assume further that $\mu = 1/2$, $\sigma = 1$ and $r = 0$ for our exposition of the method, and that $\Delta = 0$.

B.2.1 The Heat Equation

We want to approximate the solution of

$$\begin{cases} \dfrac{\partial C}{\partial t} + \dfrac{1}{2}\dfrac{\partial^2 C}{\partial x^2} = 0 \\ C(T,x) = g(x)\,, \end{cases} \tag{B.5}$$

with boundary conditions of either the Dirichlet or Neumann type. This equation is known as the *heat equation*. Note however that the heat equation is usually written as

$$\begin{cases} \dfrac{\partial u}{\partial t} - \dfrac{1}{2}\dfrac{\partial^2 u}{\partial x^2} = 0 \\ u(0,x) = g(x)\,. \end{cases} \quad \left(\text{or in some cases without the coefficient of } \dfrac{1}{2} \right)$$

The two forms are equivalent under a change of the time variable.

B.2.2 Approximations

We start with a semi-discrete approximation, i.e., we discretize only the space variable. Thus we replace (B.5) with

$$\begin{cases} \dfrac{\partial C}{\partial t} + A_h C = 0 \\ C(T,x) = g(x) \end{cases} \tag{B.6}$$

where

$$A_h C = \dfrac{1}{2h^2}\{C(t,x+h) - 2C(t,x) + C(t,x-h)\}$$

and where A_h is a matrix operator. We are led to solve system (B.6) for $x = x_0, x_1, \ldots, x_{N+1}$. We can then apply methods that are specific to differential systems (e.g. Euler, Runge–Kutta).

The exact solution to (B.6) is

$$C(t) = (\exp - A_h t)\, C(0)\,. \tag{B.7}$$

[When the μ, σ and r coefficients depend on t, this formula is no longer valid, but other analogous methods can be employed.]

To approximate the solution to (B.7), we need to approximate

$$C(t + \Delta t) = (\exp(-A_h\,\Delta t))\, C(t)\,,$$

and thus to approximate $\exp(-A_h\,\Delta t)$. We set $A_h = A$ and $\Delta t = \varepsilon$. An approximation of $e^{-\varepsilon A}$ is $(1 + \varepsilon A)^{-1}$ where

$$C(t + \varepsilon) \simeq (1 + \varepsilon A)^{-1} C(t) .$$

We can check that we thus recover the implicit scheme.

Another approximation of $\mathrm{e}^{-\varepsilon A}$ is

$$\mathrm{e}^{-\varepsilon A} \simeq \left(1 + \frac{\varepsilon}{2} A\right)^{-1} \left(1 - \frac{\varepsilon}{2} A\right)$$

(Padé's approximation). This leads to the Crank-Nicholson scheme.

We can also use mixed methods. Expanding the exponential $\mathrm{e}^{-2\varepsilon A}$, we get

$$C(t + 2\varepsilon) \simeq (1 - 2\varepsilon A + 2\varepsilon^2 A^2) C(t) .$$

The implicit scheme leads to an approximate solution γ where

$$\begin{aligned}
\gamma(t + 2\varepsilon) &= (1 + 2\varepsilon A)^{-1} \gamma(t) \\
&\simeq (1 - 2\varepsilon A + 4\varepsilon^2 A^2) \gamma(t) .
\end{aligned}$$

If we apply the implicit scheme twice (to go from t to $t + \varepsilon$, and then from $t + \varepsilon$ to $t + 2\varepsilon$), we obtain

$$\begin{aligned}
\Gamma(t + 2\varepsilon) &= (1 + \varepsilon A)^{-2} \Gamma(t) \\
&\simeq (1 - 2\varepsilon A + 3\varepsilon^2 A^2) \Gamma(t) .
\end{aligned}$$

Hence the approximation of $(1 - 2\varepsilon A + 2\varepsilon^2 A^2) C(t)$, by

$$2\Gamma(t + 2\varepsilon) - \gamma(t + 2\varepsilon) ,$$

which leads to the scheme

$$\begin{cases}
C^{n+1/3} &= (1 + 2\varepsilon A)^{-1} C^n \\
C^{n+2/3} &= (1 + 2\varepsilon A)^{-2} C^n \\
C^{n+1} &= 2C^{n+2/3} - C^{n+1/3} .
\end{cases}$$

B.3 Simulation

In this section, we give a brief overview of simulation methods that can be used to approximate solutions to stochastic differential equations as well as the expectations of random variables.

B.3.1 Simulation of the Uniform Distribution on $[0, 1]$

The probability law of a random variable X that is uniformly distributed on $[0, 1]$ is defined by $P(X \in [a, b]) = b - a$ for $0 \leq a < b \leq 1$. A sequence of *"random numbers"* is a series of random variables $X_1, X_2, \ldots, X_n, \ldots$ that are independent, identically distributed, and have the same distribution as X. We would like to simulate this sequence, i.e., we would like to obtain a deterministic sequence of numbers in $[0, 1]$ which has "the same statistical properties" as the sequence $(X_n)_{n \geq 1}$. We do not dwell on methods for simulating these sequences of random numbers here. Most programming languages provide a "random" procedure for generating random numbers. Another approach is to use low discrepancy sequences.

We refer the interested reader to Bouleau [40], Niederreiter [289] and Ripley [310]. These provide various programming methods, as well as a discussion of the meaning of the expression "the same statistical properties".

B.3.2 Simulation of Discrete Variables

To simulate a random variable X, which can take k values (a_1, $a_2, \ldots, a_k$) with probabilities $P(X = a_i) = p_i$, we can use the random variable

$$Z = a_1 \mathbf{1}_{U < p_1} + a_2 \mathbf{1}_{p_1 \leq U < p_1 + p_2} + \cdots + a_k \mathbf{1}_{p_1 + \cdots + p_{k-1} \leq U \leq 1}$$

where $\mathbf{1}_{\alpha \leq U < \beta}$ is worth 1 if $\alpha \leq U < \beta$, and 0 otherwise, and where U is a uniformly distributed random variable on $[0, 1]$, which can be simulated as outlined above.

B.3.3 Simulation of a Random Variable

Case of a Random Variable with a Continuous Density Function

Let X be a random variable with probability density function f, which is continuous. We denote by $F(x) = \int_{-\infty}^{x} f(t) \mathrm{d}t$ its cumulative distribution function. If f is strictly positive, F has an inverse mapping F^{-1}.

Exercise B.3.1. Show that, whatever the probability density function f, the variable $F(X)$ is uniformly distributed on $[0, 1]$. What is the distribution of $F^{-1}(U)$, if U is uniformly distributed on $[0, 1]$?

Show that if F is the cumulative distribution function of a random variable X (i.e., $F(x) = P(X < x)$) and if $F^-(y) = \inf\{x \,|\, y < F(x)\}$, then X has the same distribution as $F^-(U)$ where U is uniformly distributed on $[0, 1]$.

We can now simulate X by using $F^-(U)$. If we want to simulate $(X_1, X_2, \ldots, X_n)$ where the X_i are independent and identically distributed, we can use $(F^-(U_1), F^-(U_2), \ldots, F^-(U_n))$ where the $U_1, U_2, \ldots, U_n$ are independent random variables that are uniformly distributed on $[0, 1]$.

This method is often long, and requires a subroutine for calculating F^-. Therefore the accept/reject method is often used.

The Accept/Reject Method

Suppose that X is a random variable with a bounded continuous density function f with a compact support $[a, b]$.

Consider a pair of random variables (U, V) that are uniformly distributed on the rectangle $[a, b] \times [0, k]$. When the point with coordinates (U, V) is below the curve of f, we accept it, and set $X = U$. Otherwise, it is rejected and a new point is drawn at random. It is easy to check that the variable X thus defined, has probability density function f.

When the support of f is not contained in a compact set, this method is no longer valid, as there is no uniform distribution on an unbounded interval. We then use another density function g, such that

- the variable with probability density g is easy to simulate,
- $kg(x) \geq f(x)$ for a real constant k.

We then simulate a variable Y with density g and a variable U that is uniformly distributed on $[0, 1]$, and we set $Z = kUg(Y)$.

1. If $Z < f(Y)$ we set $X = Y$.
2. Otherwise, we simulate new Y and U, and go back to 1.

The Gaussian Case

Specific methods apply to this case.

Exercise B.3.2. Let U and V be two independent random variables that are uniformly distributed on $[0, 1]$. Show that

$$X = (-2 \log U)^{1/2} \cos 2\pi U$$
$$Y = (-2 \log U)^{1/2} \sin 2\pi V$$

are independent random variables that have the standard normal distribution $N(0, 1)$.

The exercise immediately yields a simulation method. A normally distributed variable with mean m and variance σ^2 can be written $m + \sigma X$, where X follows the distribution $N(0, 1)$.

Further methods are to be found in Bouleau [40] and Ripley [310].

B.3.4 Simulation of an Expectation

Let X be a random variable. We would like to simulate $E(X)$.

Using the Simulation of X

If we have at our disposal a program for simulating independent random variables with the same distribution as X, we can simulate $E(X)$ by using the law of large numbers, i.e.,

$$E(X) = \lim_n \frac{1}{n} \sum_1^n X_i \ .$$

The same method can be used to simulate $E(\psi(X))$. The stopping criterion, which determines n in such a way as to get a small enough error, is obtained via the Bienaymé–Chebyshev inequality.

