Advanced Topics in Mathematical Finance

Frühjahrs-/Sommersemester 2022

David Prömel*
Universität Mannheim

March 31, 2022

Abstract

This is a **preliminary version** of the lecture notes for the course "Advanced Topics in Mathematical Finance", which will be continuously updated during the semester and may contain many mistakes. Please be careful when using it and let me know if you find mistakes or have suggestions for improvements. The lecture course deals with central topics of mathematical finance in continuous time.

Contents

Recommended literature

- Jarrow, R. A., *Continuous-Time Asset Pricing Theory.* Springer Finance, 2018.
  [Recommended for general arbitrage theory]

- Kwok, Y.-K., *Mathematical Models of Financial Derivatives.* Springer Finance, 2008.
  [Recommended for pricing and hedging of options]

- Yan, J.-A., *Introduction to Stochastic Finance*, Springer, 2018.
  [Recommended for pricing of options and portfolio optimization in diffusion models]

- Eberlein, E., Kallsen, J. *Mathematical Finance*, Springer Finance, 2019.
  [Recommended for term structure models]

- Gatheral, J., *The Volatility Surface: A Practitioner's Guide.* John Wiley & Sons, 2006.
  [Comprehensive reference for Volatility modelling]

- Pham, H., *Continuous-time Stochastic Control and Optimization with Financial Applications*, Springer-Verlag, 2009.
  [Comprehensive reference for portfolio optimization and stochastic control]

- Filipovic, D., *Term-Structure Models: A Graduate Course*, Springer Finance, 2009.
  [Comprehensive reference for term structure modelling]

- Björk, T., *Arbitrage Theory in Continuous Time*, Oxford University Press, 2009.
  [Classical international reference for mathematical finance]

- Shreve, S. E., *Stochastic calculus for finance*, II. Springer-Verlag, 2004.
  [Classical international reference for mathematical finance]

- Sondermann, D., *Introduction to Stochastic Calculus for Finance: A New Didactic Approach*, Springer-Verlag, 2006.
  [A mathematical simpler introduction to mathematical finance]

- Kuo, H.-H., *Introduction to Stochastic Integration*, Springer, 2006.
  [Recommended for the background knowledge in stochastic calculus]

- Klenke, A., *Probability Theory.* Springer-Verlag, 2006.
  [Recommended for the background knowledge in probability theory. There is a German version of the book called "*Wahrscheinlichkeitstheorie*".]

1 Introduction - financial modeling in continuous time

Lecture 1

The aim of the course "Advanced Topics in Mathematical Finance" is to develop the theory and mathematical modeling of financial markets in continuous time and to study fundamental tasks arising in the financial industry.

The scope of this course is:

- no arbitrage theory in continuous time (fundamental theorems of asset pricing)
- market completeness
- Black-Scholes model (hedging/pricing of vanilla and exotic options)
- volatility models (stochastic and local volatility)
- optimal investment and basics of stochastic optimal control
- term structure theory for interest rates

As in the course "Mathematical Finance" (i.e. financial modeling in discrete time), we need to agree on the general structure and the main objectives of the type of financial market which we want to study. Like modeling in many other areas, financial modeling is based on philosophical concepts/beliefs and mathematical abstraction/simplification of the real world. Hence, before setting up mathematical models for financial markets, we need to agree some basic assumptions about the structure of the considered financial market.

Abstract structure of financial markets

Types of players:

- **Investors/Buyers** have money to invest.
- **Borrowers/Sellers** need money from the market.

Types of trading systems:

- **Exchanges** – regulated, transparent.

Types of things traded:

- **Shares (stocks or equities)** ("risky assets"): a fixed fraction of a company. It entitles its owner to future cash flows (dividends) from the firm's profits, and residual assets in the case of collapse. Valuation reflects the perception of further performance as well as assessment of other investors' views (e.g. will the company be taken over?). Hence the stock value changes as the views on the future events change.
- **Bonds** ("risk-free assets"): loans to corporations or governments. Some are essentially risk-free (default-free), such as the US bonds used to be, others are more risky.
- **(Financial) derivatives** on all the above: assets with future payoff depending on the future prices of underlying assets

This listing regarding financial markets is of course by no means comprehensive but sufficient for our purposes.

Fundamental concepts

In order to model a financial market we shall make some (reasonable?) assumptions how the considered financial market works. Of course, they are always subject to discussion and criticism. One should understand that these assumptions are invoked whenever we refer to the **no arbitrage principles (NA principles)**:

(i) **efficient market hypothesis, (EMH)**, roughly says that the price fully reflects all publicly available information.

(ii) **no arbitrage (NA)** (also called absence of arbitrage), roughly says, it is not possible to make a profit without risk or investment. This is often described as there is no free lunch.

(iii) **unlimited liquidity**: assets can be bought or sold in any quantity at a given market price at any moment of time.

(iv) **linear pricing rule:** the price of a linear combination of two assets (a portfolio of assets) equals the linear combination of their prices.

The NA principles are standing assumptions throughout the entire course. Recall, in a non-mathematical manner, **no arbitrage** (also called absence of arbitrage) means, there is no admissible trading strategy, which

- starts with zero initial capital and
- delivers, at some fixed time in the future, a value which is: non-negative and positive with positive probability.

In fact, financial markets are full of arbitrageurs, whose job is to find arbitrages and exploit them. These arbitragers move the prices by trading in the corresponding financial instruments, so that the arbitrage is eliminated. This justifies the no arbitrage assumption.

Remark.

- Since arbitrage does exist (albeit for a short period of time), a financial model that uses the no arbitrage principles can only be viewed as an approximation of reality.
- Each model is constructed in order to give answers to specific questions. The 'precision' with which a model approximates the reality should be sufficient to guarantee the required precision of the answers.
- This is, indeed, possible in physics, where there are well established 'scales': a model may produce an error on a microscopic scale, but do so such that this does not translate into a significant error on the macroscopic scale.
- A main difficulty in modeling for the social sciences (such as finance) is the lack of well defined 'scales'. In other words, 'very large' is often not large enough, and 'very small' is not small enough. Furthermore, social behavior could change over the years. Hence, the risk of model failure is of crucial importance in financial modeling.

Bottom line: Always keep the purpose and limitations of our models in mind!

Note that these NA principles lead to further implicit assumptions:

- *borrowing and lending rates are equal*,
- *liquid markets* (you can buy arbitrary qualities of all trading things),
- *no transaction costs* (taxes, fees, etc. ...),
- *no price impact* (trading does not effect prices),
- *short selling is allowed* (one can buy negative quantities of all financial assets),
- ... .

Remark. Of course, one has to remember that any of the above assumptions can be violated in practice. In particular, real financial markets are not as liquid as we would like, so these assumptions are again only approximations of the real world.

2 Arbitrage theory in continuous time

In this chapter we develop the foundation of mathematical finance in continuous time. In the next subsection we start with the general modeling of financial markets in a continuous-time setting.

2.1 Financial modeling in continuous time

We want to model the price evolution of risky assets (e.g. stocks) in continuous time which is a reasonable approximation of real financial markets for a couple of reasons: trading is not done on a fixed grid as required in discrete-time models, high-frequency trading leads almost to continuous-time trading in reality, some phenomena can be better explained (e.g. price bubbles) in continuous time and, in many instances, the resulting mathematics becomes "nicer" (e.g. formulas option prices).

As future prices are unknown today and cannot be precisely forecasted, the natural way is to rely on stochastic processes. We consider a financial market modeled by a $d+1$-dimensional non-negative Itô process $(S_t)_{t\in[0,T]}$ with

$$S_t = (S_t^0, S_t^1, \dots, S_t^d), \quad t \in [0,T],$$

where we usually think of these processes as

- S^0 is the price process of a *risk-free asset* (e.g. bonds, bank account, ...),
- $S^1, \dots, S^d$ are price processes of *risky assets* (e.g. prices of stocks, exchanges rates, ...).

We always assume that all stochastic processes are defined on a suitable filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with an n-dimensional Brownian motion $W = (W_t^1, \dots, W_t^n)_{t\in[0,T]}$, i.e. $W_t^1, \dots, W_t^n$ are independent one-dimensional Brownian motions, and the underlying filtration is the Brownian standard filtration $(\mathcal{F}_t)_{t\in[0,T]}$, i.e. generated by the Brownian motion W and completed to satisfy the usual condition (i.e. $(\mathcal{F}_t)_{t\in[0,T]}$ is complete and right-continuous). Note that $\mathcal{F}_0$ is trivial (meaning $\mathcal{F}_0$ only contains sets with probability 0 or 1) In this case we can suppose that $S_0^0, S_0^1, \dots, S_0^d$ are constants. Furthermore, we suppose that $\mathcal{F}_T = \mathcal{F}$.

More specifically, we assume that

$$\begin{aligned} S_t^0 &= 1 + \int_0^t r_u S_u^0 \,\mathrm{d}u, \\ S_t^i &= S_0^i + \int_0^t \mu_u^i \,\mathrm{d}u + \sum_{j=1}^n \int_0^t \sigma_u^{ij} \,\mathrm{d}W_u^j, \quad t \in [0,T], \end{aligned}$$

for $i = 1, \dots, d$, where $r, \mu^i, \sigma^{ij} \colon \Omega \times [0,T] \to \mathbb{R}$ are adapted, measurable processes satisfying the integrability conditions

$$\sup_{u\in[0,T]} |r_u| < \infty, \quad \int_0^T |\mu_u^i| \,\mathrm{d}u < \infty \qquad \text{and} \qquad \int_0^T |\sigma_u^{ij}|^2 \,\mathrm{d}u < \infty, \quad \mathbb{P}\text{-a.s.},$$

for $j = 1, \dots n$. We set $\mu := (\mu^1, \dots, \mu^d)$ and $\sigma := (\sigma^{ij})_{1\le i\le d,\, 1\le j\le n}$. Usually, r is called the **interest rate**, μ the **drift/trend** on the financial market and σ the **volatility** of the financial market. Note that the very definition of the financial model $S = (S^0, S^1, \dots S^d)$ implements the assumption that the borrowing and lending rates are equal, that the market is liquid and that there is no price impact. We also assume that the risky asset have no cash flow (dividends).

Let us briefly discuss a simple example of a continuous-time financial market to demonstrate why we need to consider such complicated stochastic processes for the modeling of financial markets.

Example 2.1. Consider the financial market $(S_t^0, S_t^1)_{t\in[0,T]}$ given by

- $S_t^0 = 1$ and
- $S_t^1 = S_0^1 + \int_0^t \mu_u^1 \,\mathrm{d}u$, for $t \in [0,T]$ (i.e. $r_s = 0$ and $\sigma = 0$).

For simplicity we restrict trading on this financial market to self-financing trading strategies $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ of the form:

- $\varphi_t^1 = f(S_t)$, for some $f \in C(\mathbb{R};\mathbb{R})$, stands for the numbers of shares of risky assets hold at time t,
- φ_t^0 stands for the numbers of risk-free assets held at time t, chosen such that $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ is self-financing.

Hence, the capital process $(V_t(\varphi))_{t\in[0,T]}$ generated by trading according to $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ satisfies

$$\begin{aligned} V_t(\varphi) &= \int_0^t \varphi_u^0 \,\mathrm{d}S_u^0 + \int_0^t \varphi_u^1 \,\mathrm{d}S_u^1 \qquad \left[\approx \sum_{i=0}^{N-1} \varphi_{t_i}^0 (S_{t_{i+1}}^0 - S_{t_i}^0) + \sum_{i=0}^{N-1} \varphi_{t_i}^1 (S_{t_{i+1}}^1 - S_{t_i}^1) \right] \\ &= \int_0^t f(S_u^1) \,\mathrm{d}S_u^1, \end{aligned}$$

for $0 \leq t_0 \leq \cdots \leq t_N \leq T$ and $N \in \mathbb{N}$.

Notice, taking a function $F\colon \mathbb{R} \to \mathbb{R}$ such that $F'(x) = f(x)$, the fundamental theorem of calculus reveals

$$F(S_T^1) - F(S_0^1) = \int_0^T f(S_u^1)\mu_u^1 \,\mathrm{d}u = \int_0^T f(S_u^1)\,\mathrm{d}S_u^1.$$

Hence, taking $f(x) := 2(x - S_0^1)$ (i.e. $F(x) = (x - S_0^1)^2$) gives

$$(S_T^1 - S_0^1)^2 = \int_0^T 2(S_u^1 - S_0^1)\,\mathrm{d}S_u^1 \geq 0,$$

and obtain that the corresponding capital process $(V_t(\varphi))_{t\in[0,T]}$ satisfies

$$V_T(\varphi) = \int_0^T 2(S_u^1 - S_0^1)\,\mathrm{d}S_u^1 = (S_T^1 - S_0^1)^2 \geq 0 \quad \text{and} \quad V_0(\varphi) = 0,$$

which is an arbitrage opportunity as soon as $\mathbb{P}(S_T^1 \neq S_0^1) > 0$.

Lecture 2

The most prominent example in the course of a model for a financial market will be the the so-called Black-Scholes model, which we will later discuss in more detail. In order to have a compact notation, we recall the definition of the stochastic exponential of a Itô process.

Definition 2.2. For an Itô process $X = (X_t)_{t\in[0,T]}$ the **stochastic exponential** $(\mathcal{E}(X)_t)_{t\in[0,T]}$ of X is defined by

$$\mathcal{E}(X)_t := \exp\left(X_t - \frac{1}{2}\langle X\rangle_t\right), \quad t \in [0,T],$$

where $(\langle X\rangle_t)_{t\in[0,T]}$ denotes the quadratic variation process of X.

Example 2.3 (Black-Scholes model). The (one-dimensional, i.e. $d = 1$) Black-Scholes model consists of

- the risk-free asset $(S_t^0)_{t\in[0,T]}$ which is often given by

  $$S_t^0 = S_0^0 \exp(rt) = S_0^0 \mathcal{E}(rI)_t, \quad t \in [0,T],$$

  where $S_0^0 > 0$, $r \in \mathbb{R}$ is the interest rate and $I = (I_t)_{t\in[0,T]}$ is the identity process $I_t = t$, which is the solution to

  $$\mathrm{d}S_t^0 = rS_t^0\,\mathrm{d}t, \quad S_0^0 = S_0^0, \quad t \in [0,T],$$

- the risky asset $(S_t^1)_{t\in[0,T]}$ given by

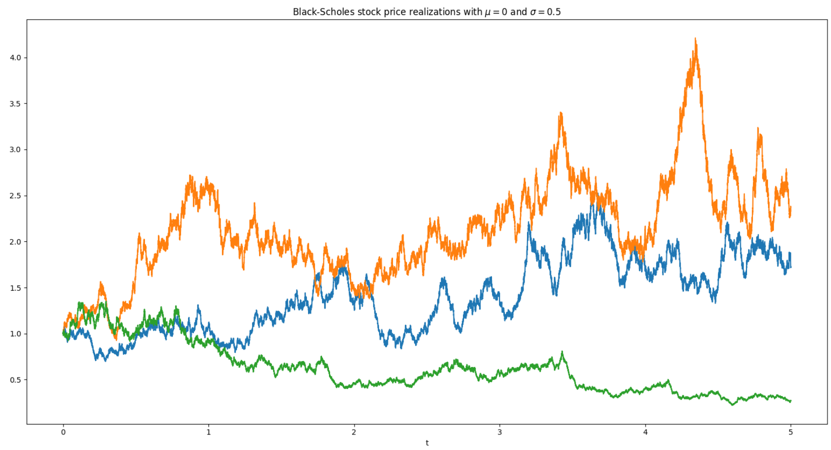
  $$S_t^1 = S_0^1 \exp(\widetilde{\mu}t + \sigma W_t) = S_0^1 \mathcal{E}(\mu I + \sigma W)_t, \quad t \in [0,T],$$

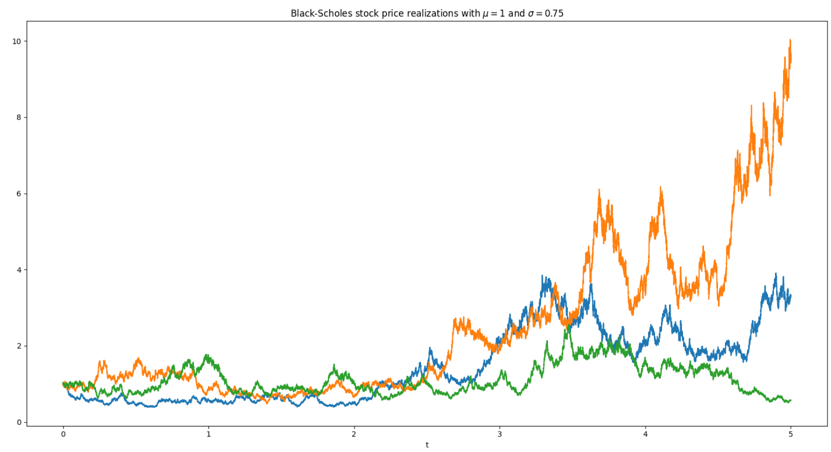
  which is the solution (check!) to

  $$\mathrm{d}S_t^1 = \mu S_t^1\,\mathrm{d}t + \sigma S_t^1\,\mathrm{d}W_t, \quad S_0^1 = S_0^1, \quad t \in [0,T],$$

  where $S_0^1 > 0$, $\mu \in \mathbb{R}$ is the drift parameter, $\sigma > 0$ is the volatility parameter, $\widetilde{\mu} := \mu - \frac{1}{2}\sigma^2$, and $(W_t)_{t\in[0,T]}$ is a Brownian motion.