When the Density of X is Known

We suppose for the sake of simplicity that the density function f of X has support $[0, 1]$. We need to calculate $\int_0^1 g(x)\mathrm{d}x$ with $g(x) = xf(x)$. The law of large numbers shows that if $x_1, x_2, \ldots, x_n$ is a sequence of numbers that are evenly spread on $[0, 1]$ (i.e., simulating a sequence of random numbers), $\frac{1}{n} \sum_1^n g(x_i) \rightarrow \int_0^1 g(x)\mathrm{d}x$ (the measure $\frac{1}{n} \sum_{i=1}^n \delta_{x_i}$ converges weakly to the Lebesgue measure).

Meanwhile, note that there are sequences $(x_1, \ldots, x_n)$ that converge faster by the method described above than when the x_i are chosen "randomly" and "independently". This is the case with the Van der Corput sequences[1].

B.3.5 Simulation of a Brownian Motion

Random Walks

Brownian motion can be approximated by a random walk[2], i.e., the distribution of B_t can be approximated by the distribution of $\frac{1}{\sqrt{n}}(X_1 + X_2 + \cdots + X_{[nt]}) = S_n$ where the X_i are independent and identically distributed random variables such that $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$, where $[\cdot]$ denotes the integer part of a number. We can then approximate $E(\psi(B_t))$ by $E(\psi(S_n))$ for any continuous bounded function ψ.

---

[1] See Bouleau [40] p. 228.
[2] See Appendix A.

Using Gaussian Variables

Another method involves using normal distributions: if $(X_i,\ i\ \leq n)$ are independent standard Gaussian variables, and if

$$S_0 = 0$$
$$S_{n+1} = S_n + \delta X_n \qquad \text{where} \quad \delta \in \mathbb{R}^+$$

then $(S_0, S_1, \ldots, S_n)$ has the same distribution as $(B_0, B_\delta, \ldots, B_{n\delta})$.

B.3.6 Simulation of Solutions to Stochastic Differential Equations

Let X_t be the solution to the stochastic differential equation

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

where B_t is a d-dimensional Brownian motion.

When μ and σ are constant, the solution is $X_t = X_0 + \exp\left[(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t\right]$, and we can then simply simulate the Brownian motion. In the general case, we need to use approximation methods such as the following.

The Euler Scheme

We discretize the stochastic differential equation above, using a scheme of the form

$$\tilde{X}_{t_{k+1}} = f\left(\tilde{X}_{t_k}, B_{t_{k+1}}, B_{t_k}\right) \qquad k \in \{0, \ldots, N-1\}$$

where the t_k subdivide $[0, T]$ into steps of size $\Delta t = \dfrac{T}{N}$. The simplest scheme is the Euler scheme,

$$\begin{cases} \tilde{X}_{t_{k+1}} = \tilde{X}_{t_k} + \mu(t_k, \tilde{X}_{t_k})(t_{k+1} - t_k) + \sigma(t_k, \tilde{X}_{t_k})(B_{t_{k+1}} - B_{t_k}) \\ \tilde{X}_0 = X_0 \,. \end{cases}$$

We can show (Maruyama [266]) that this scheme converges on quadratic average to the solution of the stochastic differential equation, in the sense that

$$\exists\, C > 0\,, \ \forall\, k \in \{0, 1, \ldots, N-1\} \qquad E\big|X_{t_k} - \tilde{X}_{t_k}^N\big|^2 \ \leq \ C\Delta t\,.$$

Numerous schemes have been introduced to improve the speed of convergence. Moreover, other criteria of convergence can be used, for example convergence in L^p spaces, or a.s. convergence.

The Milshtein Scheme

Let us consider the one dimensional case where the coefficients μ and σ do not depend on time, and where σ is of class C^1. Using Taylor's expansion to approximate $\sigma(X_t)$, we obtain

$$X(t) \simeq \mu(X(0))t + \sigma(X(0))\,(B_t - B_0)$$
$$+ \sigma(X(0))\,\sigma^T X(0))\,\int_0^t \{B(s) - B(0)\}\,\mathrm{d}B(s)\ .$$

The stochastic integral is easy to evaluate (Exercice 3.1.12). This leads us to the Milshtein scheme

$$\overline{X}_{t_{k+1}}\ =\ \overline{X}_{t_k} + \mu(\overline{X}_{t_k})\,(t_{k+1} - t_k)$$
$$+ \sigma(\overline{X}_{t_k})\,(B_{t_{k+1}} - B_{t_k})$$
$$+ \frac{1}{2}\sigma(\overline{X}_{t_k})\sigma^T(\overline{X}_{t_k})\left[B_{t_{k+1}}^2 - B_{t_k}^2 - (t_{k+1} - t_k)\right]\ .$$

We can then show (Milstein [279], Talay [350]) that the scheme converges a.s. and on quadratic average, with greater speed than the Euler scheme.

In higher dimensions than 1, the Milshtein scheme requires restrictions on the matrix σ. The reader can refer to Talay [350], Pardoux and Talay [298] or to the books Kloeden and Platen [237] and Dupuis and Kushner [131].

B.3.7 Calculating $E(f(X_t))$

We would like to give approximations of the term $E(f(X_t))$, when the process X is a solution to a stochastic differential equation

$$\mathrm{d}X_t\ =\ \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t\ .$$

The coefficients μ and σ do not depend on t. Notice that this not a restriction. In the general case, it is enough to consider the process $Y_t = (t, X_t)$ and to write down the SDE satisfied by Y_t.

Calculating the Distribution of X_t

A first method for evaluating $E(f(X_t))$ consists in calculating the distribution of X_t explicitly.

If the coefficients μ and σ are regular, and if $X(0)$ has a density function p_0, then $X(t)$ has a density distribution $p(t, \cdot)$ that solves

$$\frac{\mathrm{d}}{\mathrm{d}t}p\ =\ L^*p$$

$$p(0, \cdot)\ =\ p_0$$

where L^* is the adjoint of $L = \sum b_i(x)\dfrac{\partial}{\partial x_i} + \dfrac{1}{2}\sum a_{i,j}(x)\dfrac{\partial^2}{\partial x_i\,\partial x_j}$ with $a = \sigma\sigma^T$, i.e., $L^*p = \sum \dfrac{\partial^2}{\partial x_i\partial x_j}(a_{i,j}\,p) + \sum \dfrac{\partial}{\partial x_i}(b_i\,p)$. We can try to solve this equation numerically, but it is difficult, particularly in spaces of higher dimensions.

The Euler and Milshtein Scheme

A second method consists in using a scheme (Euler's or Milshtein's) to simulate N independent occurrences $\overline{X}_t$ of X_t, which we denote $\overline{X}_t(\omega_i)$, and in calculating

$$\frac{1}{N}\sum_{i=1}^{N} f(\overline{X}_t(\omega_i)) \,,$$

for a t of the form $\dfrac{kT}{n}$. According to the law of large numbers, this provides an approximation of $E(f(\overline{X}_t))$.

We can then show that

$$|E(f(X_T)) - E(f(\overline{X}_T))| \;\leq\; C(T)\,\frac{T}{n}\,.$$

There also exist (Talay [352]) methods that lead to second order schemes.

The General Case

To approximate expressions of the form

$$E\left(\int_0^T \Delta(X_s)\,\mathrm{d}s + g(X_T)\right),$$

we can use the process

$$Y_t \;=\; \left(\int_0^t \Delta(X_s)\,\mathrm{d}s\,,\, X_t\right),$$

and then write down the stochastic differential equation that it satisfies, and apply the methods covered in the previous subsection and described by Talay ([351] and [352]).

References

[1] Aase, K.K., Øksendal, B., (1988): Admissible investment strategies in continuous trading. Stochastic Processes and their Applications, **30**, 291–301

[2] Aftalion, F., Poncet, P., (1991): Le Matif. Presses Universitaires de France, Paris

[3] Akahori, J. (1995): Some formulae for a new type of path-dependent options. Ann. Appl. Prob., **5**, 383–388

[4] Aliprantis, C.D., Border, K.C. (1999): Infinite Dimmensional Analysis, a Hitchhiker's Guide. Springer, Second edition, Berlin Heidelberg New York Tokyo

[5] Aliprantis, C.D., Brown, D.J. Burkinshaw, O. (1989): Existence and Optimality of Competitive Equilibria. Springer-Verlag, New York-Berlin

[6] Allingham, M. (1991): Existence theorems in the Capital Asset Pricing Model. Econometrica, **59**, (4) 1169–1174

[7] Alziary, B., Decamps, J.-P., Koehl, P. F., (1997) A PDE approach to Asian options: analytical and numerical evidence. Journal of Banking and Finance, **21**, 613–640

[8] Amin, K., Khanna, A. (1994): Convergence of American option values from discrete to continuous time financial models. Mathematical Finance, **4**, 289–304

[9] Ansel, J.P., Stricker, C. (1992): Lois de martingale, densités et décomposition de Föllmer-Schweizer. Annales Inst. Henri Poincaré, **28**, 375–392

[10] Araujo, A., Monteiro, P. (1992): General equilibrium with infinitely many goods. In: Equilibrium and Dynamics, Essays in Honor of David Gale. Macmillan, New York