We will study the Black-Scholes model in more details in Section 3.1.





After setting up mathematical models for financial markets using $d+1$-dimensional Itô processes $S = (S_t^0, \ldots, S_t^d)_{t\in[0,T]}$, we introduce how one can invest in such a market. For a $\mathbb{R}^{d+1}$-valued process $\varphi = (\varphi_t^0, \varphi_t^1 \ldots, \varphi_t^d)_{t\in[0,T]}$ we say

$$\varphi \in \mathcal{L}(S) \quad \text{if } \varphi^i \in \mathcal{L}(S^i) \cap \mathcal{L}(I) \quad \text{for } i = 0, 1, \ldots, d,$$

and the corresponding integral process is defined as

$$(\varphi \cdot S)_t := \int_0^t \varphi_u \, \mathrm{d}S_u := \sum_{i=0}^d \int_0^t \varphi_u^i \, \mathrm{d}S_u^i, \quad t \in [0, T],$$

for $\varphi \in \mathcal{L}(S)$.

Definition 2.4. A **trading strategy** (or **portfolio**) is a $\mathbb{R}^{d+1}$-valued process

$$\varphi = (\varphi_t^0, \varphi_t^1 \ldots, \varphi_t^d)_{t\in[0,T]} \in \mathcal{L}(S)$$

and the corresponding **value (**or **capital) process** $V(\varphi) = (V_t(\varphi))_{t\in[0,T]}$ is given by

$$V_t(\varphi) := \varphi_t^T S_t = \sum_{i=0}^{d} \varphi_t^i S_t^i, \quad t \in [0,T],$$

where φ_t^T is the transpose of random vector φ_t.

A trading strategy φ is called **self-financing** if

$$V_t(\varphi) = V_0(\varphi) + (\varphi \cdot S)_t, \quad t \in [0,T].$$

Remark (Financial intuition). The random variable φ_t^i represents the amount, one holds, of the asset i at time $t \in [0,T]$. The stochastic integral $(\varphi \cdot S)$ represents the gains and losses associated to trading according to the trading strategy φ into the financial market S. Self-financing means that no capital is added or withdrawn from the value process. In other words, all gains/losses are created by the changes in the underlying price process S weighted by the continuously re-arranging of the positions according to φ.

Since these interpretations are not so obvious in the present abstract continuous-time setting, it might be good to recall the corresponding definitions provided in mathematical finance in discrete time, cf. the lecture course "Mathematical Finance".

In an economy it is often convenient and reasonable to consider the price of traded goods relative to the price of one single traded good. This good is then called **numeraire**. In mathematical finance, the numeraire is usually the risk-free asset $(S_t^0)_{t\in[0,T]}$ and the price processes of the risky assets relative to the risk-free asset are then called discounted prices. In principle, we could choose any asset on the financial market as numeraire and, indeed, in some situation it is useful to do so, e.g. for the pricing some type of Exotic options. We make the following definition.

Definition 2.5. The **discounted price process** $\widehat{S} = (\widehat{S}^0, \dots, \widehat{S}^d)$ is defined by

$$\widehat{S}_t := \frac{1}{S_t^0} S_t = \left(1, \frac{S_t^1}{S_t^0}, \dots, \frac{S_t^d}{S_t^0}\right), \quad t \in [0,T].$$

The **discounted value** (or **capital**) **process** $\widehat{V}(\varphi) = (\widehat{V}_t(\varphi))_{t\in[0,T]}$ associated to φ is given by

$$\widehat{V}_t(\varphi) := \frac{1}{S_t^0} V_t(\varphi) = \varphi_t^T \widehat{S}_t, \quad t \in [0,T].$$

The self-financing condition of the trading strategy is not affected by considering discounted quantities instead of undiscounted ones.

Lemma 2.6. *A trading strategy φ is self-financing, i.e.*

$$\varphi \in \mathcal{L}(S) \quad \textit{and} \quad V_t(\varphi) = V_0(\varphi) + (\varphi \cdot S)_t, \quad t \in [0,T],$$

if and only if

$$\varphi \in \mathcal{L}(\widehat{S}) \quad \textit{and} \quad \widehat{V}_t(\varphi) = \widehat{V}_0(\varphi) + (\varphi \cdot \widehat{S})_t, \quad t \in [0,T].$$

Proof. Note that $\varphi \in \mathcal{L}(\widehat{S})$ if and only if $\varphi \in \mathcal{L}(S)$ for self-financing trading strategies φ. For the sake of brevity we skip the proof of this statement.

"$\Leftarrow$" We want to show that

$$\widehat{V}_t(\varphi) = \widehat{V}_0(\varphi) + (\varphi \cdot \widehat{S})_t \quad \text{implies} \quad V_t(\varphi) = V_0(\varphi) + (\varphi \cdot S)_t, \quad t \in [0,T].$$

In particular, by assumption the value process $\widehat{V}(\varphi)$ is an Itô integral and thus

$$\widehat{V}(\varphi) = \varphi^T \widehat{S} \quad \text{is an Itô process.}$$

Using the product rule for Itô integration, we get

$$\begin{aligned} V_t(\varphi) &= \varphi_t^T S_t = (\varphi_t^T \widehat{S}_t) S_t^0 \\ &= \varphi_0^T S_0 + \big((\varphi^T \widehat{S}) \cdot S^0\big)_t + \big(S^0 \cdot (\varphi^T \widehat{S})\big)_t + \langle \varphi^T \cdot \widehat{S}, S^0 \rangle_t \\ &= \varphi_0^T S_0 + \big((\varphi^T \widehat{S}) \cdot S^0\big)_t + \big(S^0 \cdot (\varphi^T \widehat{S})\big)_t \end{aligned}$$

since $(S_t^0)_{t\in[0,T]}$ is of finite variation. Using the self-financing condition

$$\varphi_t^T \widehat{S}_t = \widehat{V}_t(\varphi) = \varphi_0^T \widehat{S}_0 + (\varphi \cdot \widehat{S})_t, \quad t \in [0,T],$$

we conclude further that

$$\begin{aligned} V_t(\varphi) &= \varphi_0^T S_0 + \big((\varphi^T \widehat{S}) \cdot S^0\big)_t + \big(S^0 \cdot (\varphi_0^T \widehat{S}_0 + (\varphi \cdot \widehat{S}))\big)_t \\ &= \varphi_0^T S_0 + \big(\varphi \cdot (\widehat{S} \cdot S^0)\big)_t + \big((\varphi^T S^0) \cdot \widehat{S})\big)_t \\ &= \varphi_0^T S_0 + \int_0^t \varphi_s \, \mathrm{d}\big((\widehat{S} \cdot S^0)_s + (S^0 \cdot \widehat{S})_s\big), \end{aligned}$$

where he associativity of the Itô integral in the second last and last line. Since, by the product rule for Itô integration

$$(\widehat{S} \cdot S^0)_s + (S^0 \cdot \widehat{S})_s = \widehat{S}_s S_s^0 - \widehat{S}_0 S_0^0,$$

we arrive at

$$\begin{aligned} V_t(\varphi) &= \varphi_0^T S_0 + \int_0^t \varphi_s \, \mathrm{d}(\widehat{S}_s S_s^0 - \widehat{S}_0 S_0^0) \\ &= \varphi_0^T S_0 + \int_0^t \varphi_s \, \mathrm{d}S_s \\ &= V_0(\varphi) + (\varphi \cdot S)_t, \end{aligned}$$

which is the claimed self-financing condition.

"$\Rightarrow$" The converse direction follows by the same arguments replacing S and $\frac{1}{S^0}$ with $\widehat{S}$ and S^0, respectively. □

The next proposition verifies that for self-financing trading strategies the amount φ^0 invested in the risk-free asset is already uniquely determined by the other positions $(\varphi^1, \dots, \varphi^d)$ and the initial capital.

Lemma 2.7. *For every trading strategy $(\varphi^1, \dots, \varphi^d) \in \mathcal{L}((\widehat{S}^1, \dots, \widehat{S}^d))$ and every $V_0 \in \mathbb{R}$ there exists a unique process φ^0 such that*

$$\varphi = (\varphi^0, \varphi^1, \dots, \varphi^d) \in \mathcal{L}(S) \textit{ is self-financing with } V_0(\varphi) = V_0.$$

Proof. Due to Lemma 2.6, $\varphi = (\varphi^0, \varphi^1, \dots, \varphi^d)$ is self-financing if and only if

$$\begin{aligned}\varphi_t^0 + (\varphi^1, \dots, \varphi^d)_t^T \left(\widehat{S}^1, \dots, \widehat{S}^d\right)_t &= \widehat{V}_t(\varphi) \\ &= \widehat{V}_0 + \left(\varphi \cdot \widehat{S}\right)_t \\ &= \widehat{V}_0 + \left((\varphi^1, \dots, \varphi^d) \cdot (\widehat{S}^1, \dots, \widehat{S}^d)\right)_t,\end{aligned}$$

for $t \in [0,T]$, because of $\widehat{S}_t^0 = 1$. The previous equation holds if and only if

$$\varphi_t^0 = \widehat{V}_0 + \left((\varphi^1, \dots, \varphi^d) \cdot (\widehat{S}^1, \dots, \widehat{S}^d)\right)_t - (\varphi^1, \dots, \varphi^d)_t^T \left(\widehat{S}^1, \dots, \widehat{S}^d\right)_t, \quad t \in [0,T].$$

Furthermore, we have $\varphi \in \mathcal{L}(S)$ (again skipping the proof for the sake of brevity). □

2.2 First fundamental theorem of asset pricing

The fundamental idea of mathematical finance is the *absence of arbitrage* in an idealized world. Roughly speaking, absence of arbitrage means that there exist no risk-free gains.

Remark (*No arbitrage*)*.* There should be no risk-free gains with zero initial capital, that is, there should be no self-financing trading strategy φ such that

- $V_0(\varphi) = 0$,
- $V_T(\varphi) \geqslant 0$ $\mathbb{P}$-a.s.,
- $\mathbb{P}(V_T(\varphi) > 0) > 0$.

The concept of no arbitrage (or more general the no arbitrage principles, see Chapter 1) is the central building block of every reasonable model of a financial market. However, in a continuous-time setting the condition of self-financing is not sufficient to ensure no arbitrage and, therefore, we need to restrict the allowed trading strategies further.

Indeed, if trading in continuous time is possible, an investor can implement a trading strategy following the idea of the so-called "doubling strategy".

Example 2.8 (Doubling strategy - simplest framework)**.** Consider the fair coin tossing game: One tosses a coin. If heads comes up, the player is paid 2 times what she bets. If tails comes up, the player loses her bet. This game can be modeled by a sequence $(X_i)_{i\in\mathbb{N}}$ of i.i.d. random variables with

$$X_i = \begin{cases} 2 & \text{with } \mathbb{P}(X_i = 2) = \frac{1}{2} \\ 0 & \text{with } \mathbb{P}(X_i = 0) = \frac{1}{2} \end{cases}.$$

The "doubling strategy" for the player is to double her bets until the first time she wins (and then to stop playing), that is, her bet at the i-th game is

$$2^i \mathbb{1}_{\{X_0 = \dots = X_{i-1} = 0\}}, \quad \text{for } i \in \mathbb{N}.$$

This strategy leads always to the gain of 1 because

$$\sum_{i=0}^{\infty} 2^i \mathbb{1}_{\{X_0 = \dots = X_{i-1} = 0\}}(\omega) X_i(\omega) - \sum_{i=0}^{\tau} 2^i = 1, \quad \text{for almost all } \omega \in \Omega,$$

where

$$\tau = \inf\{i \in \mathbb{N} \,:\, X_i \neq 0\}.$$

While the doubling strategy is excluded in discrete time by allowing trading only at finitely many times $0, 1, \dots, T$, in continuous-time one is allowed to trade infinitely many times in the time interval $[0, T]$. Even in the Black-Scholes model one can construct a self-financing trading strategy leading to an arbitrage opportunity. The standard and economically meaningful way to exclude such type of doubling strategies in continuous-time finance is to introduce a borrowing constraint.

Definition 2.9. A self-financing trading strategy $\varphi = (\varphi_t)_{t\in[0,T]}$ is called **admissible** if there exist constants $c_{-1}, c_0, \dots, c_d \in \mathbb{R}_+$ such that

$$V_t(\varphi) \geq -c_{-1} - \sum_{i=0}^{d} c_i S_t^i, \quad t \in [0, T].$$

Note: one can always choose $c := \max\{c_{-1}, c_0, \dots, c_d\}$ instead of $c_{-1}, c_0, \dots c_d$. The set of all admissible trading strategies φ is denoted by $\mathcal{A}$.

Based on the notation of self-financing and admissibility, we define arbitrage opportunities.

Definition 2.10. An admissible trading strategy $\varphi = (\varphi_t)_{t\in[0,T]}$ is called **arbitrage** or **arbitrage opportunity** if

- $V_0(\varphi) = 0$,
- $V_T(\varphi) \geqslant 0$ $\mathbb{P}$-a.s.,
- $\mathbb{P}(V_T(\varphi) > 0) > 0$.

A financial market $(S^0, \dots, S^d)$ is said to be **arbitrage-free** (or to satisfy **no arbitrage (NA)**) if there exists no arbitrage opportunity on the market $(S^0, \dots, S^d)$.

As in mathematical finance in discrete time, no arbitrage could be equivalently characterized by the existence of so-called martingale measures and this equivalence turned out to be very fundamental in the analysis of financial markets. In our present continuous-time setting the picture becomes a bit more involved but the notion of martingale measure is again essential. Later we will see that martingale measures play an crucial role for the pricing of financial derivatives.

Definition 2.11. A probability measure Q on $(\Omega, \mathcal{F})$ is called **(local) martingale measure** if the discounted price process $\widehat{S} = (\widehat{S}^0, \dots, \widehat{S}^d)$ is a (local) Q-martingale, i.e.

$$(\widehat{S}_t^i)_{t\in[0,T]} \text{ is a (local) martingale w.r.t. } Q \quad \text{for } i = 0, \dots, d.$$

A (local) martingale measure Q is called **equivalent (local) martingale measure (E(L)MM)** if $Q \sim \mathbb{P}$.

Lecture 3

First let us observe that the discounted value process associated to an admissible trading strategy is a supermartingale under every equivalent local martingale measure.

Lemma 2.12. *Let $(\varphi_t)_{t\in[0,T]}$ be an admissible trading strategy and let Q be an ELMM. Then, the discounted value process*

$$\widehat{V}_t(\varphi) = \widehat{V}_0(\varphi) + (\varphi \cdot \widehat{S})_t, \quad t \in [0, T],$$

is a Q-supermartingale.

Proof. Let Q be an ELMM, i.e. $\widehat{S}$ is a local Q-martingale. If $(\varphi_t)_{t\in[0,T]}$ is an admissible trading strategy, $(\varphi_t)_{t\in[0,T]}$ is, in particular, self-financing, which implies that

$$\widehat{V}(\varphi) = \widehat{V}_0(\varphi) + (\varphi \cdot \widehat{S}) \quad \text{is a local } Q\text{-martingale.}$$

Since φ is admissible, there is some constant $c > 0$ such that

$$M := \widehat{V}(\varphi) + c\sum_{i=0}^{d} \widehat{S}^i \geq 0$$

is a non-negative continuous local martingale. Therefore, M is a Q-supermartingale, which completes the proof. □

It turns out that that the existence of an equivalent local martingale measure ensure that the financial market is still arbitrage-free.

Proposition 2.13. *If there exists an equivalent local martingale measure Q for $(S_t)_{t\in[0,T]}$, then the financial market $(S_t)_{t\in[0,T]}$ is arbitrage-free.*

Proof. Suppose that $\widehat{S}$ is a local Q-martingale. Let φ be an admissible trading strategy with

$$V_0(\varphi) = 0 \quad \text{and} \quad V_T(\varphi) \geq 0.$$

Since φ is admissible, by Lemma 2.12 the discounted value process $(\widehat{V}_t(\varphi))_{t\in[0,T]}$ is a Q-supermartingale. Therefore, one gets

$$0 \leq \mathbb{E}^Q[\widehat{V}_T(\varphi)] \leq \mathbb{E}^Q[\widehat{V}_0(\varphi)] = 0,$$

which reveals that

$$\widehat{V}_T(\varphi) = 0 \quad \text{and thus } V_T(\varphi) = 0.$$

In words, the financial market $(S_t)_{t\in[0,T]}$ is arbitrage-free. □

In case financial markets are modeled with discrete time $0, 1, \dots, T$, also the reverse direction of Proposition 2.13 holds true (both directions were called the "first fundamental theorem of asset pricing"). In our present setting, where the financial market is modeled with continuous time $[0,T]$, no arbitrage is not sufficient to imply the existence of an equivalent (local) martingale measure. To restore the existence of an equivalent (local) martingale measure, we shall exclude "no unbounded profit with bounded risk".

However, also in continuous-time financial modeling, one still can obtain a first fundamental theorem of asset pricing based on a weaker notion of "no arbitrage".

Definition 2.14. A sequence $(\varphi^n)_{n\in\mathbb{N}}$ of admissible trading strategies generates an **unbounded profit with bounded risk** if

- $V_0(\varphi^n) = x > 0$,
- $V_t(\varphi^n) \geqslant 0$, $t \in [0,T]$, $\mathbb{P}$-a.s.,
- $\lim_{m\to\infty} \left(\sup_{n\in\mathbb{N}} \mathbb{P}(V_T(\varphi^n) > m)\right) > 0$.

The market $(S_t)_{t\in[0,T]}$ satisfies **no unbounded profit with bounded risk (NUPBR)** if there exists no unbounded with bounded risk.

Remark. A sequence of admissible trading strategies whose value processes are always non-negative with initial capital $x > 0$ generates a unbounded profit with bounded risk if it never loses more than the strictly positive initial investment x and generates unboundedly large profits at time T with a strictly positive probability.

The notion of no arbitrage and no bounded profit with bounded risk can be combined to the notion of no free lunch with vanishing risk.

Definition 2.15. A non-negative random variable X with $\mathbb{P}(X > 0) > 0$ is called **free lunch with vanishing risk** if there exists a sequence $(\varphi^n)_{n\in\mathbb{N}}$ of admissible trading strategies and a sequence $(v_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$ with $v_n \to 0$ such that

$$V_0(\varphi^n) \leq v_n \quad \text{and} \quad V_T(\varphi^n) \geq X \quad \text{for all } n \in \mathbb{N}.$$

The market $(S_t)_{t\in[0,T]}$ satisfies **no free lunch with vanishing risk (NFLVR)** if there exists no such X on $(S_t)_{t\in[0,T]}$.

Remark 2.16. Every arbitrage opportunity $(\varphi_t)_{t\in[0,T]}$ is also a free lunch with vanishing risk, by setting $X := V_T(\varphi)$, $\varphi^n = \varphi$ and $v_n = 0$. In other words, NFLVR implies NA. The converse statement does not hold true.

Theorem 2.17. *The market $(S_t)_{t\in[0,T]}$ satisfies NA and NUPBR if and only if $(S_t)_{t\in[0,T]}$ satisfies NFLVR.*

For a proof see Delbaen and Schachermayer (1994, Corollary 3.8). The condition of NFLVR allows to recover the first fundamental theorem of asset pricing (FTAP): NFLVR is indeed equivalent to the existence of an equivalent local martingale measure.

Theorem 2.18 (First fundamental theorem of asset pricing)**.** *The market $(S_t)_{t\in[0,T]}$ satisfies NFLVR if and only if there exists an EMM Q.*

The proof of the first fundamental theorem of asset pricing is rather involved and lengthy. It is mainly based on functional analysis, in particular, on the Hahn-Banach theorem. We do here only the "easy" direction.

Elements of the proof. "$\Leftarrow$" Suppose there exists an EMM Q. Furthermore, suppose that X is a free lunch with vanishing risk, i.e. there exists a sequence $(\varphi^n)_{n\in\mathbb{N}}$ of admissible trading strategies and a sequence $(v_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}_+$ with $v_n \to 0$ such that

$$V_0(\varphi^n) \leq v_n \quad \text{and} \quad V_T(\varphi^n) \geq X \quad \text{for all } n \in \mathbb{N}.$$

We want to show that X is not a free lunch with vanishing risk. For this purpose we define

$$\widehat{X} := \frac{X}{S_T^0}.$$

Since $(\widehat{V}_t(\varphi^n))_{t\in[0,T]}$ is a Q-supermartingale by Lemma 2.12, we get

$$\mathbb{E}^Q[\widehat{X}] \leq \mathbb{E}^Q[\widehat{V}_T(\varphi^n)] \leq \mathbb{E}^Q[\widehat{V}_0(\varphi^n)] \leq \frac{v_n}{S_0^0} \to 0 \quad \text{as} \quad n \to \infty,$$

which reveals $\mathbb{E}^Q[\widehat{X}] = 0$ (recall $\widehat{X} \geq 0$) and thus $Q(\widehat{X} = 0) = 1$ as $\widehat{X} \geq 0$. However, since $Q \sim \mathbb{P}$ this contradicts

$$\mathbb{P}(X > 0) > 0,$$

that means, there does not exist a free lunch with vanishing risk.

"$\Rightarrow$" This direction is much more involved and outside the scope of the course. A proof can be found in Delbaen and Schachermayer (1994) or in (Jarrow, 2021, Section 2.6, Theorem 2.5). $\square$

Remark. In general, the first fundamental theorem of asset pricing states that the NFLVR is equivalent to the existence of an equivalent *local* martingale measure. In order to ensure that these local martingale measures are indeed martingale measures, one needs to add a further assumptions on the underlying financial market and chose the class of admissible trading strategies carefully, as we have already done.

2.3 Pricing and hedging of financial derivatives

The aim of this subsection is to study the *pricing and hedging of financial derivative* X on our financial market $(S_t)_{t\in[0,T]}$, assuming that the $(S_t)_{t\in[0,T]}$ satisfies no free lunch with vanishing risk (NFLVR).

Definition 2.19. A **(contingent) claim** (or a **financial derivative**) X is an $\mathcal{F}_T$-measurable random variable. The **discounted claim** $\widehat{X}$ is given by

$$\widehat{X} := \frac{X}{S_T^0}.$$

Here T is called the **maturity** of the claim X and $\widehat{X}$, respectively.

While from a mathematical perspective financial derivatives are nothing else than random variables, on real financial markets they are contracts whose future payoff depends on the price processes of one or several risky assets.

Example 2.20 (Financial Derivatives)**.** The most common ones are

- **Call options**: $(S_T^1 - K)^+ := \max\{0, S_T^1 - K\}$ with *strike price* $K \geq 0$ and maturity T,
- **Put options**: $(K - S_T^1)^+$ with *strike price* $K \geq 0$ and maturity T.

Financial derivatives of the form $g(S_T)$, where $g\colon \mathbb{R} \to \mathbb{R}$, are often referred to as *"Vanilla options"*. Financial derivatives with a more complex payoff are usually referred to as *"Exotic options"*, for example:

- **Barrier options** are options whose existence depends upon the underlying asset's price hitting (or not) a barrier, e.g., an up-and-out Put option

$$(K - S_T^1)^+ \mathbb{1}_{\{\sup_{0\leq t\leq T} S_t^1 \leq B\}}, \quad \text{for some } B \geq 0,$$

- **Lookback options** are options depending on the

$$\text{running maximum } M_T := \sup_{0\leq t\leq T} S_t^1 \quad \text{or} \quad \text{running minimum } m_T := \inf_{0\leq t\leq T} S_t^1$$

of the underlying price process $(S_t^1)_{t\in[0,T]}$, e.g., lookback rate Put options

$$(M_t - K)^+ \quad \text{or} \quad (m_t - K)^+,$$

- **Asian options** are options depending on some type of average (arithmetic, geometric, ...) of the underlying price process $(S_t^1)_{t\in[0,T]}$, e.g., a continuously sampled arithmetic-average rate call
$$\left(\frac{1}{T}\int_0^T S_t^1\,\mathrm{d}t - K\right)^+.$$

The list of financial derivatives provided in Example 2.20 is by no means exhausting. We will discuss some Exotic options in more detail later in Section 3.3.

Lecture 4

A main task of mathematical finance is to determine an arbitrage-free price of a contingent claim X as well as to find a trading strategy φ which eliminates the risk caused by selling the claim X. This hedging problem can be conveniently solved for *replicable claims*.

The basic idea to determine the arbitrage-free price of a replicable claim X is the well-known *law of one price*: "two contracts providing the identical payoff have the same price". Consequently, the price of the claim X should be equal to the initial capital required to implement its replicating trading strategy. However, while on arbitrage-free financial markets in a discrete-time setting, the value process $(V_t(\varphi))_{t\in[0,T]}$ was completely determined by its terminal value $V_T(\varphi)$. This is not true on arbitrage-free markets in continuous-time. Recall, that the discounted value process $(\widehat{V}_t(\varphi))_{t\in[0,T]}$ is in general "only" a supermartingale under every EMM Q.

Example. For example in the Black-Scholes model (with $r = 0$) there exists an admissible trading strategy φ such that
$$V_0(\varphi) = 0 \quad \text{and} \quad V_T(\varphi) = -1.$$
Moreover, the admissible trading strategy $\psi = (-1, 0)$ satisfies
$$V_0(\psi) = -1 \quad \text{and} \quad V_T(\psi) = -1.$$

In other words, the initial value V_0 of a value process is not uniquely determined by its terminal value V_T. As a consequence, the law of one price does not hold in general. In order to ensure the law of one price, we need an stronger notion of admissibility.

Definition 2.21. A self-financing trading strategy φ is called **strongly admissible** if the trading strategies φ and $-\varphi$ are both admissible.

While strongly admissible trading strategies ensure the law of one price, as we will see below, they are economically speaking not really justified as strongly admissibility requires an upper bound of the corresponding value processes. We will only use the notion of strongly admissibility if it is strictly necessary.

Lemma 2.22. *Suppose that $(S_t)_{t\in[0,T]}$ satisfies NFLVR. If φ is a strongly admissible trading strategy, then the discounted value process $(\widehat{V}_t(\varphi))_{t\in[0,T]}$ is a Q-martingale, for every EMM Q.*

Proof. Let Q be an EMM. Since φ is strongly admissible, Lemma 2.12 says that $\widehat{V}(\varphi)$ and $\widehat{V}(-\varphi)$ are Q-supermartingales. Note that
$$\widehat{V}_t(-\varphi) = (-\varphi_t)^T\widehat{S}_t = -\varphi_t^T\widehat{S}_t = -\widehat{V}_t(\varphi), \quad t\in[0,T].$$

Hence, for $s < t$ we obtain

$$\mathbb{E}^Q[\widehat{V}_t(\varphi)|\mathcal{F}_s] \le \widehat{V}_s(\varphi), \quad \text{and} \quad -\mathbb{E}^Q[\widehat{V}_t(\varphi)|\mathcal{F}_s] = \mathbb{E}^Q[\widehat{V}_t(-\varphi)|\mathcal{F}_s] \le \widehat{V}_s(-\varphi) = -\widehat{V}_s(\varphi),$$

which implies that $\widehat{V}(\varphi)$ is a Q-martingale. □

Proposition 2.23 (Law of one price)**.** *Suppose that $(S_t)_{t\in[0,T]}$ satisfies NA. Let φ and ψ two be strongly admissible trading strategies with $V_T(\varphi) = V_T(\psi)$. Then, one has*

$$V_t(\varphi) = V_t(\psi), \quad \textit{for all } t \in [0,T].$$

In particular, if φ is a strongly admissible trading strategy with $V_T(\varphi) = S_T^i$ for some $i \in \{0,\dots,d\}$, then $V_t(\varphi) = S_t^i$ for all $t \in [0,T]$.

Proof. Proof by contradiction: Assume there exists a $t \in [0,T]$ such that

$$\mathbb{P}(V_t(\varphi) \neq V_t(\psi)) > 0.$$

Without loss of generality let us assume that

$$\mathbb{P}(A) > 0 \quad \text{with } A := \{V_t(\psi) > V_t(\varphi)\}.$$

In order to obtain a contradiction, we want to construct an arbitrage opportunity $\delta = (\delta^0, \delta^1, \dots, \delta^d)$. Since A is $\mathcal{F}_t$-measurable, we can define the trading strategy

$$(\delta^1,\dots,\delta^d)_s := \begin{cases} 0 & \text{if } s \le t \\ ((\varphi^1,\dots,\varphi^d) - (\psi^1,\dots,\psi^d))\mathbb{1}_A & \text{if } s > t \end{cases}.$$

By Proposition 2.7, we can turn $(\delta^1,\dots,\delta^d)$ into a self-financing trading strategy $\delta = (\delta^0, \delta^1, \dots, \delta^d)$ with $V_0(\delta) = 0$.

Hence, we get

$$\widehat{V}_0(\delta) = 0, \quad \widehat{V}_s(\delta) = 0 \quad \text{for } s \le t$$

and for $s > t$,

$$\begin{aligned} \widehat{V}_s(\delta) &= \left((\varphi - \psi)\mathbb{1}_{A\times(t,T]} \cdot \widehat{S}\right)_s = \mathbb{1}_{A\times(t,T]}((\varphi-\psi)\cdot\widehat{S})_s \\ &= (((\varphi-\psi)\cdot\widehat{S})_s - ((\varphi-\psi)\cdot\widehat{S})_t)\mathbb{1}_A \\ &= (\widehat{V}_s(\varphi) - \widehat{V}_s(\psi) - \widehat{V}_t(\varphi) + \widehat{V}_t(\psi))\mathbb{1}_A \\ &\ge (\widehat{V}_s(\varphi) - \widehat{V}_s(\psi))\mathbb{1}_A \end{aligned} \tag{2.1}$$

as $\left(-\widehat{V}_t(\varphi) + \widehat{V}_t(\psi)\right)\mathbb{1}_A \ge 0$. Hence, $(\delta_t)_{t\in[0,T]}$ is admissible since φ and ψ are strongly admissible.