[11] Araujo, A., Monteiro, P., (1989): Existence without uniform conditions. Journal of Economic Theory, **48**, 416–427

[12] Arrow, K. (1953): Le rôle des valeurs boursières pour la répartition la meilleure des risques. Econométrie, Colloq. Internat. Centre National de

la Recherche Scientifique, **40** (Paris 1952), 41–47 ; Discussion, 47–48. C.N.R.S., Paris

[13] Arrow, K.J., Debreu, G. (1954): Existence of an equilibrium for a competitive economy. Econometrica, **22**, 265–290

[14] Arrow, K.J., Hahn, F.J. (1971): General Competitive Analysis. Holden-Day, San Francisco

[15] Artzner, P., Delbaen, F. (1989): Term structure of interest rates: the martingale approach. Advances in Applied Mathematics, **10**, 95–129

[16] Augros, J.C. (1990): Les Options sur Taux d'Intérêt. Economica, Paris

[17] Avellaneda, M., Levy, A., Parás, A. (1995): Pricing and hedging derivative securities in markets with uncertain volatilities. Appl. Math. Finance, **2** , 73–88

[18] Bachelier, L. (1900): Théorie de la spéculation. Ann. Sci. Ecole Normale Supérieure, **17**, 21–86

[19] Back, K., Pliska, S.R., (1991): On the fundamental theorem of asset pricing with an infinite state space. Journal of Mathematical Economics, **20**, 1–18

[20] Balasko Y. (1988): Fondements de la Théorie de l'Équilbre Général. Economica, Paris

[21] Ball, C., Torous, W. (1983): Bond price dynamics and options. Journal of Financial and Quantitative Analysis, **18**, 517–531

[22] Bardhan, I. (1993): Stochastic income in incomplete markets: hedging and optimal policies. Unpublished Manuscript

[23] Barron, E., Jensen, R. (1990): A stochastic control approach to the pricing of options. Mathematics of Operations Research, **1**, 49–79

[24] Basak, S. (2000): A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk. Journal of Economic Dynamics and Control, **24**, 63–95

[25] Basak, S., Cuoco, D. (1998): An equilibrium model with restricted stock market participation. Review of Financial Studies, **11**, 309–341

[26] Bellamy, N., Jeanblanc, M. (1998): Incomplete markets with jumps. Finance and Stochastics, **4**, 209–222

[27] Bensaid, B., Lesne, J.P., Pagès, H., Scheinkman, J. (1992): Derivative asset pricing with transaction costs. Mathematical Finance, **2**, 63–86

[28] Bensoussan, A. (1984): On the theory of option pricing. Acta Applicandae Mathematicae, **2**, 139–158

[29] Bensoussan, A., Elliott, R. (1995): Attainable claims in a Markov market. Mathematical Finance, **5**, 121–131

[30] Bewley, T.F. (1969): A Theorem on the Existence of Competitive Equilibria in a Market with a Finite Number of Agents and whose Commodity Space is L^∞. Discussion paper n° 6904, Université Catholique de Louvain

[31] Bewley, T.F. (1972): Existence of equilibria in economies with infinitely many commodities. Journal of Economic Theory, **4**, 514–540

[32] Biais, B., Björk, T., Cvitanić, J., El Karoui, N., Jouini, E., Rochet, J.C. (1996): In: Runggaldier W. (ed), Financial Mathematics, Bressanone. Lecture notes in mathematics 1656, Springer

[33] Bingham, N.H. and Kiesel, R. (1998). Risk-neutral Valuation. Springer, Berlin

[34] Björk, T.(1998), Arbitrage Theory in Continuous Time. Oxford University Press, Oxford

[35] Björk, T. (201), A Geometric View of Interest Rate Theory. In: Jouini, E. and Cvitanić, J. and Musiela, M. (eds) Option pricing, Interest rates and risk management, 241-277, Cambridge University Press, Cambridge

[36] Black, F. (1972): Capital market equilibrium with restricted borrowing. Journal of Business, **45**, 444–454

[37] Black, F., Scholes, M. (1973): The pricing of options and corporate liabilities. Journal of Political Economy, **3**, 637–654

[38] Border, K.C. (1985): Fixed Point Theorems with Applications to Economics and Game Theory. Cambridge University Press, Cambridge

[39] Borodin, A.N., Salminen, P. (1997): Handbook of Brownian Motion. Facts and Formulae. Birkhäuser. Basel–Boston–London

[40] Bouleau, N. (1986): Probabilités de l'Ingénieur. Variables Aléatoires et Simulation. Hermann, Paris

[41] Bouleau, N., Lamberton, D. (1989): Residual risks and hedging strategies in Markovian markets. Stochastic Processes and Applications, **33**, 131–150

[42] Bowie, J., Carr, P. (1994): Static Simplicity. Risk Magazine, **7**, 45–49

[43] Breeden, D. (1979): An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics, **7**, 265–296

[44] Brennan, M.J., Schwartz, E. (1979): A continuous time approach to the pricing of bonds. Journal of Banking and Finance, **3**, 133–155

[45] Brigo, D., Mercurio, F. (2001): Interest Rate Models : Theory and Practice. Springer–Verlag, Berlin

[46] Brockhaus, O., Farkas, M., Ferraris, A., Long, D., Overhaus, M. (2000): Equity Derivatives and Market Risk Models. Risk books, London

[47] Carr, P. (1995): European Put Call Symmetry. Unpublished manuscript

[48] Carr P., Chou, A. (1997): Breaking barriers. Risk Magazine, **10**, 12–18

[49] Cass, D. (1984): Competitive Equilibrium in Incomplete Financial Markets. Working paper, Economics Department, University of Pennsylvania

[50] Cessenat, M., Ladanois, G., Lions, P.L., Pardoux, E., Sentis, R. (1989): Méthodes Probabilistes pour les Équations de la Physique. Eyrolles, Paris

[51] Chamberlain, G. (1983): A characterisation of the distributions that imply mean–variance utility functions. Journal of Economics Theory, **29**, 185–201

[52] Chamberlain, G. (1985): Asset Pricing in Multiperiod Securities Markets. Working paper, Economics Department, University of Wisconsin, Madison

[53] Chatelain, M., Stricker, C. (1994): On componentwise and vector stochastic integration. Mathematical Finance, **4**, 57–66

[54] Chérif, T. (1996): Modélisation Stochastique en Finance. Modèles de Taux et Evaluation d'Actifs Financiers Contingents. Thesis, Université Paris VI, Paris

[55] Chesney, M., Cornwall, J., Jeanblanc-Picqué, M., Kentwell, G., Yor, M. (1997): Parisian barrier options : a discussion. Risk Magazine, January, 77–79

[56] Chesney, M., Jeanblanc-Picqué, M., Yor, M. (1997) Brownian excursion and Parisian barrier options. Annals Appl. Prob. **29**, 165–184

[57] Chesney, M., Marois, B., Wojakowski, R. (1995): Les Options de Change. Economica, Paris

[58] Chung, K.L. (1968): A Course in Probability Theory. Academic Press, New York

[59] Chung, K.L., Williams, R. (1983): An Introduction to Stochastic Integration. Birkhäuser, Boston

[60] Ciarlet, P.G. (1988): Une Introduction à l'Analyse Numérique Matricielle et à l'Optimisation. Masson, Paris

[61] Cochrane, J.H. (2001): Asset Pricing, Princeton University Press

[62] Colwell, D., Elliott, R., Kopp, E. (1991): Martingale representation and hedging policies. Stochastic Process and Applications, **38**, 335–345

[63] Constantinides, G. (1986): Capital market equilibrium with transaction Costs. Journal of Political Economy, **94**, 842–862

[64] Cont, R. and Tankov, P.,(2004): Financial modeling with jump processes, Chapman & Hall/CRC.

[65] Conze, A., Viswanathan, R. (1991): Path dependent options: the case of lookback Options. Journal of Finance, **46**, 1893–1907

[66] Conze, A., Viswanathan, R. (1991): European path dependent options: the case of geometric averages. Finance, **12**, 7–22

[67] Courtadon, G. (1982): The pricing of options on default-free bonds. Journal of Financial and Quantitative Analysis, **17**, 75–100

[68] Cox, J., Huang, C. (1988): Optimal consumption and portfolio policies when asset prices follow a diffusion process. Journal of Economic Theory, **49**, 33–83

[69] Cox, J., Ingersoll, J., Ross, S. (1981): The relation between forward and futures prices. Journal of Financial Economics, **9**, 321–341

[70] Cox, J., Ingersoll, J., Ross, S. (1985): An intertemporal general equilibrium model of asset prices. Econometrica, **53**, 363–384

[71] Cox, J., Ross, S., Rubinstein, M. (1979): Option pricing. A simplified approach. Journal of Financial Economics, **7**, 229–263

[72] Cox, J., Rubinstein, M. (1985): Options Markets. Prentice-Hall, Englewood Cliffs

[73] Cuoco, D. (1997): Optimal consumption and equilibrium prices with portfolio constraints and stochastic income. Journal of Economic Theory, **72**, 33–73

[74] Cvitanić, J. (2001): Theory of Portfolio Optimization in Markets with Frictions. In: Jouini, E. and Cvitanić, J. and Musiela, M.(eds) Option pricing, Interest rates and risk management, 577-631, Cambridge University Press, Cambridge

[75] Cvitanić, J., Karatzas, I. (1993): Hedging contingent claims with constrained portfolios. Annals of Applied Probability, **3**, 652–681

[76] Cvitanić, J., Karatzas, I. (1995): On dynamic measures of risk. SIAM Journal of Control and Optimization, **33**, 29–66

[77] Cvitanić, J., Karatzas, I. (1996): Contingent claim valuation and hedging with constrained portfolio. In: Davis, M., Duffie, D., Fleming, W., Shreve, S. (eds) IMA volume in Math, **65** 13–33

[78] Cvitanić, J., Pham, H., Touzi, N. (2000): Super-replication in stochastic volatility models under portfolio constraints. Journal of Applied Probability, **36**, 523–545

[79] Cvitanić, J. and Zapatero, F. (2004): Introduction to Economics and Mathematics of Financial Markets, MIT Press

[80] Dalang, R.C., Morton, A., Willinger, W. (1989): Equivalent martingale measures and no arbitrage in stochastic securities market models. Stochastics and Stochastics Reports, **29**, 185–202

[81] Dana, R.A. (1993): Existence and uniqueness of equilibria when preferences are additively separable. Econometrica, **61**, 4, 953–957

[82] Dana, R.A. (1993): Existence, uniqueness and determinacy of Arrow–Debreu equilibria in finance models. Journal of Mathematical Economics, **22**, 563–579

[83] Dana, R.A. (1999): Existence, Uniqueness and determinacy of equilibrium in CAPM with a riskless asset. Journal of Mathematical Economics, **32**, 167–175

[84] Dana, R.A. (2001): Risk aversion and uniqueness of equilibrium. An application to financial markets. Review of Economic Design, **6**, 155–173

[85] Dana, R.A., Pontier, M. (1992): On the existence of a stochastic equilibrium, a remark. Mathematics of Operations Research, **17**, 148–164

[86] Dana, R.A., Le Van, C., Magnien, F. (1999): On different notions of arbitrage and existence of equilibrium. Journal of Economic Theory, **87**, 169–193

[87] Davis, M.H.A. (1997): Option pricing in incomplete markets. In: Dempster, M.H.A., Pliska, S.R. (eds) Mathematics of Derivative Securities, 216–227 Publication of the Newton Institute, Cambridge University Press

[88] Davis, M.H.A., Normann, A.R. (1990): Portfolio selection with transaction costs. Mathematics of Operations Research, **15**, 676–713

[89] Davis, M., Panas, V., Zariphopoulou, T. (1993): European option pricing with transaction costs. SIAM Journal of Control and Optimization, **31**, 470–493

[90] Debreu, G. (1953): Une Économie de l'Incertain. Working paper, Electricité de France

[91] Debreu, G. (1959): Théorie de la Valeur. Dunod, Paris

[92] Debreu, G. (1970): Economies with a finite set of equilibria. Econometrica, **38** (3), 387–392

[93] Debreu, G. (1972): Smooth preferences. Econometrica, **40** 4, 603–615

[94] Delbaen, F. (1992): Representing martingale measures when the asset prices are continuous and bounded. Mathematical Finance, **2**, 107–130

[95] Delbaen, F., Grandits, P., Rheinländer, T., Samperi, D., Schweizer, M. Stricker, C. (2002): Exponential hedging and entropic penalties. Mathematical Finance, **12**, 99–124

[96] Delbaen, F., Schachermayer, W. (1993): A General Version of the Fundamental Theorem of Asset Pricing. Mathematische Annalen, **300**, 463–520

[97] Delbaen, F., Schachermayer, W. (1994): Arbitrage and free-lunch with bounded risk for unbounded continuous processes. Mathematical Finance, **4**, 343–348

[98] Delbaen, F. and Schachermayer, W. (2005), The Mathematics of Arbitrage, Springer, Berlin

[99] De Marzo, P., Skiadas, N. (1998): Aggregation, determinacy, and informational efficiency for a class of economies with asymmetric information, Journal of Economic Theory, **80**, May, 123–152

[100] De Marzo, P., Skiadas, N. (1999): On the uniqueness of fully informative rational expectations equilibria. Economic Theory, **13**, January, 1–24

[101] Detemple, J., Murthy, S. (1994): Intertemporal asset pricing with heterogeneous beliefs. Journal of Economic Theory, **62**, 294–320

[102] Detemple, J., Zapatero, F. (1992): Optimal consumption–portfolio policies with habit formation. Mathematical Finance, **4**, 251–274

[103] Dierker, E. (1972): Two remarks on the number of equilibria of an economy. Econometrica, **40**, 951–953

[104] Dierker, E. (1974): Topological Methods in Walrasian Economics. Springer–Verlag, Berlin

[105] Di Masi, G., Platen, E., Runggaldier, W. (1991): Pricing of Options on Assets with Markovian Volatilities under Discrete Observations. Unpublished manuscript

[106] Di Masi, G., Platen, E., Runggaldier, W. (1991): Hedging of Options Under Discrete Observation on Assets With Stochastic Volatility. In : Seminar on Stochastic Analysis, Random Fields and Applications (E.Bolthausen, M.Dozzi, F.Russo, eds.). Progress in Probability, Vol.36, Birkhauser Verlag, 1995, pp. 359-364.

[107] Dothan, M. (1978): On the term structure of interest rates. Journal of Financial Economics, **7**, 229–269

[108] Dothan, M. (1990): Prices in Financial Markets. Oxford University Press, New York

[109] Dudley, R.M. (1977): Wiener functionals as Itô integrals. Annals of Probability, **5**, 140–141

[110] Duffie, D. (1986): Stochastic equilibria: existence, spanning number and the "no expected financial gain from trade" hypothesis. Econometrica, **54**, 1161–1183

[111] Duffie, D. (1987): Stochastic equilibria with incomplete financial markets. Journal of Economic Theory, **41**, 405–416

[112] Duffie, D. (1988): An extension of the Black–Scholes model of security valuation. Journal of Economic Theory, **46**, 194–204

[113] Duffie, D. (1989): Futures Markets. Prentice Hall, Englewood Cliffs

[114] Duffie, D. (1991): The theory of value in security markets. In: Hildenbrand, Sonnenshein (eds), Handbook of Mathematical Economics, Volume IV. North Holland, Amsterdam

[115] Duffie, D. (1992): Dynamic Asset Pricing Theory. Princeton University press

[116] Duffie, D., Epstein, L. (1992): Stochastic differential utility and asset pricing. Econometrica, **60**, 353–394

[117] Duffie, D., Fleming, W., Soner, H.M., Zariphopoulou, T. (1997): Hedging in incomplete markets with HARA utility. The Journal of Economics Dynamics and Control, **21**, 753–782

[118] Duffie, D., Geoffard, P., Skiadas, C. (1994): Efficient and equilibrium allocations with stochastic differential utility. Journal of Mathematical Economics, **23**, 133–146

[119] Duffie, D., Huang, C.F. (1985): Implementing Arrow–Debreu equilibria by continuous trading of few long-lived securities. Econometrica, **53**, 1337–1356

[120] Duffie, D., Kan, R. (1996): A yield-factor model of interest rates. Mathematical Finance, **5**, 379–406

[121] Duffie, D., Protter, P. (1992): From discrete to continuous time finance: weak convergence of the financial gain process. Mathematical Finance, **2**, 1–16

[122] Duffie, D., Shafer, W. (1985): Equilibrium in incomplete markets I : a basic model of generic existence. Journal of Mathematical Economics, **14**, 285–300

[123] Duffie, D., Shafer, W. (1985): Equilibrium in incomplete markets II: generic existence in stochastic economies. Journal of Mathematical Economics, **14**, 285–300

[124] Duffie, D., Skiadias, C. (1994): Continuous time security pricing: a utility gradient approach. Journal of Mathematical Economics, **23**, 107–131

[125] Duffie, D., Stanton, R. (1992): Pricing continuously resettled contingent claims. Journal of Economic Dynamics and Control, **16**, 561–574

[126] Duffie, D., Sun, T. (1989): Transaction costs and portfolio choice in a discrete-continuous time setting. Journal of Economics Dynamics and Control, **14**, 35–51

[127] Dumas, B. (1989): Two-person dynamic equilibrium in the capital market, Review of Financial Studies, **2**, 157–188

[128] Dumas, B., Luciano, E. (1991): An exact solution to a dynamic portfolio choice problem under transactions costs. The Journal of Finance, **2**, 577–587

[129] Dumas, B., Uppal, R., Wang, T., (2000): Efficient intertemporal allocations with recursive utility. Journal of Economic Theory, **93**, 240–259

[130] Dunn, K., Singleton, K. (1986): Modeling the term structure of interest rates under nonseparable utility and durability of goods. Journal of Financial Economics, **17**, 27–55

[131] Dupuis, P.G., Kushner, H.J. (1992): Numerical Methods for Stochastic Control Processes in Continuous Time. Springer–Verlag, Berlin

[132] Dybig, P.H., Huang, G. (1988): Non negative wealth, absence of arbitrage and feasible consumption plans. The Review of Financial Studies, **4**, 377–401