Furthermore, using (2.1) and $\widehat{V}_T(\varphi) = \widehat{V}_T(\psi)$, we get

$$\begin{aligned} \widehat{V}_T(\delta) &= (\widehat{V}_T(\varphi) - \widehat{V}_T(\psi) - \widehat{V}_t(\varphi) + \widehat{V}_t(\psi))\mathbb{1}_A \\ &= (\widehat{V}_t(\psi) - \widehat{V}_t(\varphi))\mathbb{1}_A, \end{aligned}$$

which implies

$$V_T(\delta) > 0 \quad \text{on } A \quad \text{and} \quad V_T(\delta) = 0 \quad \text{on } A^c.$$

Therefore, δ is an arbitrage opportunity.

For the second statement consider the trading strategy ψ with

$$\psi^j := \begin{cases} 1 & \text{for } j = i \\ 0 & \text{otherwise} \end{cases}, \quad \text{for } j = 1, \dots, d.$$

□

Based on the law of price we can study the pricing problem of a replicable claim.

Definition 2.24. A claim X is called **replicable** (or **attainable**) if there exists an strongly admissible trading strategy $(\varphi)_{t\in[0,T]}$ such that

$$X = V_T(\varphi).$$

The trading strategy $(\varphi)_{t\in[0,T]}$ is called **replicating trading strategy** of X.

Selling a claim X on the financial market $(S^0, \dots, S^d)$ creates a new tradable asset, say, $(S^{d+1})_{t\in[0,T]}$ denotes the price process of the claim X. To ensure that the now extended market $(S^0, \dots, S^d, S^{d+1})$ still fulfills NFLVR, we know by the first fundamental theorem of asset pricing that the discounted price process $(\widehat{S}^{d+1}_t)_{t\in[0,T]}$ has to be a Q-martingale for at least one EMM Q. For replicable claims X the corresponding price process S^{d+1} is even unique.

Proposition 2.25. *Suppose that the financial market $S = (S_t)_{t\in[0,T]}$ satisfies NFLVR. Let X be a non-negative replicable claim X with replicating trading strategy φ. Then, one has:*

(i) There exists a unique price process $(S^{d+1}_t)_{t\in[0,T]}$ with $S^{d+1}_T = X$ such that the extended market

$$(S^0_t, \dots, S^d_t, S^{d+1}_t)_{t\in[0,T]} \quad \textit{satisfies NFLVR.}$$

In particular, $(S^0_t, \dots, S^d_t, S^{d+1}_t)_{t\in[0,T]}$ is arbitrage-free and

$$S^{d+1}_t = V_t(\varphi) \geq 0 \quad \textit{for all} \quad t \in [0,T].$$

(ii) If Q is an EMM for the market $(S^0, \dots, S^d)$, then $\widehat{X} \in L^1(Q)$ and

$$\widehat{S}^{d+1}_t = \mathbb{E}^Q[\widehat{X}|\mathcal{F}_t], \quad t \in [0,T],$$

for $(S^{\mathrm{d}+1}_t)_{t\in[0,T]}$ as in (i).

Proof. (i) Existence: Let φ be a replicating trading strategy of X and Q be an EMM for the market $(S^0, \dots, S^d)$, which exists by the 1. FTAP. Set

$$S^{d+1}_t := V_t(\varphi), \quad t \in [0,T].$$

Since φ is strongly admissible and $\psi := 0$ is also a strongly admissible trading strategy with

$$V_T(\varphi) = X \geq 0 = V_T(\psi),$$

the law of one price (Proposition 2.23) implies

$$V_t(\varphi) \geq V_t(\psi) = 0, \quad \text{for all } t \in [0,T],$$

i.e. $S_t^{d+1} \geq 0$. Since φ is strongly admissible, we also know by Lemma 2.22, that

$$\widehat{V}(\varphi) = \widehat{S}^{d+1} \quad \text{is a } Q\text{-martingale.}$$

Hence, Q is an EMM for the extended market

$$(S^0, \ldots, S^d, S^{d+1})$$

and thus, by the 1. FTAP the extended market satisfies NFLVR.

Uniqueness: The uniqueness of $(S_t^{d+1})_{t\in[0,T]}$ follows directly by the law of one price (Proposition 2.23) since by assumption $S_T^{d+1} = X$.

(ii) Since $\widehat{S}^{d+1}$ is a Q-martingale, we have

$$\widehat{X} = \widehat{S}_T^{d+1} \in L^1(Q)$$

and

$$\widehat{S}_t^{d+1} = \mathbb{E}^Q[\widehat{S}_T^{d+1}|\mathcal{F}_t] = \mathbb{E}^Q[\widehat{X}|\mathcal{F}_t], \quad t \in [0,T].$$

□

Definition 2.26. Let X be a replicable claim with replicating trading strategy φ. The process $(V_t(X))_{t\in[0,T]}$ such that

$$V_t(X) := S_t^{d+1}, \quad t \in [0,T],$$

is called **value** (or **price**) **process** of X, where $(S_t^{d+1})_{t\in[0,T]}$ is defined as in Proposition 2.25. The corresponding replicating strategy φ is also called **hedging** (or **replicating) strategy**.

Remark 2.27. If a bank sells the replicable claim X and invests according to the replicating trading strategy φ, the bank's terminal wealth is

$$-X + V_T(\varphi) = 0,$$

that means, the bank has no risk of making losses. In other words, the trading strategy φ hedges the bank against the risk of losses.

In general, not every claim is replicable on a financial market satisfying NFLVR.

Definition 2.28. A financial market S is called **complete** if every bounded discounted claim $\widehat{X}$ (i.e. $\widehat{X}$ is bounded) is replicable.

The completeness of a financial market can be characterized in terms of equivalent martingale measures.

Theorem 2.29 (Second fundamental theorem of asset pricing (2. FTAP))**.** *Suppose that the financial market $S = (S_t)_{t\in[0,T]}$ satisfies NFLVR. The following conditions are equivalent:*

(i) The market S is complete.

(ii) There exists a unique equivalent martingale measure Q.

In this case, every claim X such that

$$|\widehat{X}| \le c\Big(\sum_{i=0}^{d} \widehat{S}^i_T\Big) \quad \textit{for some constant } c \ge 0,$$

is replicable. (Even all claims X with $\widehat{X} \in L^1(Q)$ are replicable.)

Elements of the proof. "⇒" The existence of a unique EMM Q is the easy direction: By the 1. FTAP there exists an EMM Q as the market S satisfies NFLVR. For $A \in \mathcal{F}$ we define the discounted claim $\widehat{X} = \mathbb{1}_A$, which is non-negative and bounded. Since the market is complete, $\widehat{X}$ is replicable, that is, there exists a strongly admissible trading strategy φ such that $\widehat{X} = \widehat{V}_T(\varphi)$. By Proposition 2.25

$$\widehat{V}_0(\varphi) = \mathbb{E}^Q[\widehat{V}_T(\varphi)] = \mathbb{E}^Q[\widehat{X}] = Q(A)$$

for all EMM Q. Because $\widehat{V}_0(\varphi)$ does not dependent on Q, we deduce that Q is unique.

"⇐" That the existence of an EMM Q implies the NFLVR is the content of the 1. FTAP. That the uniqueness of the EMM implies the completeness of the market is more involved and we do not present here the proof. It requires advanced results from stochastic analysis. In particular, the key ingredient of the proof is that the uniqueness of the EMM Q ensures that the *martingale representation theorem* holds for the martingale $\widehat{S}$. A more general version of the martingale representation theorem, covering continuous martingales, can be found in the book Revuz and Yor (1999). We drop further details here and refer again to Delbaen and Schachermayer (1994) for a complete poof. □

Lecture 5

As a corollary of the second fundamental theorem of asset pricing, we obtain a formula for the price of claims on complete markets satisfying NFLVR.

Corollary 2.30. *Suppose that the financial market $(S^0, \dots, S^d)$ satisfies NFLVR and it is complete. Let X be a claim such that*

$$|\widehat{X}| \le c\Big(\sum_{i=0}^{d} \widehat{S}^i_T\Big) \quad \textit{for some constant } c \ge 0,$$

or, more general, such that $\widehat{X} \in L^1(Q)$. Then, the value process $(V_t(X))_{t\in[0,T]}$ is given by

$$V_t(X) = S^0_t \mathbb{E}^Q\Big[\frac{X}{S^0_T}\Big|\mathcal{F}_t\Big], \quad t \in [0,T],$$

where Q is the unique equivalent martingale measure for $(S^0, \dots, S^d)$.

Proof. By the 2. FTAP the claim X is replicable and there exists a unique EMM Q. Hence, the assertion follows by Proposition 2.25. □

3 Pricing and hedging in Black-Scholes models

In this section we study in more detail the one-dimensional Black-Scholes model and the pricing and hedging problem of financial derivatives (Vanilla and Exotic options) within the Black-Scholes framework. However, the theory of Chapter 2 and the ideas, which we shall develop in this chapter, also apply in more general multi-dimensional diffusion models for complete markets or, in other words, in complete financial markets modeled by stochastic differential equations. The one-dimensional Black-Scholes just serves as most famous prototypical example of such complete markets.

3.1 Black-Scholes model

Let us recall that the (one-dimensional, i.e. $d = 1$) Black-Scholes model $(S_t^0, S_t^1)_{t\in[0,T]}$ consists of

- a risk-free asset $(S_t^0)_{t\in[0,T]}$ given by
$$S_t^0 := S_0^0 \exp(rt) = S_0^0 \mathcal{E}(rI)_t, \quad t \in [0, T],$$
where $S_0^0 > 0$, $r \in \mathbb{R}$ is the interest rate and $I = (I_t)_{t\in[0,T]}$ is the identity process $I_t = t$,
- a risky asset $(S_t^1)_{t\in[0,T]}$ given by
$$S_t^1 := S_0^1 \exp(\widetilde{\mu}t + \sigma W_t) = S_0^1 \mathcal{E}(\mu I + \sigma W)_t, \quad t \in [0, T],$$
where $S_0^1 > 0$, $\mu \in \mathbb{R}$ is the drift parameter, $\sigma > 0$ is the volatility parameter,
$$\widetilde{\mu} := \mu - \frac{1}{2}\sigma^2$$
and $(W_t)_{t\in[0,T]}$ is a Brownian motion.

Throughout this chapter we will always consider the Black-Scholes model as our underlying financial market.

Remark 3.1. Recall that $(S_t^1)_{t\in[0,T]}$ is the solution to the stochastic differential equations (SDEs)
$$\mathrm{d}S_t^1 = \mu S_t^1 \,\mathrm{d}t + \sigma S_t^1 \,\mathrm{d}W_t, \quad S_0^1 = S_0^1, \quad t \in [0, T],$$
which can be heurstically re-formulated to
$$\frac{\mathrm{d}S_t^1}{S_t^1} = \mu \,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \quad S_0^1 = S_0^1, \quad t \in [0, T].$$
From the last formulation one can see that the expected return during the time "dt" is
$$\mathbb{E}\Big[\frac{\mathrm{d}S_t^1}{S_t^1}\Big] = \mu \,\mathrm{d}t.$$
In words, μ models the expected returns/trend and σ the volatility of the returns.

As a first step we need to verify that the Black-Scholes model is indeed a reasonable model for a financial market, which means, we want to check that it fulfills no free lunch with vanishing risk (NFLVR).

Theorem 3.2.

(i) The Radon-Nikodym density
$$\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} := \mathcal{E}\Big(-\frac{\mu - r}{\sigma}W\Big)_T := \exp\Big(-\frac{\mu - r}{\sigma}W_T - \frac{1}{2}\Big(\frac{\mu - r}{\sigma}\Big)^2 T\Big)$$
defines an equivalent martingale measure $Q \sim \mathbb{P}$.

(ii) The discounted price process $(\widehat{S}_t^1)_{t\in[0,T]}$ is given by

$$\widehat{S}_t^1 = \widehat{S}_0^1 \exp\left(\sigma\widetilde{W}_t - \frac{\sigma^2}{2}t\right) = \widehat{S}_0^1\mathcal{E}(\sigma\widetilde{W})_t$$

for a Q-Brownian motion $(\widetilde{W}_t)_{t\in[0,T]}$ with

$$\widetilde{W}_t = W_t + \frac{\mu - r}{\sigma}t, \quad t \in [0,T].$$

(iii) The Black-Scholes model $(S_t^0, S_t^1)_{t\in[0,T]}$ satisfies NFLVR. In particular, the Black-Scholes model $(S_t^0, S_t^1)_{t\in[0,T]}$ is arbitrage-free.

Proof. (i)+(ii) The stochastic process $(L_t)_{t\in[0,T]}$, given by

$$L_t := \mathcal{E}\Big(-\frac{\mu - r}{\sigma}W\Big)_t, \quad t \in [0,T],$$

is a continuous strictly positive martingale with $\mathbb{E}[L_T] = \mathbb{E}[L_0] = 1$ (check!). Hence, $(L_t)_{t\in[0,T]}$ is a density process for an equivalent probability measure $Q \sim \mathbb{P}$ and by Girsanov's theorem the stochastic process

$$\widetilde{W}_t = W_t + \frac{\mu - r}{\sigma}t = W_t - \int_0^t \Big(-\frac{\mu - r}{\sigma}\Big)\,\mathrm{d}s, \quad t \in [0,T],$$

is a Q-Brownian motion. Furthermore, recalling $\widetilde{\mu} := \mu - \frac{1}{2}\sigma^2$, one has

$$\begin{aligned}
\widehat{S}_t^1 &= \widehat{S}_0^1 \exp\left(\sigma W_t + (\widetilde{\mu} - r)t\right) \\
&= \widehat{S}_0^1 \exp\left(\sigma\widetilde{W}_t - \sigma\frac{\mu - r}{\sigma}t + (\widetilde{\mu} - r)t\right) \\
&= \widehat{S}_0^1 \exp\left(\sigma\widetilde{W}_t + (\widetilde{\mu} - \mu)t\right) \\
&= \widehat{S}_0^1 \exp\left(\sigma\widetilde{W}_t - \frac{\sigma^2}{2}t\right) = \widehat{S}_0^1\mathcal{E}(\sigma\widetilde{W})_t,
\end{aligned}$$

which completes the proof of (ii). Moreover, $(\widehat{S}_t^1)_{t\in[0,T]}$ is a continuous martingale, which completes the proof of (i).

(iii) Since there exists an EMM Q, the Black-Scholes model $(S_t^0, S_t^1)_{t\in[0,T]}$ fulfills NFLVR by the first FTAP. Hence, the Black-Scholes model is, in particular, arbitrage-free. □

The Black-Scholes model is not only arbitrage-free and there is even no free lunch with vanishing risk, it is also a model for a complete financial market, that means, every bounded discounted claim can be replicated by trading on the financial market.