[133] Eberlein, E., Jacod, J., (1997): On the range of option pricing. Finance and Stochastics, **1**, 131–140

[134] Einstein, A. (1905): On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat. Ann. Physik, **17**

[135] Ekeland, I. (1979): Eléments d'Économie Mathématique. Hermann, Paris

[136] El Karoui, N. (1993): Lecture notes, Ecole Polytechnique

[137] El Karoui, N., Jeanblanc-Picqué, M., Lacoste, V. (2005): Optimal Portfolio Management with American Capital Guarantee. Journal of Control and Dynamic Theory, **29**, 449-468

[138] El Karoui, N., Jeanblanc-Picqué, M., Shreve, S. (1998): Robustness of the Black and Scholes Formula. Mathematical Finance, **8**, 93-126

[139] El Karoui, N., Jeanblanc-Picqué, M., Viswanathan, R., (1991): Bounds for options. In: Rutgers (ed) Proceedings US French Workshop on Applied Stochastic Analysis, Lecture Notes in Control and Information Sciences, **117**, 224–237. Springer–Verlag

[140] El Karoui, N., Lacoste, V. (1992): Multifactor Models of the Term Structure of Interest Rates. Congrès de l'association Française de Finance, Juilet 1992, Paris

[141] El Karoui, N., Myneni, R., Viswanathan, R. (1992): Arbitrage Pricing and Hedging of Interest Rate Claims with State Variables, I Theory. Unpublished manuscript, Laboratoire de Probabilités, Paris 6

[142] El Karoui, N., Myneni, R., Viswanathan, R. (1992): Arbitrage Pricing and Hedging of Interest Rate Claims with State Variables, II Applications. Unpublished manuscript, Laboratoire de Probabilités, Paris 6

[143] El Karoui, N., Peng, S., Quenez, M.C. (1997): Backward stochastic dif-
ferential equation finance and optimization. Mathematical Finance, **7**,
1–71

[144] El Karoui, N., Quenez, M-C. (1991): Programmation dynamique et
évaluation des actifs contingents en marché incomplet. C.R. Acad. Sci.
Paris, 851–854

[145] El Karoui, N., Quenez, M-C. (1995): Dynamic programming and pric-
ing of contingent claims in an incomplete market. SIAM J. control and
Optim., **33**, 29–66

[146] El Karoui, N., Rochet, J.C. (1990): A pricing formula for options on
coupon-bonds. Modèles Mathématiques en Finance, INRIA, Ecole CEA–
EDF–INRIA

[147] El Karoui, N., Rouge, R. (2000): Pricing via utility maximization and
entropy. Mathematical Finance, **10**, 259–276

[148] El Karoui, N., Saada, D. (1992): The Ho and Lee Model Revisited.
Unpublished manuscript, Laboratoire de Probabilités, Paris 6

[149] Elliot, R.J., Kopp P.E. (1998): Mathematics of Financial Markets.
Springer-Verlag, New York

[150] Elliott, R.J. and Van der Hoek, J. (2004) Binomial Models in Finance,
Springer Finance, Berlin

[151] Epstein, L., Miao., J. (2003): A two-person dynamic equilibrium under
ambiguity. Journal of Economic Dynamics and Control, **27**, 1253-1288

[152] Epstein, L. and Wang, T. (1994): Intertemporal Asset Pricing under
Knightian Uncertainty. Econometrica **62**, 283-322

[153] Epstein, L., Zin, S. (1989): Substitution, risk aversion and the temporal
behavior of consumption and asset returns : a theoritical framework.
Econometrica, **57**, 937–969

[154] Fitzpatrick, B., Fleming, W. (1991): Numerical methods for optimal
investment/consumption model. Mathematics of Operations Research,
16, 822–841

[155] Fleming, W., Rishel, R. (1975): Deterministic and Stochastic Control.
Springer–Verlag, Berlin

[156] Fleming, W. and Soner, M. (1993): Controlled Markov Processes and
Viscosity Solutions. Springer-Verlag, Berlin

[157] Fleming, W., Zariphopoulou, T. (1991): An optimal invest-
ment/consumption problem with borrowing. Mathematics of Operations
Research, **16**, 802–822

[158] Florenzano, M., Gourdel, P. (1996): Incomplete markets in infinite hori-
zon: debt constraints versus node constraints. Mathematical Finance, **6**,
167–196

[159] Florenzano, M., Le Van, C. (2000): Finite Dimensional Convexity and
Optimisation. Springer, Studies in Economic Theory, Berlin Heidelberg
New York

[160] Foldes. L. (1990): Conditions for optimality in the infinite-horizon portfolio-cum-saving problem with semi-martingale investments. Stochastics and Stochastics Reports, **29**, 133–170

[161] Föllmer, H., Leukert, P., (1999): Quantile hedging. Finance and Stochastics, **III(3)**, 251–273

[162] Föllmer, H. and Schied, A., (2004): Stochastic Finance : An Introduction in Discrete Time. 2nd Edition. DeGruyter, Berlin

[163] Föllmer, H., Sondermann, D. (1986): Hedging of non redundant contingent claims. In: Hildenbrand, W., Mas-Colell, A. (eds) Contributions to Mathematical Economics, in Honor of Gérard Debreu. North-Holland, Amsterdam

[164] Fouque, J-P. and Papanicolaou, G. and Sircar, X. (2000): Derivatives in financial markets with stochastic volatilities. Cambridge University press, Cambridge

[165] Frachot, A., Lesne, J.P. (1993): Econometrics of the Linear Gaussian Model of Interest Rates. Preprint, Banque de France

[166] Frittelli, M. (2000): The minimal entropy martingale measure and the valuation problem in incomplete markets. Mathematical Finance, **10**, 215–225

[167] Gabay, D. (1982): Stochastic processes in models of financial markets: the valuation equation of finance and its applications. In: Proceedings, IFIP Conference on Control of Distributed Systems. Pergamon Press, Toulouse

[168] Gale, D. (1955): The law of supply and demand. Mathematica Scandinavica, **3**, 155–169

[169] Geman, H. (1989): The Importance of the Forward Neutral Probability in a Stochastic Approach of Interest Rates. Working paper, ESSEC

[170] Geman, H., El Karoui, N., Rochet, J.C. (1995): Changes of numéraires, changes of probability measure and option pricing. The Journal of Applied Probability, **32**, 443–458

[171] Geman, H., Yor, M. (1992): Quelques relations entre processus de Bessel, options asiatiques et fonctions confluentes géomètriques. CRAS **314** Série 1, 471–474

[172] Geman, H., Yor, M. (1993): Bessel processes, Asian options and perpetuities. Mathematical Finance, **4**, 345–371

[173] Geman, H., Yor, M. (1996): Pricing and hedging double barrier options: a probabilistic approach. Mathematical Finance, **6**, 365–378

[174] Genon-Catalot, V., Jacod, J. (1993): On the estimation of the diffusion coefficient for multidimensional diffusion processes. Annales de l'Institut Henri Poincaré, **29**, 119–152

[175] Gerber, H., Pafumi, G. (2000): Pricing dynamic investment fund protection. North American Acturial Journal **4**, 28–41

[176] Gozzi, F., Vargioglu, T. (2002): Super-replication of European multi-asset derivatives with bounded stochastic volatility. Math. Methods in Oper. Reseach, **55**, 69–91

[177] Harrison, M., Kreps, D. (1979): Martingales and arbitrage in multi-period securities markets. Journal of Economic Theory, **20**, 381–408

[178] Harrison, J.M., Pliska, S.R. (1981): Martingales and stochastic integrals in the theory of continuous trading. Stochastic Processes and their Applications, **11**, 215–260

[179] Harrison, J.M., Pliska, S.R. (1983): A stochastic calculus model of continuous trading: complete markets. Stochastic Processes and their Applications, **15**, 313–316

[180] Haug, E.G. (1998): The Complete Guide to Option Pricing Formulas. McGraw-Hill, New York

[181] He, H., (1990) Essays in Dynamic Portfolio Optimization and Diffusion Estimations. Ph.D. Thesis, Sloan School of Management, MIT

[182] He, H.(1990): Convergence from discrete to continuous time contingent claims prices. The Review of Financial Studies, **3**, 523–546

[183] He, H. (1991): Optimal consumption–portfolio policies: a convergence from discrete to continuous time models. J. Econ. Theory **55**, 340–363

[184] He, H., Pagès, H. (1993): Labor Income, Borrowing Constraints and Equilibrium Asset Prices. Economic Theory, **3** (4), 663–696

[185] He, H., Pearson, N.D. (1991): Consumption and portfolio policies with incomplete markets and short sale constraints : the finite dimensional case. Mathematical Finance, **1**, 1–10

[186] He, H., Pearson, N.D. (1991): Consumption and portfolio policies with incomplete markets and short sale constraints : the infinite dimensional case. Journal of Economic Theory, **54**, 259–305

[187] Heath, D., Jarrow, R., Morton, A.J. (1990): Bond pricing and the term structure of interest rates. A discrete time approximation. Journal of Financial and Quantitative Analysis, **25**, 419–440

[188] Heath, D., Jarrow, R., Morton A.J. (1990): Contingent claim valuation with a random evolution of interest rates. The Review of Futures Markets, **9**, 54–76

[189] Heath, D., Jarrow, R., Morton, A.J. (1992): Bond pricing and the term structure of interest rates. A new methodology for contingent claims valuation. Econometrica, **60**, 77–106

[190] Heynen, R., Kat, H. (1995): Crossing barriers. Risk Magazine, **7**, 46–49

[191] Hildenbrand, W., Kirman, P. (1989): Introduction to Equilibrium Analysis. North-Holland, Amsterdam

[192] Hiriart-Urruty, J.B.I., Lemaréchal C. (1996): Convex analysis and minimization algorithms, Springer-verlag, Second edition, Berlin Heidelberg New York Tokyo

[193] Ho, T.S.Y., Lee, S. (1986): Term structure movements and pricing interest rate contingent claims. Journal of Finance, **41**, 1011–1029

[194] Hodges, S.D., Neuberger, A. (1989): Optimal replication of contingent claims under transaction costs. Review of Futures Markets, **8**, 222–239

[195] Hofmann, N., Platen, C., Schweizer, M. (1992): Option pricing under incompleteness and stochastic volatility. Mathematical Finance, **3**, 153–188

[196] Huang, C. (1987): An intertemporal general equilibrium asset pricing model : the case of diffusion information. Econometrica, **55**, 117–142

[197] Huang, C., Litzenberger, R.H., (1988): Foundations for Financial Economics. North-Holland, Amsterdam

[198] Huang, C., Pagès, H. (1992): Optimal consumption and portfolio policies with an infinite horizon. Existence and convergence. The Annals of Probability, **2**, 36–69

[199] Hugonnier, J. (1999): The Feynman-Kac formula and pricing occupation time derivatives. International Journal of Theoretical and Applied Finance, **2(2)**, 153–178

[200] Hull, J. (2000): Options, Futures, and Other Derivative Securities. Prentice-Hall, Englewood Cliffs, New Jersey

[201] Hull, J., White, A. (1987): The pricing of options on assets with stochastic volatilities. Journal of Finance, **42**, 281–300

[202] Hull, J., White, A. (1990): Pricing interest rate derivatives. The Review of Financial Studies, **3**, 573–592

[203] Hunt, P.J. and Kennedy, J.E. (2000) Financial Derivatives in Theory and Practice. Wiley, Chichester

[204] Ikeda, N., Watanabe, S. (1981): Stochastic Differential Equations and Diffusion Processes. North-Holland, New–York

[205] Jacod, J. (1979): Calcul Stochastique et Problèmes de Martingales, Lecture Notes in Mathematics, 714. Springer–Verlag, Berlin

[206] Jamshidian, F. (1989): The Multifactor Gaussian Interest Rate Model and Implementation. Preprint, World Financial Center, Merrill Lynch, New York

[207] Jamshidian, F. (1989): An exact bond option formula. Journal of Finance, **44**, 205–209

[208] Jamshidian, F. (1991): Bond and option evaluation in the Gaussian interest rate model. Res. Finance, **9**, 131–170

[209] Jamshidian, F. (1993): Option and Futures Evaluation with Stochastic Interest rate and Spot Yield. . Mathematical Finance **3**, 149–159

[210] Jamshidian, F. (1993): Option and future evaluation with deterministic volatilities. Mathematical Finance, **3**, 149–160

[211] Jarrow, R.A., Madan, D. (1991): Characterisation of complete markets in a Brownian filtration. Mathematical Finance, **3**, 31–43

[212] Jarrow, R.A., Madan, D. (1999): Hedging contingent claims on semimartingales. Finance and Stochastics, **3**, 111–134

[213] Jarrow, R.A., Rudd, A. (1983): Option Pricing. Irwin, Chicago

[214] Jarrow, R. and Turnbull (1996): Derivative securities. Southwestern college, Cincinnati

[215] Jeanblanc-Picqué, M., Pontier, M. (1990): Optimal portfolio for a small investor in a market model with discontinuous prices. Applied Mathematics and Optimization, **22**, 287–310

[216] Jeanblanc, M. and Yor, M. and Chesney, M. (2007): Mathematical Models for financial Markets. Springer, Berlin

[217] Jensen, B.A., Nielsen, J. (1998): The Structure of Binomial Lattice Models for Bonds. Surveys in applied and industrial mathematics,**5**, 361–386

[218] Johnson, H. (1987): Options on the Maximum or the Minimum of Several Assets. The Journal of Financial and Quantitative Analysis, **22** (3), 277–283

[219] Jouini E., C. Napp, (2006): Consensus consumer and intertemporal asset pricing with heterogeneous beliefs, to appear in Review of Economic studies

[220] Jouini E., C. Napp, (2006): Heterogeneous beliefs and asset pricing in discrete time: An analysis of doubt and pessimism, Journal of Economics Dynamics and Control, **30**, 1233–1260

[221] Jouini, E., Kallal, H. (1995): Martingales, arbitrage and equilibrium in securities markets with transactions costs. Journal of Economic Theory, **66**, 178–197

[222] Jouini, E., Kallal, H. (1995): Arbitrage in securities markets with short-sales constraints. Mathematical Finance, **5**, 178–197

[223] Jouini, E. and Napp, C. (2001): Market models with frictions: Arbitrage and Pricing Issues. In: Jouini, E. and Cvitanić, J. and Musiela, M. (eds) Option pricing, Interest rates and risk management, 43-66. Cambridge University Press

[224] Kabanov, Y. (2001): Arbitrage theory. In: Jouini, E. and Cvitanić, J. and Musiela, M. (eds) Option pricing, Interest rates and risk management, 3-42. Cambridge University Press

[225] Kabanov, Y., Kramkov, D. (1994): No arbitrage and equivalent martingale measures: an elementary proof of the Harrison–Pliska theorem. Theory of Probability and its Applications, **39**, 523–526

[226] Kabanov, Y.M., Safarian, M.M. (1997): On Leland's strategy of option pricing with transactions costs. Finance and Stochastics, **1**, 239–250

[227] Kabanov, Y. and Stricker, Ch. (2001) The Harrison-Pliska Arbitrage Pricing Theorem under Transaction Costs. J. Math. Econ., **35** (2) , 185–196.

[228] Kallianpur, G., Karandikar, R.L. (1999): Introduction to Option Pricing Theory. Birkhäuser

[229] Karatzas, I. (1988): On the pricing of American options. Applied Mathematics and Optimization, **17**, 37–60

[230] Karatzas, I. (1997): Lectures on the Mathematics of Finance. CRM monograph series, Volume 8, AMS

[231] Karatzas, I., Lehoczky, J., Shreve, S. (1987): Optimal portfolio and consumption decisions for a small investor on a finite horizon. SIAM Journal of Control and Optimization, **25**, 1557–1586

[232] Karatzas, I., Lehoczky, J., Shreve, S. (1990): Dynamic equilibria in multi-agent economy: construction and uniqueness. Mathematics of Operations Research, **15**, 80–128

[233] Karatzas, I., Shreve, S. (1998): Brownian Motion and Stochastic Calculus. Springer Verlag, Berlin

[234] Karlin, S. (1981): A Second Course in Stochastic Processes. Academic Press, New York

[235] Kat, H.M. (2001): Structured Equity Derivatives. Wiley, Chichester

[236] Kind, P., Liptser, R., Runggaldier, W. (1991): Diffusion approximation in path dependent models and applications to option pricing. The Annals of Applied Probability, **1**, 379–405

[237] Kloeden, R.E., Platen, E. (1991): The Numerical Solution of Stochastic Differential Equations. Springer–Verlag, Berlin

[238] Knight, F.B. (1981): Essentials of Brownian Motion and Diffusions, Math. Surveys, vol.18. American Mathematical Society, Providence, R.I

[239] Konno, H., Pliska, S.R. , Suzuki, K. (1993): Optimal portfolio with asymptotic criteria. Annals of Operations Research, **45**, 187–204

[240] Korn, R. (1998): Optimal Portfolio. World Scientific, Singapore

[241] Kramkov, D.O. (1996): Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. Prob. Theory and related fields, **105**, 459–479

[242] Kreps, D. (1981): Arbitrage and equilibrium in economies with infinitely many commodities. Journal of Mathematical Economics, **8**, 15–35

[243] Kreps, D. (1982): Multiperiod securities and the efficient allocation of risk : a comment on the Black–Scholes option pricing model, 203–232. In: McCall, J. (ed) The Economics of Uncertainty and Information. University of Chicago Press, Chicago, Illinois

[244] Kreps, D. (1990): A Course in Microeconomic Theory. Princeton University Press, Princeton

[245] Krylov, N. (1980): Controlled Diffusion Processes. Springer–Verlag, Berlin

[246] Kuhn, H.W. (1956): On a theorem of Wald, 265–273. In: Kuhn, H.W., Tucker, A.W. (eds) Linear Inequalities and Related Systems, Annals of Mathematical Studies, **38**. Princeton University Press, Princeton

[247] Kuhn, H.W. (1956): A note on the law of supply and demand. Mathematica Scandinavica, **4**, 143–146

[248] Kunitomo, N., Ikeda, M. (1992): Pricing options with curved boundaries. Mathematical Finance, **4**, 275–298

[249] Kusuoka, S. (1995): Limit theorem on option replication cost with transaction costs. Annals of Applied Probability, **5**, 198–221