Lemma 3.3. *The one-dimensional Black-Scholes model $(S_t^0, S_t^1)_{t\in[0,T]}$ is complete.*

Proof. Let X be a claim (i.e. random variable) such that $|\widehat{X}| \le m \in \mathbb{R}_+$ and Q is the EMM defined in Theorem 3.2. By the martingale representation theorem for a Brownian motion, there exists an $H \in \mathcal{H}^2(\widetilde{W})$ such that

$$\widehat{X} = \mathbb{E}^Q[\widehat{X}] + (H \cdot \widetilde{W})_T.$$

Let us define the stochastic process $(\widehat{S}^2_t)_{t\in[0,T]}$ by

$$\widehat{S}^2_t := \widehat{V}_t(X) = \mathbb{E}^Q[\widehat{X}|\mathcal{F}_t], \quad t \in [0,T],$$

which is a Q-martingale. Hence, we have

$$\widehat{S}^2_t = \widehat{S}^2_0 + (H \cdot \widetilde{W})_t, \quad t \in [0,T].$$

Setting $\varphi^1 := \frac{H}{\sigma \widehat{S}^1}$, we observe that

$$\widehat{S}^2_0 + \left(\varphi^1 \cdot \widehat{S}^1\right)_t = \widehat{S}^2_0 + \left((\varphi^1 \sigma \widehat{S}^1) \cdot \widetilde{W}\right)_t = \widehat{S}^2_0 + \left(H \cdot \widetilde{W}\right)_t = \widehat{S}^2_t, \quad t \in [0,T],$$

since $(\widehat{S}^1_t)_{t\in[0,T]}$ is a stochastic exponential satisfying

$$\widehat{S}^1_t = \widehat{S}^1_0 + \left(\widehat{S}^1 \cdot (\sigma \widetilde{W})\right)_t, \quad t \in [0,T].$$

Now we enhance φ^1 to a self-financing trading strategy $\varphi = (\varphi^0, \varphi^1)$ by Proposition 2.7.

Because of $\widehat{V}(\varphi) = \widehat{S}^2$ we have

$$|\widehat{V}(\varphi)| \le m \quad \text{and} \quad \widehat{V}_T(\varphi) = \widehat{X},$$

which means that $(\varphi_t)_{t\in[0,T]}$ is a strongly admissible trading strategy replicating $\widehat{X}$. Hence, the one-dimensional Black-Scholes model is complete. □

A useful observation for studying the hedging problem of financial derivatives, is to understand the growth rate of locally risk-free capital processes.

Definition 3.4. Let φ be an admissible trading strategy. A value (or capital) process $(V_t(\varphi))_{t\in[0,T]}$ is **locally risk-free** if there exists an adapted and integrable process $(\beta_t)_{t\in[0,T]}$ such that

$$V_t(\varphi) = V_0(\varphi) + \int_0^t \beta_u \,\mathrm{d}u, \quad t \in [0,T].$$

Lemma 3.5. *Every locally risk-free value process $(V_t(\varphi))_{t\in[0,T]}$ satisfies*

$$V_t(\varphi) = V_0(\varphi) + \int_0^t r V_u(\varphi) \,\mathrm{d}u, \quad t \in [0,T].$$

Proof. See Exercise 3.3 on Problem sheet 3. □

3.2 Pricing and hedging of Vanilla options

As we have seen, the one-dimensional Black-Scholes model satisfies NFLVR and is a complete model for a financial market. These properties ensure that it is reasonable to study the pricing and hedging problem for financial options in this Black-Scholes setting. The purpose of this subsection is to derive the prices of simple European vanilla options $X = g(S^1_T)$ for a function $g\colon (0,\infty) \to \mathbb{R}_+$. To do so, there are two natural and commonly used approaches:

Approach 1: PDE approach Heuristic idea: As the Black-Scholes model is complete there exists an strongly admissible trading strategy $\varphi = (\varphi^0, \varphi^1)$ such that

$$V_t(X) = V_t(\varphi), \quad t \in [0, T].$$

In order to find φ (and thus $V_t(X)$ by the law of one price), keeping the Markov structure in mind, we make the ansatz that there twice continuously differentiable $v\colon [0, T) \times (0, \infty) \to \mathbb{R}$ such that

$$V_t(\varphi) = V_t(X) = v(t, S_t^1).$$

Applying Itô formula to get an partial differential equation (PDE) determining v and $\varphi_t^1 = \frac{\partial}{\partial x} v(t, S_t^1)$ (delta-hedging strategy).

Approach 2: Martingale approach Heuristic idea: There exists a unique EMM Q for the Black-Scholes model and thus the arbitrage-free price process $(V_t(X))_{t\in[0,T]}$ of the claim X is given by

$$V_t(X) = S_t^0 \mathbb{E}^Q[\widehat{X}|\mathcal{F}_t], \quad t \in [0, T].$$

Now we calculate the conditional expectation, to get a twice continuously differentiable $v\colon [0, T) \times (0, \infty) \to \mathbb{R}$ such that

$$V_t(X) = v(t, S_t^1), \quad t \in [0, T].$$

To the corresponding hedging strategy $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$, we apply Itô formula to $\widehat{V}_t(X)$ to get $\varphi_t^1 = \frac{\partial}{\partial x} v(t, S_t^1)$ (delta-hedging strategy).

In the following we shall implement in the martingale approach to derive the prices and hedging strategies of simple European vanilla options.

Proposition 3.6. *Let $X = g(S_T^1)$ be a claim with*

$$g\colon (0, \infty) \to \mathbb{R}_+ \quad \text{such that } g(x) \le c(1 + x) \text{ for some } c \ge 0.$$

Then, one has:

(i) X is replicable.

(ii) The value process $(V_t(X))_{t\in[0,T]}$ satisfies

$$V_t(X) = v(t, S_t^1), \quad t \in [0, T],$$

with

$$v(t, x) := \exp(-r(T - t)) \int_{\mathbb{R}} g(\exp(y)) \varphi_{\log(x)+(r-\sigma^2/2)(T-t),\sigma^2(T-t)}(y) \, \mathrm{d}y, \tag{3.1}$$

where φ_{μ,σ^2} denotes the density function of normal distribution $\mathcal{N}(\mu, \sigma^2)$, i.e.

$$\varphi_{\mu,\sigma^2}(y) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(- \frac{1}{2} \Big(\frac{\mu - y}{\sigma} \Big)^2 \Big), \quad y \in \mathbb{R}.$$

(iii) The value function $v\colon [0, T) \times (0, \infty) \to \mathbb{R}$ is twice continuously differentiable.

Proof. (i) Since
$$0 \le g(S_T^1) \le c(1 + S_T^1) \le c(S_T^0 + S_T^1),$$
the claim $X = g(S_T^1)$ is replicable by the 2. FTAP (Theorem 2.29). Lecture 6

(ii) By Corollary 2.30, the discounted value process $(\widehat{V}_t(X))_{t\in[0,T]}$ is given by
$$\begin{aligned}\widehat{V}_t(X) &= \mathbb{E}^Q[\widehat{X}|\mathcal{F}_t]\\ &= \exp(-rT)\mathbb{E}^Q[g(S_T^1)|\mathcal{F}_t]\\ &= \exp(-rT)\mathbb{E}^Q\Big[g\Big(S_t^1 \exp\Big((r-\frac{\sigma^2}{2})(T-t)+\sigma(\widetilde{W}_T-\widetilde{W}_t)\Big)\Big)\Big|\mathcal{F}_t\Big], \quad t\in[0,T].\end{aligned}$$

Because S_t^1 is $\mathcal{F}_t$-measurable, $\widetilde{W}_T - \widetilde{W}_t$ is independent of $\mathcal{F}_t$ and $\widetilde{W}_T - \widetilde{W}_t \sim \mathcal{N}(0, T-t)$ under Q, we get
$$\begin{aligned}\widehat{V}_t(X) &= \exp(-rT)\int_{\mathbb{R}} g(S_t^1 \exp((r-\frac{\sigma^2}{2})(T-t)+\sigma y))\varphi_{0,T-t}(y)\,\mathrm{d}y\\ &= \exp(-rT)\int_{\mathbb{R}} g(\exp(\log(S_t^1)+(r-\frac{\sigma^2}{2})(T-t)+\sigma y))\varphi_{0,T-t}(y)\,\mathrm{d}y.\end{aligned}$$

Using the substitution
$$z = \log(S_t^1) + (r-\frac{\sigma^2}{2})(T-t)+\sigma y,$$
we obtain
$$\widehat{V}_t(X) = \exp(-rT)\int_{\mathbb{R}} g(\exp(z))\varphi_{\log(S_t^1)+\left(r-\frac{\sigma^2}{2}\right)(T-t),\sigma^2(T-t)}(z)\,\mathrm{d}z.$$

(iii) Using the dominated convergence theorem, the differentiability of v follows by interchanging integration and differentiation. □

Using the value function v associated to the vanilla option $g(S_T^1)$, we can also explicitly derive its replicating trading strategy. Recall, that such a replicating trading strategy has to exist because the Black-Scholes model is complete.

Proposition 3.7. *Let $X = g(S_T^1)$ be a claim with*
$$g\colon (0,\infty) \to \mathbb{R}_+ \quad \text{such that } g(x) \le c(1+x) \text{ for some } c \ge 0$$
and associated value function $v\colon [0,T)\times(0,\infty)\to\mathbb{R}$ as defined in (3.1). Then, one has:

(i) The replicating trading strategy $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ of X is given by
$$\begin{aligned}\varphi_t^0 &:= \exp(-rt)\Big(v(t,S_t^1) - S_t^1\frac{\partial}{\partial x}v(t,S_t^1)\Big),\\ \varphi_t^1 &:= \frac{\partial}{\partial x}v(t,S_t^1).\end{aligned}$$

(ii) The value function v fulfills the Black-Scholes PDE
$$\frac{\partial}{\partial t}v(t,x) + rx\frac{\partial}{\partial x}v(t,x) + \frac{1}{2}\sigma^2x^2\frac{\partial^2}{\partial x^2}v(t,x) - rv(t,x) = 0,$$
for $t \in [0,\infty)\times(0,\infty)$.

Proof. (i) Proposition 3.6 (ii) implies that

$$\widehat{V}_t(X) = \widehat{v}(t, \widehat{S}^1_t) \quad \text{with} \quad \widehat{v} = \exp(-rt)v(t, x\exp(rt))$$

and that $\widehat{v}$ is twice continuously differentiable. Applying Itô formula yields

$$\begin{aligned}\widehat{V}_t(X) &= \widehat{v}(t, \widehat{S}^1_t)\\ &= \widehat{v}(0, \widehat{S}^1_0) + \int_0^t \frac{\partial}{\partial t}\widehat{v}(s, \widehat{S}^1_s)\,\mathrm{d}s + \int_0^t \frac{\partial}{\partial x}\widehat{v}(s, \widehat{S}^1_s)\,\mathrm{d}\widehat{S}^1_s + \frac{1}{2}\int_0^t \frac{\partial^2}{\partial x^2}\widehat{v}(s, \widehat{S}^1_s)\,\mathrm{d}\langle \widehat{S}^1\rangle_s\end{aligned}$$

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

$$\langle \widehat{S}^1\rangle_s = \langle(\sigma\widehat{S}^1 \cdot \widetilde{W})\rangle_s = \int_0^s (\widehat{S}^1_u\sigma)^2\,\mathrm{d}u,$$

leads to

$$\widehat{V}_t(X) = \widehat{v}_0(0, \widehat{S}^1_0) + \int_0^t \frac{\partial}{\partial x}\widehat{v}(s, \widehat{S}^1_s)\,\mathrm{d}\widehat{S}^1_s + \int_0^t \left(\frac{\partial}{\partial t}\widehat{v}(s, \widehat{S}^1_s) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\widehat{v}(s, \widehat{S}^1_s)(\widehat{S}^1_s\sigma)^2\right)\mathrm{d}s.$$

Since $\widehat{S}^1$ and $\widehat{V}(X)$ are Q-martingales and $\widehat{v}_0(0, \widehat{S}^1_0) = \widehat{V}_0(X)$, we observe that

$$\int_0^t \left(\frac{\partial}{\partial t}\widehat{v}(s, \widehat{S}^1_s) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\widehat{v}(s, \widehat{S}^1_s)(\widehat{S}^1_s\sigma)^2\right)\mathrm{d}s = \widehat{V}_t(X) - \widehat{V}_0(X) - \int_0^t \frac{\partial}{\partial x}\widehat{v}(s, \widehat{S}^1_s)\,\mathrm{d}\widehat{S}^1_s \qquad (3.2)$$

is a continuous local martingale of finite variation, which means it is identical to zero. This means

$$\int_0^t \left(\frac{\partial}{\partial t}\widehat{v}(s, \widehat{S}^1_s) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\widehat{v}(s, \widehat{S}^1_s)(\widehat{S}^1_s\sigma)^2\right)\mathrm{d}s = 0, \quad \text{for all } t \in [0, T],$$

and

$$\widehat{V}_t(X) = \widehat{V}_0(X) + \int_0^t \frac{\partial}{\partial x}\widehat{v}(s, \widehat{S}^1_s)\,\mathrm{d}\widehat{S}^1_s, \quad \text{for all } t \in [0, T].$$

Let $\varphi = (\varphi^0, \varphi^1)$ be the self-financing trading strategy with $\varphi^1 := \frac{\partial}{\partial x}\widehat{v}(s, \widehat{S}^1_s)$ and initial capital $\widehat{V}_0(X)$, that is, we have

$$\widehat{V}(X) = \widehat{V}_0(X) + (\varphi^1 \cdot \widehat{S}^1) = \widehat{V}(\varphi).$$

Since $0 \leq V(\varphi) = V(X) \leq c(S^0 + S^1)$, the trading strategy φ is strongly admissible and thus a replicating trading strategy. Furthermore, one has

$$\varphi^1_t = \frac{\partial}{\partial x}\widehat{v}(t, \widehat{S}^1_t) = \frac{\partial}{\partial x}\widehat{v}(t, \exp(-rt)S^1_t) = \frac{\partial}{\partial x}v(t, S^1_t), \quad t \in [0, T],$$

and

$$\varphi^0_t = \widehat{V}_t(\varphi) - \varphi^1_t\widehat{S}^1_t = \widehat{v}(t, \widehat{S}^1_t) - \widehat{S}^1_t\frac{\partial}{\partial x}v(t, S^1_t) = \exp(-rt)\left(v(t, S^1_t) - S^1_t\frac{\partial}{\partial x}v(t, S^1_t)\right).$$

(ii) Because of (3.2), it holds

$$\frac{\partial}{\partial t}\widehat{v}(t, \widehat{S}^1_t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\widehat{v}(t, \widehat{S}^1_t)(\widehat{S}^1_t\sigma)^2 = 0$$

for almost all $t \in [0, T]$ and thus for all $t \in [0, T]$ by continuity. Hence, for a fixed t is holds

$$\frac{\partial}{\partial t}\widehat{v}(t,x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\widehat{v}(t,x)(x\sigma)^2 = 0$$

for $\mathcal{L}(\widehat{S}_t^1)$-almost all $x > 0$ and thus for all $x > 0$ since the distribution of $\widehat{S}_t^1$ is equivalent to the Lebesgue measure on $(0, \infty)$ and by continuity. Therefore, we conclude that

$$\begin{aligned}
0 &= \frac{\partial}{\partial t}\widehat{v}(t,x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\widehat{v}(t,x)(x\sigma)^2 \\
&= \frac{\partial}{\partial t}\Big(\exp(-rt)v(t, \exp(rt)x)\Big) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\widehat{v}(t,x)(x\sigma)^2 \\
&= -r\exp(-rt)v(t, x\exp(rt)) \\
&\quad + \exp(-rt)\frac{\partial}{\partial t}v(t, x\exp(rt)) + xr\frac{\partial}{\partial x}v(t, x\exp(rt)) \\
&\quad + \frac{1}{2}x^2\sigma^2\exp(rt)\frac{\partial^2}{\partial x^2}v(t, x\exp(rt)).
\end{aligned}$$

Setting $\widetilde{x} := x\exp(rt)$ gives

$$0 = \exp(-rt)\Big(-rv(t,\widetilde{x}) + \frac{\partial}{\partial t}v(t,\widetilde{x}) + r\widetilde{x}\frac{\partial}{\partial x}v(t,\widetilde{x}) + \frac{1}{2}\widetilde{x}^2\sigma^2\frac{\partial^2}{\partial x^2}v(t,\widetilde{x})\Big),$$

which reveals the Black-Scholes PDE

$$0 = -rv(t,\widetilde{x}) + \frac{\partial}{\partial t}v(t,\widetilde{x}) + r\widetilde{x}\frac{\partial}{\partial x}v(t,\widetilde{x}) + \frac{1}{2}\widetilde{x}^2\sigma^2\frac{\partial^2}{\partial x^2}v(t,\widetilde{x}).$$

□

The most frequently traded examples of vanilla options are Call and Put options. For these options we obtain fairly simple explicit expressions for the value processes and replicating strategies in Black-Scholes setting.