[250] Lamberton, D., Lapeyre, B. (1997): Introduction to Stochastic Calculus Applied to Finance. Chapman & Hall, London

[251] Lehoczky, J., Sethi, S., Shreve, S. (1983): Optimal consumption and investment policies allowing consumption constraints and bankrupty. Math. of Operations Research, **8**, 613–636

[252] Leroy, S.F., Werner, J. (2001): Principles of Financial Economics. Cambridge University Press, Cambridge

[253] Levine, D.K., Zame, W. (1992): Debt Constraints and Equilibrium in Infinite Horizon Economies with Incomplete Markets. Review of Economic Studies, **60**, 865–888

[254] Lévy, P. (1948): Processus Stochastiques et Mouvement Brownien. Gauthier–Villars, Paris

[255] Linetsky, V. (1997): Step options. Mathematical Finance, **9**, 55–96

[256] Lintner, J. (1965): The valuation of risky assets and the selection of risky investment in stock portfolios and capital budgets. Review of Economics and Statistics, **47**, 13–37

[257] Lipton, A. (2001): Mathematical Methods for Foreign Exhange. World Scientific, Singapore

[258] Longstaff, F., Schwartz, E. (1992): Interest rate volatility and the term structure: a two factor general equilibrium model. The Journal of Finance, **47**, 1259–1282

[259] Lucas, R. (1978): Asset prices in an exchange economy. Econometrica, **46**, 1429–1445

[260] Luenberger, D.G. (1969): Optimisation by Vector Space Methods. Wiley, New York

[261] Madan, D. (2001): Purely discontinuous asset price process. In: Jouini, E. and Cvitanić, J. and Musiela, M.(eds) Option pricing, Interest rates and risk management. 67-104, Cambridge University Press

[262] Magill, M.G.P., Quinzii, M. (1993): Theory of Incomplete Markets. MIT Press, Boston

[263] Magill, M.G.P., Quinzii M. (1994): Infinite horizon incomplete markets. Econometrica, **62**, 853–880

[264] Magill, M., Shafer, W. (1991): Incomplete markets. In: Hildenbrand, W., Sonnenschein, H. (eds) Handbook of Mathematical Economics. North Holland, Amsterdam

[265] Martellini, L. and Priaulet, Ph. (2000): Fixed income securities for interest rates risk pricing and hedging. J. Wiley and Sons, Chichester

[266] Maruyama, (1955) Continuous Markov processes and stochastic equations. Rendiconti del Circolo Matematico Palermo, **4**, 48–90

[267] Mas-Colell, A. (1985): The Theory of General Economic Equilibrium – A Differentiable Approach. Cambridge University Press, Cambridge

[268] Mas-Colell, A., Winston, M.D., Green, J. (1995): Microeconomic Theory. Oxford University Press, Oxford

[269] Mas-Colell, A., Zame, W. (1991): Infinite dimensional equilibria. In: Hildenbrand, Sonnenshein (eds) The Handbook of Mathematical Economics Volume IV. North Holland, Amsterdam

[270] McKenzie, L. (1959): On the existence of general equilibrium for a competitive market. Econometrica, **27**, 54–71

[271] Mel'nikov, A.V. (1999): Financial Markets. Stochastic Analysis and the pricing of derivative securities. American Mathematical Society, Providence

[272] Merton, R. (1971): Optimum consumption and portfolio rules in a continuous time model. Journal of Economic Theory, **3**, 373–413

[273] Merton, R. (1973): An intertemporal Capital Asset Pricing Model. Econometrica, **41**, 867–888

[274] Merton, R. (1973): Theory of rational option pricing. Bell Journal of Economics and Management Science, **4**, 141–183

[275] Merton, R. (1976): Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics, **3**, 125–144

[276] Merton, R. (1991): Continuous Time Finance. Basil Blackwell, Oxford

[277] Michel, P. (1989): Cours de Mathématiques pour Économistes. Economica, Paris

[278] Mikosch, T. (1999): Elementary Stochastic calculus with finance in view. World Scientific, Singapore

[279] Milstein, G.N. (1974): Approximate integration of stochastic differential equations. Theory of Probability and Applications, **19**, 557–562

[280] Miyahara, Y. (1997): Canonical martingale measures of incomplete assets markets. In: Watanabe, S., Yu, V. Prohorov, V., Fukushima, M., Shiryaev, A.N. (eds) Proceedings of the Seventh Japan Russia Symposium, World Scientific

[281] Morton, A.J. (1989): Arbitrage and Martingales. Ph. D. Thesis. Cornell University

[282] Morton, A.J., Pliska, S.R. (1995): Optimal portfolio management with fixed transaction costs. Mathematical Finance, **5**, 337–356

[283] Mulinacci, S., Pratelli, M. (1998): Functional convergence of Snell envelopes: applications to American options approximations. Finance and Stochastics, **2**, 311–327

[284] Müller, S. (1987): Arbitrage Pricing of Contingent Claims, Lecture Notes in Economics and Mathemetical Systems no 254. Springer Verlag, Berlin

[285] Musiela, M., Rutkowski, M. (1997): Martingale Methods in Financial Modelling. Springer, Berlin–Heidelberg

[286] Muth, J. (1961): Rational expectations and the theory of price movements. Econometrica, **29**, 315–335

[287] Myneni, R. (1992): The pricing of American options. The Annals of Applied Probability, **2**, 1–23

[288] Negishi, T. (1960): Welfare economics and existence of an equilibrium for a competitive economy. Metroeconomica, **12**, 92–97

[289] Niederreiter, H. (1992): Random Number Generation and Quasi-Monte Carlo Methods. Society for Industrial and Applied Mathematics, Philadelphia

[290] Nielsen, L.T. (1989): Asset market equilibrium with short-selling. Review of Economic Studies, **56**, 467–474

[291] Nielsen, L.T. (1989): Existence of equilibrium in CAPM. Journal of Economic Theory, **52**, 223–231

[292] Nielsen, L.T. (1990): Equilibrium in CAPM without a riskless asset. Review of Economic Studies, **57**, 315–324

[293] Nikaido, H. (1956): On the classical multilateral exchange problem. Metroeconomica, **8**, 135–145

[294] Øksendal, B. (1998): Stochastic Differential Equations. Springer–Verlag, Berlin

[295] Overhaus, M., Ferraris, A., Knudsen, T., Milward, R., Nguyen-Ngoc, L., Schindlmayr, G., (2002): Equity Derivatives, Theory and Applications. Wiley Finance, New York

[296] Page, F.H. Jr (1996): Arbitrage and asset prices. Mathematical Social Sciences, **31**, 183–208

[297] Pagès, H. (1989): Three Essays in Optimal Consumption. Ph.D. Thesis, MIT

[298] Pardoux, E., Talay, D. (1985): Discretization and simulation of stochastic differential equations. Acta Applicandae Mathematicae, **3**, 23–47

[299] Pham, H. (1998): Méthodes d'Evaluation et Couverture d'Options en Marché Incomplet. Ensae lecture notes, option formation par la recherche

[300] Pliska, S.R. (1986): A stochastic calculus model of continuous trading: optimal portfolios. Mathematics of Operations Research, **11**, 371–382

[301] Pliska, S.R. (1997): Introduction to Mathematical Finance. Blackwell, Oxford

[302] Prigent, J.L. (1999): Incomplete markets : convergence of option values under the risk minimal martingale measure. Advances in Applied Probability, **31**, 1–20

[303] Prigent, J-L. (2003): Weak Convergence of financial markets, Springer Finance, Berlin

[304] Protter, P. (2005): Stochastic Integration and Stochastic Differential Equations. Second edition. Springer–Verlag, Berlin

[305] Radner, R. (1972): Existence of equilibrium of plans, prices and price expectations in a sequence of markets. Econometrica, **40**, 289–303

[306] Rebonato, R. (1997): Interest-rate Option Models. Wiley, Second edition, Chichester

[307] Revuz, D., Yor, M., (1999): Continuous Martingales and Brownian Motion. Third edition. Springer–Verlag, Berlin

[308] Rich, D.R. (1994): The mathematical foundations of barrier option-pricing theory. Advances in Futures and Options Research, **7**, 267–311

[309] Riedel, F., (2001): Existence of Arrow-Radner Equilibrium with endogenously complete markets under incomplete information. Journal of Economic Theory, **97**, 109–122

[310] Ripley, B.D. (1987): Stochastic Simulation. Wiley, New York

[311] Roberts, G.O., Shortland, C.F. (1997): Pricing barrier options with time-dependent coefficients. Mathematical Finance, **7**, 83–93

[312] Rockafellar, R.T.(1970): Convex Analysis. Princeton University Press, Princeton

[313] Rogers, L.C.G., Shi, Z. (1995): The value of an Asian option. J. Appl. Prob., **32**, 1077–1088

[314] Rogers, L.C.G., Talay, D. (1997): Numerical Methods in Finance. Cambridge University Press, Cambridge

[315] Rogers, L.C.G., Williams, D. (1988): Diffusions, Markov Processes and Martingales. Wiley, New York

[316] Ross, S. (1976): The arbitrage theory of capital asset pricing. Journal of Economic Theory, **13**, 341–360

[317] Ross, S. (1978): A simple approach to the valuation of risky streams. Journal of Business, **51**, 453–475

[318] Rubinstein, M. (1976): The valuation of uncertain income streams and the pricing of options. Bell Journal of Economics, **7**, 407–425

[319] Rubinstein, M., Reiner, E., (1991): Breaking down the barriers. Risk Magazine, **9**, 28–35

[320] Runggaldier, W.J., Schweizer M. (1995): Convergence of option values under incompleteness. In: Bolthausen, E., Dozzi, M., Russo, F. (eds) Seminar on Stochastic Analysis, Random Fields and Applications, Ascona 1993, Progress in Probability, Birkhauser, 365-384.