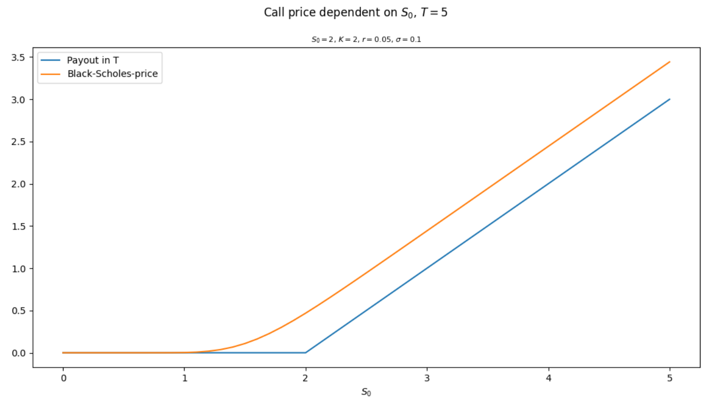
Proposition 3.8 (Black-Scholes formula). *Let $(S_T^1 - K)^+$ be a Call option with strike price $K \geq 0$ and maturity T and denote by $C(t, S_t^1) := V_t((S_T^1 - K)^+)$, $t \in [0, T]$, the corresponding value process. Then, one has*

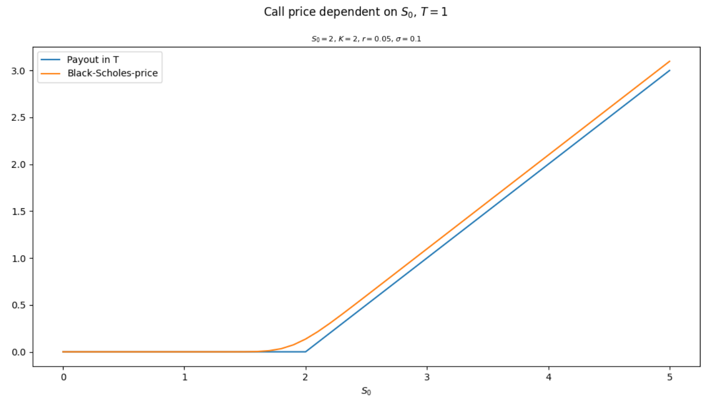
$$C(t, S_t^1) = S_t^1\Phi(d_1) - K\exp(-r(T-t))\Phi(d_2),$$

where Φ is the distribution function of the standard normal distribution $\mathcal{N}(0,1)$,

$$d_1 := \frac{1}{\sigma\sqrt{T-t}}\Big(\log\Big(\frac{S_t^1}{K}\Big) + \Big(r + \frac{\sigma^2}{2}\Big)(T-t)\Big) \quad \textit{and} \quad d_2 := d_1 - \sigma\sqrt{T-t}.$$

The following graphics show the Black-Scholes price of a Call option with strike $K = 2$ as a function in the risky asset price with different maturities.





Proof. [The proof is left as a self-study exercise.] Since $(S_T^1 - K)^+ = g(S_T^1)$ with $g(x) := (x-K)^+$, Proposition 3.6 states

$$C(t, S_t) := V_t((S_T^1 - K)^+) = \exp(-r(T-t)) \int_{\mathbb{R}} (e^y - K)^+ \varphi_{\log(S_t^1)+(r-\sigma^2/2)(T-t), \sigma^2(T-t)}(y)\, \mathrm{d}y,$$

for $t \in [0, T]$. Using the substitution

$$x := S_t^1 \quad \text{and} \quad y := \log(x) + (r - \frac{\sigma^2}{2})(T-t) + \sigma\sqrt{T-t}\, z,$$

leads to

$$C(t, x) = \exp(-r(T-t)) \int_{\mathbb{R}} \left(x \exp\left((r - \frac{\sigma^2}{2})(T-t) + \sigma\sqrt{T-t} z\right) - K \right)^+ \varphi_{0,1}(z)\, \mathrm{d}z.$$

Setting $\tau := T - t$, $\widetilde{r} := r - \frac{\sigma^2}{2}$ and $l := \frac{\log(K/x) - \widetilde{r}\tau}{\sigma\sqrt{\tau}}$, we calculate

$$\begin{aligned}
C(t,x) &= \exp(-r\tau)\int_l^\infty \Big(x\exp\big(\widetilde{r}\tau + \sigma\sqrt{\tau}z\big) - K\Big)\varphi_{0,1}(z)\,\mathrm{d}z \\
&= x\exp\big(-\frac{1}{2}\sigma^2\tau\big)\int_l^\infty \exp\big(\sigma\sqrt{\tau}z\big)\frac{1}{\sqrt{2\pi}}\exp(-\frac{z^2}{2})\,\mathrm{d}z \\
&\quad - \exp(-r\tau)K\int_l^\infty \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{z^2}{2}\Big)\,\mathrm{d}z \\
&= x\int_l^\infty \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{(z-\sigma\sqrt{\tau})^2}{2}\Big)\,\mathrm{d}z - \exp(-r\tau)K\Phi(-l) \\
&= x\int_{l-\sigma\sqrt{\tau}}^\infty \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{z^2}{2}\Big)\,\mathrm{d}z - \exp(-r\tau)K\Phi(-l) \\
&= x\Phi(-l+\sigma\sqrt{\tau}) - \exp(-r\tau)K\Phi(-l),
\end{aligned}$$

where we used in the third line that

$$(z-\sigma\sqrt{\tau})^2 = z^2 - 2\sigma\sqrt{\tau}z + \sigma^2\tau \quad \text{and} \quad \Phi([l,\infty)) = \Phi([-\infty,-l)).$$

Plugging in $S_t^1 = x$ reveals the Black-Scholes formula for a Call option. □

Remark 3.9. Using the Black-Scholes formula (Proposition 3.8) and Proposition 3.7, one can also obtain a formula for the replicating trading strategy $\varphi = (\varphi^0, \varphi^1)$ of the Call option $(S_T^1 - K)^+$, which is given by

$$\begin{aligned}
\varphi_t^0 &= \exp(-rT)K\Phi\Big(\frac{1}{\sigma\sqrt{T-t}}\Big(\log\big(\frac{S_t^1}{K}\big) + (r - \frac{\sigma^2}{2})(T-t)\Big)\Big), \\
\varphi_t^1 &= \Phi\Big(\frac{1}{\sigma\sqrt{T-t}}\Big(\log\big(\frac{S_t^1}{K}\big) + (r + \frac{\sigma^2}{2})(T-t)\Big)\Big),
\end{aligned}$$

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

Relying on the so-called Put-Call Parity, one can immediately derive the Black-Scholes formula and the replicating trading strategy for Put options as well. Notice that the Put option $(K - S_T^1)^+$ can be synthesized using a Call option, the risky asset S^1 and the risk-less asset S^0 (e.g. a bond), that is

$$\underbrace{(K - S_T^1)^+}_{\text{put}} = \underbrace{(S_T^1 - K)^+}_{\text{call}} - \underbrace{S_T^1}_{\text{stock}} + K\exp(-rT)\underbrace{S_T^0}_{\text{bond}}.$$

This leads to the following lemma (exercise).

Lemma 3.10 (Put-Call Parity). *Let $C(t, S_t^1)$ and $P(t, S_t^1)$ be the value at time t of a Call option and of a Put option, respectively, both with the same strike price $K > 0$ and same maturity T. Then, the **Put-Call Parity** holds:*

$$P(t, S_t^1) = C(t, S_t^1) - S_t^1 + K\exp(-r(T-t)), \quad t \in [0,T].$$

Proof. See Exercise 3.2 on Problem sheet 3. □

Greeks: Considering the value process $(V_t(X))_{t\in[0,T]}$ of a financial option $X = g(S_T^1)$, we have seen it is of the form

$$V_t(X) = v(t, S_t^1; r, \sigma), \quad t \in [0,T],$$

in the Black-Scholes model. In the analysis of the value process and its corresponding replicating trading strategies, the partial derivatives of v play a crucial role. In the financial industry these partial derivatives are labeled by Greek letters and called **Greeks**. In other words, the Greeks measure the sensitivity of the value of a financial option to changes in one parameter while holding the other parameters fixed. The common examples are:

- $\Delta := \frac{\partial}{\partial x} v(t, S_t^1; r, \sigma)$ measures the sensitivity w.r.t. price changes. Δ is also the "Delta hedging strategy".
- $\Gamma := \frac{\partial^2}{\partial x^2} v(t, S_t^1; r, \sigma)$ measures the sensitivity of the Delta hedging Δ w.r.t. price changes. Gamma measures the discretization error of Delta hedging.
- $\rho := \frac{\partial}{\partial r} v(t, S_t^1; r, \sigma)$
- $\Theta := \frac{\partial}{\partial t} v(t, S_t^1; r, \sigma)$
- $\mathcal{V} := \frac{\partial}{\partial \sigma} v(t, S_t^1; r, \sigma)$ ("Vega")

3.3 Pricing and hedging of exotic options

Lecture 7

So far we have studied the pricing and hedging of vanilla options, that are options of the form $g(S_T^1)$. These rather simple claims allow to derive explicit formulas for their prices and replicating strategies in the Black-Scholes setting. Vanilla options cover, of course, not all types of financial options traded on real financial markets. Notice that there are many more financial derivatives with more complex payoffs, which can depend on the whole past price trajectory of the underlying price process $(S_t^1)_{t\in[0,T]}$. More complex financial derivatives are usually called *exotic options*, see Example 2.20. Often, exotic options do not allow to derive explicit pricing and hedging formulas and thus numerical methods are required to solve the associated pricing and hedging problems.

Recall, that the Black-Scholes model satisfies NFLVR and is a complete market, that implies that every exotic option can be replicated by trading on the underlying financial market. Moreover, by the second FTAP there exists a unique EMM Q. Hence, as for vanilla options, there are two main approaches to price an option:

- *PDE approach*: derive the replicating trading strategy by finding the pricing PDE for the value function.
- *martingale approach*: calculate the value process using the EMM Q and then derive the replicating trading strategy.

Let us briefly discuss these approaches for two examples (Asian options and Barrier options).

Asian options are options depending on some type of average (arithmetic, geometric, ...) in a discrete or continuous manner, of the underlying price process $(S_t^1)_{t\in[0,T]}$. As a prototypical example let us consider the Asian option A with payoff

$$A := g(S_T^1, I_T) \quad \text{with } I_t := \int_0^t f(u, S_u^1)\,\mathrm{d}u, \quad t \in [0,T],$$

for continuous functions $g\colon \mathbb{R}^2 \to \mathbb{R}_+$ and $f\colon [0,T]\times\mathbb{R} \to \mathbb{R}_+$. Note that this Asian option is outside the scope of Subsection 3.1. However, we know that the claim A is replicable in the Black-Scholes model. Hence, let us try to solve (heuristically and without details) the pricing and hedging problem for the Asian option $g(S_T^1, I_T)$ by finding its value functions using the PDE approach.

Recall, the risky asset $(S_t^1)_{t\in[0,T]}$ in the Black-Scholes model follows the dynamics

$$\mathrm{d}S_t^1 = \mu S_t^1\,\mathrm{d}t + \sigma S_t^1\,\mathrm{d}W_t, \quad t \in [0,T],$$

and we observe that

$$\mathrm{d}I_t = f(t, S_t^1)\,\mathrm{d}t, \quad t \in [0,T].$$

Since $(S_t^1, I_t)_{t\in[0,T]}$ is a two-dimensional Markov process ("At time t the value (S_t^1, I_t) provides all accessible information about the future, i.e. $\mathbb{E}^Q[(S_T^1, I_T)|\mathcal{F}_t] = \mathbb{E}^Q[(S_T^1, I_T)|(S_t^1, I_t)]$.") as a solution of two dimensional SDE, we make the ansatz

$$V_t(A) = \exp(-r(T-t))\mathbb{E}^Q[A|\mathcal{F}_t] = v(t, S_t^1, I_t) \quad \text{for some } v\colon [0,T]\times\mathbb{R}^2 \to \mathbb{R}.$$

Applying Itô formula to $v(t, S_t, I_t)$, we get

$$\begin{aligned}
&\mathrm{d}v(t, S_t^1, I_t)\\
&\quad= \frac{\partial}{\partial t}v(t,S_t^1,I_t)\,\mathrm{d}t + \frac{\partial}{\partial x}v(t,S_t^1,I_t)\,\mathrm{d}S_t^1 + \frac{1}{2}\sigma^2(S_t^1)^2\frac{\partial^2}{\partial x}v(t,S_t^1,I_t)\,\mathrm{d}t + \frac{\partial}{\partial I}v(t,S_t^1,I_t)\,\mathrm{d}I_t\\
&\quad= \frac{\partial}{\partial x}v(t,S_t^1,I_t)\,\mathrm{d}S_t^1 \qquad\qquad (3.3)\\
&\qquad+ \Big(\frac{\partial}{\partial t}v(t,S_t^1,I_t) + \frac{1}{2}\sigma^2(S_t^1)^2\frac{\partial^2}{\partial x}v(t,S_t^1,I_t) + \frac{\partial}{\partial I}v(t,S_t^1,I_t)f(t,S_t^1)\Big)\,\mathrm{d}t.
\end{aligned}$$

To eliminate the local risk, we choose the delta hedging strategy $\varphi^1 = \frac{\partial}{\partial x}v(t, S_t^1, I_t)$. Indeed, then we have that the value process $(V_t(A) - (\varphi\cdot S^1)_t)_{t\in[0,T]}$ is locally risk-free, since (3.3) reveals

$$\begin{aligned}
&\mathrm{d}v(t,S_t^1,I_t) - \frac{\partial}{\partial x}v(t,S_t^1,I_t)\,\mathrm{d}S_t^1 \qquad\qquad (3.4)\\
&\quad= \Big(\frac{\partial}{\partial t}v(t,S_t^1,I_t) + \frac{1}{2}\sigma^2(S_t^1)^2\frac{\partial^2}{\partial x}v(t,S_t^1,I_t) + \frac{\partial}{\partial I}v(t,S_t^1,I_t)f(t,S_t^1)\Big)\,\mathrm{d}t.
\end{aligned}$$

By Lemma 3.5 we also have

$$\mathrm{d}v(t,S_t^1,I_t) - \frac{\partial}{\partial x}v(t,S_t^1,I_t)\,\mathrm{d}S_t^1 = r\big(v(t,S_t^1,I_t) - \frac{\partial}{\partial x}v(t,S_t^1,I_t)S_t^1\big)\,\mathrm{d}t, \quad t\in[0,T]. \qquad (3.5)$$

Hence, combining (3.4) and (3.5) leads to the following pricing PDE for the Asian option A:

$$\frac{\partial}{\partial t}v + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2}v + rx\frac{\partial}{\partial x}v + f(t,x)\frac{\partial}{\partial I}v - rv = 0, \quad t \in [0,T), \quad x, I \in \mathbb{R}_+,$$

where we suppressed the dependence of v on t, x, I.

However, this pricing PDE for the value process can, in general, not be solved explicitly and, thus, one often needs numerical methods to actually solve the pricing and hedging problem for Asian options.

Barrier options are options whose existence depends upon the underlying asset's price hitting (or not) a barrier, e.g. up-and-out Call option X with payoff

$$C^{\text{up\&out}} := (S_T^1 - K)^+ \mathbb{1}_{\{\max_{t\in[0,T]} S_t < B\}} = (S_T^1 - K)^+ \mathbb{1}_{\{M_T^1 < B\}} \quad \text{for some } B \geq K \geq 0,$$

where we set

$$M_t^1 := \max_{u\in[0,t]} S_u^1, \quad \text{for } t \in [0,T].$$

For this barrier option, we can use the martingale approach. To this end, we need to calculate the value process

$$V_t(C^{\text{up\&out}}) = \exp(-r(T-t))\mathbb{E}^Q[C^{\text{up\&out}}|\mathcal{F}_t], \quad t \in [0,T].$$

Notice that in the Black-Scholes model the source of randomness is the Brownian motion $(W_t)_{t\in[0,T]}$, which is normally distributed $W_t \sim \mathcal{N}(0,t)$, from which one can derive related distributions e.g. the distribution of its running maximum or running minimum. For the up-and-out Call options, we are lucky and can obtain indeed an explicit formula for its price and its corresponding replicating strategy.

Lemma 3.11. *For $K \leq B$ consider the up-and-out Call option $C^{\text{up\&out}}$ with payoff*

$$C^{\text{up\&out}} = (S_T^1 - K)^+ \mathbb{1}_{\{M_T^1 < B\}}.$$

The value process $(V_t(C^{\text{up\&out}}))_{t\in[0,T]}$ is given by

$$V_t(C^{\text{up\&out}}) = v(t, S_t^1, M_t^1), \quad t \in [0,T],$$

with

$$\begin{aligned}
v(t,x,m) = x\mathbb{1}_{\{m<B\}}&\Big\{\Phi\Big(\delta_+^{T-t}\Big(\frac{x}{K}\Big)\Big) - \Phi\Big(\delta_+^{T-t}\Big(\frac{x}{B}\Big)\Big) \\
&- \Big(\frac{B}{x}\Big)^{1+\frac{2r}{\sigma^2}}\Big(\Phi\Big(\delta_+^{T-t}\Big(\frac{B^2}{Kx}\Big)\Big) - \Phi\Big(\delta_+^{T-t}\Big(\frac{B}{x}\Big)\Big)\Big)\Big\} \\
&- \exp(-r(T-t))K\mathbb{1}_{\{m<B\}}\Big\{\Phi\Big(\delta_-^{T-t}\Big(\frac{x}{K}\Big)\Big) - \Phi\Big(\delta_-^{T-t}\Big(\frac{x}{B}\Big)\Big) \\
&- \Big(\frac{x}{B}\Big)^{1-\frac{2r}{\sigma^2}}\Big(\Phi\Big(\delta_-^{T-t}\Big(\frac{B^2}{Kx}\Big)\Big) - \Phi\Big(\delta_-^{T-t}\Big(\frac{B}{x}\Big)\Big)\Big)\Big\},
\end{aligned}$$

where Φ is the distribution function of the standard normal distribution $\mathcal{N}(0,1)$ and

$$\delta_\pm^\tau(s) := \frac{1}{\sigma\sqrt{\tau}}\Big(\log s + \Big(r \pm \frac{\sigma^2}{2}\Big)\tau\Big), \quad s, \tau > 0.$$

The replicating trading strategy $(\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ *is given by*

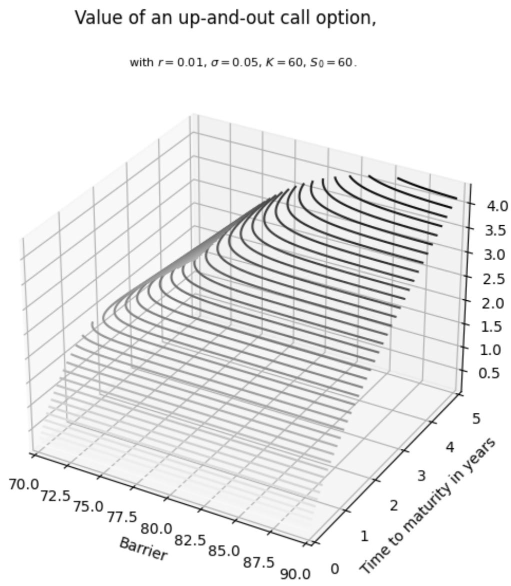
$$\varphi_t^1 := \begin{cases} \frac{\partial}{\partial x} v(t, S_t^1, M_t^1) & \text{if } \max_{s\in[0,t]} S_s^1 < B \\ 0 & \text{otherwise} \end{cases},$$

$$\varphi_t^0 := \begin{cases} \exp(-rt)\big(v(t, S_t^1, M_t^1) - S_t^1 \frac{\partial}{\partial x} v(t, S_t^1, M_t^1)\big) & \text{if } \max_{s\in[0,t]} S_s^1 < B \\ 0 & \text{otherwise} \end{cases},$$

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

The price of the up-and-out barrier Call option $C^{\text{up\&out}}$ *is zero if* $B \le K$.

See (Shreve, 2004, Section 7.3.3) for a derivation pricing formula presented in Lemma 3.11.



Remark. The price of up-and-out Call option converges to the price of a Call option if the barrier B tends to ∞.

4 Volatility modeling

In order to calculate the prices and hedging strategies for financial options in the Black-Scholes model, *the crucial parameter is the volatility* σ, which can not be directly observed on a financial market. Notice that the interest rate r is basically observable on a financial market and the parameters of a financial derivative (e.g. its maturity T or strike price K) are known when buying or selling an options. Furthermore, recall that the drift parameter μ does not appear neither in the pricing formulas nor in the hedging strategies of financial derivatives.

In practical applications the volatility σ needs to be estimated from collected asset price data or from data of prices of similar products traded on the underlying financial market. We briefly discuss two standard approaches.

Historical volatility: The first approach to determine the volatility σ on a real world financial market is to statistically estimate (or forecast) it from the in the past observed asset prices. Thus an estimation leads to the so-called **historical volatility**.

Let us assume we observe the prices $(S^1_{t_i})_{i=0,1,\ldots,n}$ at equidistant time points in the interval $[0,T]$, say $t_i = i\delta$ with delta $\delta := \frac{T}{n} > 0$ and the prices $(S^1_{t_i})_{i=0,1,\ldots,n}$ are supposed to come from a Back-Scholes model with unknown volatility $\sigma > 0$, i.e.

$$S^1_t = S^1_0 \exp(\widetilde{\mu} t + \sigma W_t), \quad t \in [0,T].$$

Hence, we see that

$$S^1_{t_{i+1}} = S^1_{t_i} \exp(\widetilde{\mu}\delta + \sigma(W_{t_{i+1}} - W_{t_i})), \quad i = 0, 1, \ldots, n-1,$$

which leads to (independent and identically distributed) logarithmic relative increments

$$R_i := \log\left(\frac{S^1_{t_{i+1}}}{S^1_{t_i}}\right) = \widetilde{\mu}\delta + \sigma(W_{t_{i+1}} - W_{t_i}).$$

Consequently, the variance

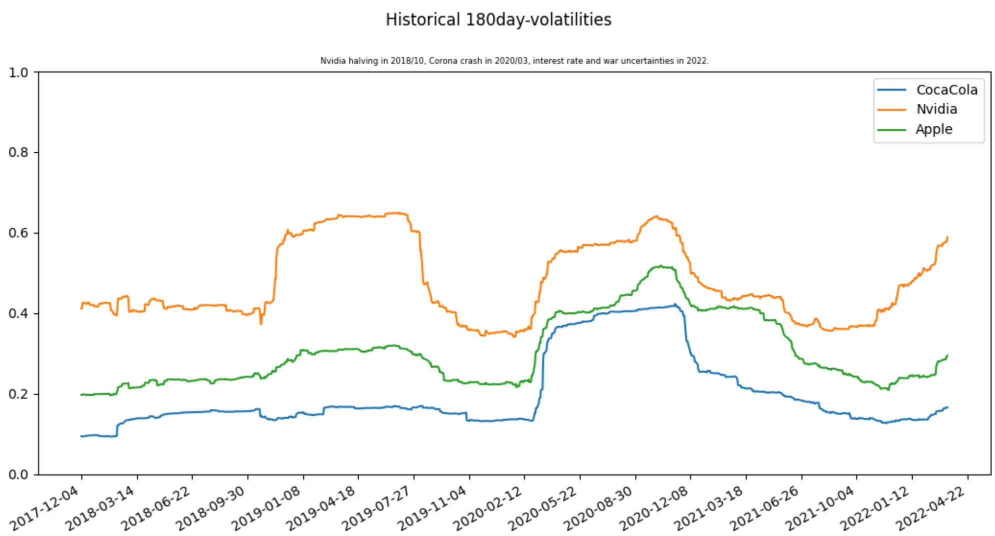
$$\mathrm{Var}(R_i) = \sigma^2\delta$$

and we know a good estimator for the $\mathrm{Var}(R_i)$ is the sample variance

$$\widehat{v} := \frac{1}{n-1}\sum_{i=0}^{n-1}(R_i - \overline{R})^2 \quad \text{with } \overline{R} := \frac{1}{n}\sum_{i=0}^{n-1} R_i,$$

which is an unbiased estimator. Hence, a natural estimator of the volatility σ is the Black-Scholes model is given by

$$\widehat{\sigma} := \sqrt{\frac{\widehat{v}}{\delta}}.$$



Implied volatility: The second approach to determine the volatility σ on a real world financial market is to derived it the prices of currently traded options, e.g. Call or Put options.

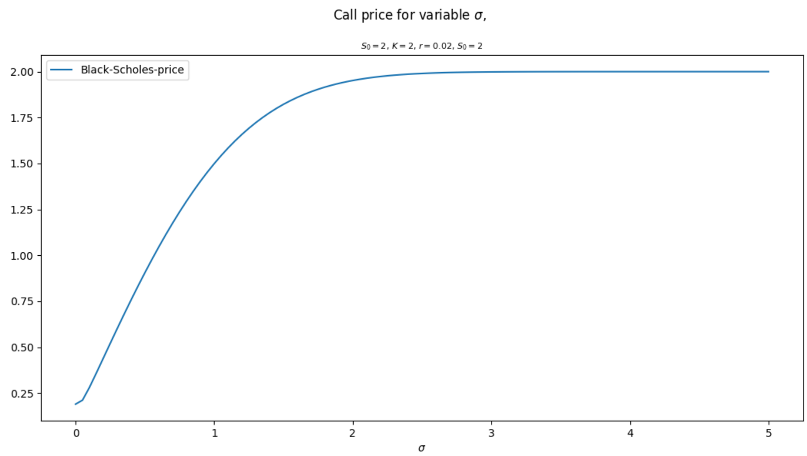
Recall that, assuming the Black-Scholes model does correctly model the underlying financial market, the current prices of Call options are given by the Black-Scholes formula (Proposition 3.8):

$$C(t, S_t^1; r, \sigma, K, T) = C(t, S_t^1) = S_t^1 \Phi(d_1) - K \exp(-r(T-t))\Phi(d_2), \quad t \in [0, T]$$

with

$$d_1 := \frac{1}{\sigma\sqrt{T-t}}\left(\log\left(\frac{S_t^1}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right) \quad \text{and} \quad d_2 := d_1 - \sigma\sqrt{T-t}.$$

In particular, we see that the Call price $C(t, S_t^1; r, \sigma, K, T)$ is a function of the model parameters r, σ, the option parameters K, T, the current time t and price S_t^1. As discussed before, we know all input parameters expect the volatility σ and we notice that $\sigma \mapsto C(t, S_t^1; r, \sigma, K, T)$ is strictly increasing.



Hence, observing the Call price C_{market} and setting

$$C(t, S_t^1; r, \sigma_{\text{implied}}, K, T) = C_{\text{market}},$$

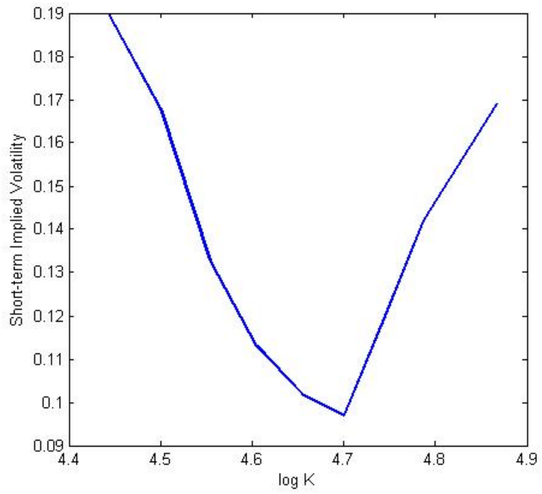
we can invert resulting Black-Scholes formula to obtain the so-called **implied volatility** σ_{implied}, which fits to the observed Call prices. Empirical studies have shown that approach based on implied volatility is better suited than the historical volatility for predicting the volatility of future asset's prices. The pricing of other options can be based on an implied volatility deduced from related options. Extracting an implied volatility is nowdays the primary use of the Black-Scholes formula by practitioners.

If the Black-Scholes models fits well the real world, then the implied volatility would be the same constant for all options with different strike prices or maturities. However, using the methods of implied volatility based on the Blach-Scholes model, empirical evidence leads to a so-called

- *"volatility smile"* (the implied volatility for Call prices with different strike prices tends to rise for deeply in- or out-of-money options), and

- *"volatility skew"* (the implied volatility for Call prices with different strike prices is lopsided),

as can be seen in the figure below. The following graphic shows the implied volatility of Calls on the S&P 500 index, for the smallest available maturity, plotted as a function of the log-strike.



Lecture 8

It is not surprising that the assumption of a constant volatility σ, as made in the classical Black-Scholes model, is an oversimplification of the real world. Empirical data from financial markets lead to various stylized facts about the actual behavior of the volatility, such as:

- *Volatility varies in time.*

- *Volatility clustering*: high volatility is followed by high volatility, low volatility by follow low one.

- *Volatility is mean-reverting*: while the volatility is fluctuating, it tends to return to its mean.

- *Volatility has heavy tails*: the log returns of many risky assets have heaver tails than the normal distribution.

Note that one can make many more empirical observations about the true nature of the volatility on financial markets. Of course, the exact behavior depends on the actually considered class of risky assets and the specific financial market, see for example Gatheral (2006) or Cont (2001).

Remark. To extend the Black-Scholes model covering a time-dependent deterministic volatility parameter turns out the be a fairly straightforward task. Let $\sigma\colon [0,T] \to \mathbb{R}_+$ and $\mu\colon [0,T] \to \mathbb{R}$ be deterministic functions. The one-dimensional Black-Scholes model with time-dependent volatility is given by

$$\mathrm{d}S_t^1 = \mu_t S_t^1 \,\mathrm{d}t + \sigma_t S_t^1 \,\mathrm{d}W_t, \quad S_0^1 = s_0^1, \quad t \in [0,T], \tag{4.1}$$

where $(W_t)_{t\in[0,T]}$ is a Brownian motion and $s_0^1 \in \mathbb{R}_+$. Note that the arguments presented in Section 3.1, 2.3 and 3.3 generalize to the Black-Scholes model (4.1) since the deterministic nature of the volatility $(\sigma_t)_{t\in[0,T]}$ does not effect the probabilistic arguments. However, many of the frequently observed properties of volatility are still not captured by the model (4.1). Therefore, over the past 40 years various probabilistic models for the volatility process were developed.