[321] Sandmann, K., Sondermann, D. (1993): A Term Structure Model and the Pricing of Interest Rate Derivatives. The review of futures Markets, **12** (2), 391–423

[322] Santos, M., Woodford, M. (1997): Rational asset pricing models. Econometrica, **65**, 19–55

[323] Scarf, H.E. (1967): The computation of equilibrium prices, an exposition. In: Arrow, K.J., Intrilligator, M. (eds) The Handbook of Mathematical Economics, Volume II. North Holland, Amsterdam

[324] Scarf, H.E. (1967): On the computation of equilibrium prices. In: Fellner, W.J. (ed), Ten Economic Studies in the tradition of Irving Fisher, 207–230. Wiley, New York

[325] Schachermayer, W. (1992): A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time. Insurance: Mathematics and Economics, **11**, 291–301

[326] Schachermayer, W. (1994): Martingale measures for discrete time processes with infinite horizon. Mathematical Finance, **4**, 25–56

[327] Schaefer, S., Schwartz, E. (1984): A two factor model of the term structure: an approximate analytical solution. Journal of Financial and Quantitative Analysis, **19**, 413–424

[328] Schuger, K. (1996): On the existence of equivalent τ-measures in finite discrete time. Stochastic Processes and their Applications, **61**, 109–128

[329] Schweizer, M. (1988): Hedging of Options in a General Semimartingale Model. Ph.D. Thesis. Zurich

[330] Schweizer, M. (1991): Option hedging for semimartingales. Stochastic Process and Applications, **37**, 339–363

[331] Schweizer, M. (1992): Mean–variance hedging for general claims. The Annals of Probability, **2**, 171–179

[332] Serrat A. (2001): A dynamic equilibrium model of international portfolio holdings, Econometrica, **69**, 1467–1489

[333] Sharpe, W. (1964): Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance, **19**, 425–442

[334] Shirakawa, H. (1991): Interest rate option pricing with poisson-gaussian forward rate curve processes. Mathematical Finance, **1**, 77–94

[335] Shirakawa, H. (1994): Optimal consumption and portfolio selection with incomplete markets and upper and lower bounds constraints. Mathematical Finance, **4**, 1–24

[336] Shiryaev, A. (1999): Essential of Stochastic Finance. World Scientific, Singapore

[337] Shreve, S.E. (2004): Stochastic Calculus Models for Finance, discrete time, Springer

[338] Shreve, S.E. (2004): Stochastic Calculus Models for Finance, Springer

[339] Shreve, S., Soner, M. (1991): Optimal investment and consumption with two bonds and transaction costs. Mathematical Finance, **1**, 53–84

[340] Shreve, S., Soner, M. (1994): Optimal investment and consumption with transaction costs. Ann. Appl. Probab., **4**, 609–692

[341] Shreve, S.E., Soner, H.M., Cvitanić J. (1995): There is no nontrivial hedging portfolio for option pricing with transaction costs. Ann. Appl. Prob., **5**, 327–355

[342] Shreve, S., Xu, G. (1992): A duality method for optimal consumption and investment under short-selling prohibition. I General market coefficients. The Annals of Probability, **2**, 87–112

[343] Shreve, S., Xu, G. (1992): A duality method for optimal consumption and investment under short-selling prohibition. II Constant market coefficients. The Annals of Probability, **2**, 314–328

[344] Stanton, R. (1989): Path Dependent Payoffs and Contingent Claims Valuation: Single Premium Deferred Annuities. Unpublished manuscript

[345] Steele, M. (2001): Stochastic calculus and Financial applications. Second edition. Springer Verlag, Berlin

[346] Stokey, N., Lucas, R.E., Prescott, E.C. (1989): Recursive Methods in Economics Dynamics. Harvard University Press, Boston

[347] Stricker, C. (1984): Integral representation in the theory of continuous trading. Stochastics, **13**, 249–265

[348] Stricker, C. (1989): Arbitrage et lois de martingale. Annales Inst. Henri Poincaré, **26**, 451–460

[349] Sulem A. (1992): Cours de DEA, Université Paris 6.

[350] Talay, D. (1983): Résolution trajectorielle et analyse numérique des équations différentielles stochastiques. Stochastics, **9**, 275–306

[351] Talay, D. (1986): Discrétisation d'une EDS et calcul approché d'espérance de fonctionnelles de la solution. Math. Modelling and Numerical Analysis, 20–1, 141

[352] Talay, D. (1991): Simulation and numerical analysis of stochastic differential systems. In: Kree, P., Wedig, W. (eds) Effective Stochastic Analysis. Springer–Verlag, New–York, Heidelberg, Berlin

[353] Taleb, N. (1997): Dynamic hedging. Wiley, New-York

[354] Tallon, J.M. (1995): Théorie de l'équilibre général avec marchés financiers incomplets. Revue Economique, **46**, 1207–1239

[355] Touzi, N. (1998): Ensae lecture notes, option formation par la recherche

[356] Uzawa, H. (1956): Note on the Existence of an Equilibrium for a Competitive Economy. Unpublished, Department of Economics, Stanford University

[357] Uzawa, H. (1960): Walras tatonnement in the theory of exchange. Review of Economic Studies, **27**, 182–194

[358] Varadhan, S.R.S. (1980): Lectures on Diffusion Problems and Partial Differential Equations. Tata Institute of Fundamental Research, Bombay

[359] Varian, H.R. (1988): Le principe d'arbitrage en économie financière. Annales d'Economie et de Statistique, **10**, 1–22

[360] Vasicek, O. (1977): An equilibrium characterization of the term structure. Journal of Financial Economics, **5**, 177–188

[361] Wald, A. (1936): Über einige Gleichungssysteme der mathematischen Ökonomie. Zeitschrift für Nationalökonomie, **7**, 637–670. Translated into English. On some systems of equations of mathematical economics. Econometrica, **19** (1951), 368–403

[362] Walras, L. (1874–7): Eléments d'économie politique pure. Corbaz, Lausanne. Translated as: Elements of Pure Economics. Irwin, Chicago (1954)

[363] Wang, J., The term structure of interest rates in a pure exchange economy with heterogeneous investors. Journal of Financial Economics, **41**, 75–110

[364] Werner, J. (1985): Equilibrium in economies with incomplete financial markets. Journal of Economics Theory, **36**, 110–119

[365] Werner, J. (1987): Arbitrage and the existence of competitive equilibrium. Econometrica, **55**, 1403–1418

[366] Wiener, N. (1923): Differential space. Journal of Mathematical Physics, **2**, 131–174

[367] Williams, D. (1991): Probability with Martingales. Cambridge University Press, Cambridge

[368] Willinger, W., Taqqu, M. (1991): Towards a convergence theory for continuous stochastic securities market models. Mathematical Finance, **1**, 55–99

[369] Wilmott, P. (1998) Derivatives; the Theory and Practice of Financial Engineering. University Edition, John Wiley, Chichester

[370] Wilmott, P. (2001): Paul Wilmott Introduces Quantitative Finance. John Wiley, Chichester

[371] Wilmott, P., Dewynne, J., Howson, S. (1994): Options Pricing. Mathematical Models and Computation. Oxford Financial Press, Oxford

[372] Xia, J. and Yan, J-A (2001): Some remarks on arbitrage pricing theory. In: Yong, J. (ed) International conference on Mathematical Finance: Recent developmentss in Mathematical Finance. 218-227. World Scientific, Singapore

[373] Yor, M. (1992): Some Aspects of Brownian Motion, Part I: Some Special Functionals. Lectures in Mathematics, ETH Zürich, Birkäuser, Basel

[374] Yor, M. (1995): Local Times and Excursions for Brownian Motion : a Concise Introduction. Lecciones en Matemáticas, Facultad de Ciencias, Universidad Central de Venezuela, Caracas

[375] Yor, M. (1997): Some Aspects of Brownian Motion, Part II: Some Recent Martingale Problems. Lectures in Mathematics, ETH Zürich, Birkäuser, Basel

[376] Yor, M., Chesney, M., Geman, H., Jeanblanc-Picqué, M. (1997): Some combinations of Asian, Parisian and barrier options. In: Dempster, M., Pliska, S.R. (eds) Mathematics of Derivative Securities. The Newton Institute, Cambridge University Press, Cambridge

[377] Zapatero, F. (1998): Effects of financial innovation on market volatility when beliefs are heterogeneous. Journal of Economic Dynamics and Control, **22**, 597–626

[378] Zariphopoulou, T. (1992): Consumption investment models with constraints. SIAM Journal of Control, **32**, 59–85

[379] Zhang, P.G. (1997): Exotic Options. World Scientific, Singapore

Index