4.1 Local volatility models

To capture the volatility smile and/or the volatility skew, one can model the asset price by a stochastic differential equation

$$\mathrm{d}S_t^1 = \mu(t, S_t^1)S_t^1\,\mathrm{d}t + \sigma(t, S_t^1)S_t^1\,\mathrm{d}W_t, \quad S_0^1 = S_0^1, \quad t \in [0,T], \tag{4.2}$$

where $\mu\colon [0,T] \to \mathbb{R}$ and $\sigma\colon [0,T] \times \mathbb{R} \to \mathbb{R}_+$ are deterministic and sufficiently smooth functions such that there exists a unique solution to (4.2), cf. e.g. Theorem A.31. As before, $S_0^1 \in \mathbb{R}_+$ and $(W_t)_{t\in[0,T]}$ is a Brownian motion. The financial models of the form (4.2) is called **local volatility model** or **level-dependent volatility model**. For notational simplicity, we assume that the interest rate $r = 0$, i.e. $S_t^0 = 1$ for $t \in [0,T]$.

Assuming μ, σ are sufficiently nice, we obtain a unique equivalent martingale measure Q by the Radon-Nikodym density

$$\frac{\mathrm{d}Q}{d\mathbb{P}} := \mathcal{E}\Big(- \int_0^\cdot \frac{\mu}{\sigma}(u, S_u^1)\,\mathrm{d}W_u \Big)_T = \exp\Big(- \int_0^T \frac{\mu}{\sigma}(u, S_u^1)\,\mathrm{d}W_u - \frac{1}{2}\int_0^T \Big(\frac{\mu}{\sigma}(u, S_u^1)\Big)^2\,\mathrm{d}u \Big).$$

By Girsanov's theorem, we have

$$\mathrm{d}S_t^1 = \sigma(t, S_t^1)S_t^1\,\mathrm{d}\widetilde{W}_t, \quad S_0^1 = s_0^1, \quad t \in [0,T],$$

for a Q-Brownian motion $(\widetilde{W}_t)_{t\in[0,T]}$ with

$$\widetilde{W}_t := W_t + \int_0^t \frac{\mu}{\sigma}(u, S_u)\,\mathrm{d}u.$$

In particular, the local volatility model (S^0, S^1) satisfies NFLVR, it is a complete market and the pricing and hedging theory of Subsection 2.3 applies to it. However, notice that the previous conclusions only hold under suitable conditions on μ and σ!

Breeden and Litzenberger (1978) observed that the distribution of S_T^1 is fully determined if one knows the prices of Call options $(S_T^1 - K)^+$ for all strike prices K. Indeed, let us assume that we know all prices $C(S_0^1; K, T)$ of the Call options $(S_T^1 - K)^+$ for all $K \in \mathbb{R}_+$ and that the distribution of the asset price S_T^1 has a density function f (which is feasible in a local volatility model) under the equivalent martingale measure Q. Keeping in mind that $\widehat{S}_t^1 = S_t^1$ for $t \in [0,T]$, we know that

$$C(S_0^1; K, T) = \mathbb{E}^Q[(S_T^1 - K)^+] = \int_K^\infty (x - K)f(x)\,\mathrm{d}x$$

and differentiating twice with respect to K reveals the *Breeden-Litzenberger formula*

$$f(K) = \frac{\partial^2 C}{\partial K^2}(S_0^1; K, T), \quad K \geq 0,$$

see Exercise 5.3. So, if one really could observe the prices of Call options for *every* maturity $t \in [0,T]$ and *every* strike price $K \in \mathbb{R}_+$, the marginal distributions of the discounted price process $(\widehat{S}_t^1)_{t\in[0,T]}$ under the equivalent martingale measure Q would be fully determined. Given all the marginal distributions of $(\widehat{S}_t^1)_{t\in[0,T]}$, one can actually show that there exists a unique function $\sigma\colon [0,T]\times\mathbb{R}\to\mathbb{R}$ such that

$$\mathrm{d}\widehat{S}_t^1 = \sigma(t,\widehat{S}_t^1)\widehat{S}_t^1\,\mathrm{d}\widetilde{W}_t, \quad t\in[0,T].$$

In the case of local volatility models, Dupire (1994) derived a explicit formula for the volatility function σ, based on Call prices:

$$\sigma^2(t,x) = \frac{2}{K^2}\frac{\partial C}{\partial T}\left(\frac{\partial^2 C}{\partial^2 K}\right)^{-1}(t,x). \tag{4.3}$$

Note that the right-hand side of (4.3) can be computed from the observed Call prices. The formula (4.3) is called *Dupire's formula.* Of course, in reality Call options are not traded with every strike price and every maturity. So, the volatility function σ is usually estimated by using a parametric form for the volatility functions or by an interpolation algorithm. The estimated volatility function is called **implied volatility functions.**

Example 4.1 (CEV Model)**.** Taking the leverage effect into consideration (meaning there is a negative correlation between volatility and asset's returns), Cox (1975) and Cox and Ross (1976) proposed the *constant elasticity of variance model (CEV model)*

$$\mathrm{d}S_t^1 = \mu S_t^1\,\mathrm{d}t + \sigma(S_t^1)^{\frac{\alpha}{2}}\,\mathrm{d}W_t, \quad t\in[0,T],$$

where $\alpha\in(0,2]$ is a constant, called the elasticity factor. The CEV model only captures the skew of the implied volatility but not its smile.

While local volatility models allow for a fast calibration of the underlying model and lead to a feasible computational complexity in case of option pricing, empirical tests showed that the resulting implied volatility functions do not always fit the volatility surface well and are unstable over time.

4.2 Stochastic volatility models

The discussed stylized facts on the volatility of financial markets and to overcome the shortcoming of local volatility models motivated researchers to describe the volatility process by rather general stochastic processes. This led to introduce so-called **stochastic volatility models** for financial markets. In these models $(S_t^0, S_t^1)_{t\in[0,T]}$, it is often assumed that the risky asset follows a Black-Scholes type dynamics such as

$$\mathrm{d}S_t^1 = \mu_t S_t^1\,\mathrm{d}t + \sigma_t S_t^1\,\mathrm{d}W_t^1, \quad S_0^1 = S_0^1, \quad t\in[0,T], \tag{4.4}$$

where the volatility process itself is given by a stochastic process

$$\mathrm{d}\sigma_t = \alpha(t,\sigma_t)\,\mathrm{d}t + \beta(t,\sigma_t)\,\mathrm{d}W_t^2, \quad \sigma_0=\sigma_0, \quad t\in[0,T]. \tag{4.5}$$

Here W^1, W^2 are two Brownian motions with correlation $\rho\in(-1,1)$, $S_0^1, \sigma_0\in\mathbb{R}_+$. $\mu\colon[0,T]\to\mathbb{R}$ is a deterministic function (for simplicity) and $\alpha,\beta\colon[0,T]\to\mathbb{R}$ are suitable deterministic functions such that the SDEs (4.4) and (4.5) possesses unique solutions. Note that the correlation coefficient ρ is usually negative reflecting the negative correlation of volatility and prices.

As before in the Black-Scholes model, we assume that the risk-free asset $(S_t^0)_{t\in[0,T]}$ is given by

$$S_t^0 = rS_t^0 \,\mathrm{d}t, \quad S_0^0 = 1, \quad t \in [0,T],$$

for a constant interest rate $r \in \mathbb{R}$.

Remark 4.2. Note that the market model (4.4) consists of one risky asset but two sources of randomness (here: W^1 and W^2). This ensures that the market is arbitrage-free but leads to an incomplete market, i.e. not every bounded claim is replicable. The rule of thumb is that a $(d+1)$-dimensional market model with m sources of randomness is arbitrage-free if $d \le m$ and complete if $d \ge m$.

Since the stochastic volatility models is incomplete market, pricing and hedging of financial derivatives by arbitrage arguments is, in general, not possible anymore. In case of incomplete market, the option pricing will depend on the risk preferences of an investor. Nevertheless, let us study the pricing of a Vanilla option with payoff

$$X := g(S_T^1) \quad \text{for} \quad g \in C(\mathbb{R}_+;\mathbb{R}) \quad \text{with } |g(x)| \le C(1+|x|),$$

by constructing a hedging trading strategy φ, i.e. let us try to implement the "PDE approach". This allows us determine the effects of the stochastic volatility on the prices of financial derivatives.

Remark. Recall, in the classical Black-Scholes model, there is only one source of randomness and, therefore, the claim X can be hedged by trading in the one risky asset, that is, we could find a locally risk-free trading strategy (φ^0, φ^1) which replicates the payoff of the option at maturity:

$$\mathrm{d}V_t(X) - \varphi_t^0 \,\mathrm{d}S_t^0 - \varphi_t^1 \,\mathrm{d}S_t^1 = r\big(V_t(X) - \varphi_t^0 S_t^0 - \varphi_t^1 S_t^1\big)\,\mathrm{d}t \quad \text{and} \quad V_T(\varphi) = g(S_T^1).$$

In the present stochastic volatility model (S^0, S^1), the volatility is random itself and it needs to be hedged in order to form a locally risk-free trading strategy. Hence, let us make the theoretical assumptions:

(i) there exists a sufficiently smooth value function $v\colon [0,T]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ such that the value process $(V_t(X))_{t\in[0,T]}$ is given by

$$V_t(X) = v(t, S_t^1, \sigma_t), \quad t \in [0,T],$$

(ii) there is a value process $(V_t^\sigma)_{t\in[0,T]}$ associated to the volatility process $(\sigma_t)_{t\in[0,T]}$, in which one can trade and which is of the form

$$V_t^\sigma = v^\sigma(t, S_t^1, \sigma_t), \quad t \in [0,T],$$

for a sufficiently smooth value function $v^\sigma\colon [0,T]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$.

The aim is now to construct a locally risk-free process $(\Pi_t)_{t\in[0,T]}$ of the form

$$\Pi_t = V_t(X) - \varphi_t^0 S_t^0 - \varphi_t^1 S_t^1 - \varphi_t^2 V_t^\sigma, \quad t \in [0,T], \tag{4.6}$$

for a suitable self-financing trading strategy $\varphi = (\varphi^0, \varphi^1, \varphi^2)$. In other words, φ would be a potential hedging strategy for the option X.

Let us start by characterizing φ^1 and φ^2. Since $\varphi = (\varphi^0, \varphi^1, \varphi^2)$ is self-financing and using the above Assumption (i) and (ii), we have

$$\begin{aligned}
\mathrm{d}\Pi_t &= \mathrm{d}V_t(X) - \varphi_t^0 \mathrm{d}S_t^0 - \varphi_t^1 \mathrm{d}S_t^1 - \varphi_t^2 \mathrm{d}V_t^\sigma \\
&= \mathrm{d}v(t, S_t^1, \sigma_t) - \varphi_t^0 \mathrm{d}S_t^0 - \varphi_t^1 \mathrm{d}S_t^1 - \varphi_t^2 \mathrm{d}v^\sigma(t, S_t^1, \sigma_t).
\end{aligned}$$

Apply Itô formula to $v(t, S_t^1, \sigma_t)$ and $v^\sigma(t, S_t^1, \sigma_t)$ and suppressing their dependence on t, S_t^1, σ_t in the following, we see that

$$\begin{aligned}
\mathrm{d}\Pi_t = &\frac{\partial v}{\partial t}\mathrm{d}t + \frac{\partial v}{\partial x}\mathrm{d}S_t^1 + \frac{\partial v}{\partial \sigma}\mathrm{d}\sigma_t + \frac{\partial^2 v}{\partial x \partial \sigma}\,\mathrm{d}\langle S^1, \sigma\rangle_t + \frac{1}{2}\frac{\partial^2 v}{\partial x^2}\mathrm{d}\langle S^1\rangle_t + \frac{1}{2}\frac{\partial^2 v}{\partial \sigma^2}\mathrm{d}\langle \sigma\rangle_t \\
&- \varphi_t^0\,\mathrm{d}S_t^0 \\
&- \varphi_t^1 \mathrm{d}S_t^1 \\
&- \varphi_t^2\Big(\frac{\partial v^\sigma}{\partial t}\mathrm{d}t + \frac{\partial v^\sigma}{\partial x}\mathrm{d}S_t^1 + \frac{\partial v^\sigma}{\partial \sigma}\mathrm{d}\sigma_t + \frac{\partial^2 v^\sigma}{\partial x \partial \sigma}\,\mathrm{d}\langle S^1, \sigma\rangle_t + \frac{1}{2}\frac{\partial^2 v^\sigma}{\partial x^2}\mathrm{d}\langle S^1\rangle_t + \frac{1}{2}\frac{\partial^2 v^\sigma}{\partial \sigma^2}\mathrm{d}\langle \sigma\rangle_t\Big).
\end{aligned}$$

Recalling the dynamics (4.4) and (4.5) and that W^1, W^2 have the correlation ρ (i.e. $\langle W^1, W^2\rangle_t = \rho t$), we note that

$$\mathrm{d}\langle S^1\rangle_t = (\sigma_t S_t^1)^2\,\mathrm{d}t, \quad \mathrm{d}\langle \sigma\rangle_t = \beta^2(t,\sigma_t)\,\mathrm{d}t \quad \text{and} \quad \mathrm{d}\langle S^1, \sigma\rangle_t = \sigma_t S_t^1 \beta(t,\sigma_t)\rho\,\mathrm{d}t$$

and thus we get

$$\begin{aligned}
\mathrm{d}\Pi_t = &\Big(\frac{\partial v}{\partial t} + \sigma_t S_t^1 \beta\rho\frac{\partial^2 v}{\partial x \partial \sigma} + \frac{1}{2}(\sigma_t S_t^1)^2\frac{\partial^2 v}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2 v}{\partial \sigma^2}\Big)\,\mathrm{d}t \\
&- \varphi_t^2\Big(\frac{\partial v^\sigma}{\partial t} + \sigma_t S_t^1 \beta\rho\frac{\partial^2 v^\sigma}{\partial x \partial \sigma} + \frac{1}{2}(\sigma_t S_t^1)^2\frac{\partial^2 v^\sigma}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2 v^\sigma}{\partial \sigma^2}\Big)\,\mathrm{d}t \\
&- \varphi_t^0 r S_t^0\,\mathrm{d}t \\
&+ \Big(\frac{\partial v}{\partial x} - \varphi_t^2\frac{\partial v^\sigma}{\partial x} - \varphi_t^1\Big)\,\mathrm{d}S_t^1 \\
&+ \Big(\frac{\partial v}{\partial \sigma} - \varphi_t^2\frac{\partial v^\sigma}{\partial \sigma}\Big)\,\mathrm{d}\sigma_t.
\end{aligned}$$

To make the capital process Π locally risk-free, we must choose φ^1 and φ^2 such that

$$\frac{\partial v}{\partial x} - \varphi_t^2\frac{\partial v^\sigma}{\partial x} - \varphi_t^1 = 0 \tag{4.7}$$

to eliminate the risky $\mathrm{d}S_t^1$ terms, and

$$\frac{\partial v}{\partial \sigma} - \varphi_t^2\frac{\partial v^\sigma}{\partial \sigma} = 0 \tag{4.8}$$

to eliminate the risky $\mathrm{d}\sigma_t$ terms. Recalling that in an arbitrage-free market (with risk-free asset given by $\mathrm{d}S_t^0 = rS_t^0\,\mathrm{d}t$) every locally risk-free capital process fulfills

$$\mathrm{d}\Pi_t = r\Pi_t\,\mathrm{d}t,$$

equation (4.7) and (4.8) leads to

$$\begin{aligned}
\mathrm{d}\Pi_t &= \Big(\frac{\partial v}{\partial t} + \sigma_t S_t^1\beta\rho\frac{\partial^2 v}{\partial x\partial\sigma} + \frac{1}{2}(\sigma_t S_t^1)^2\frac{\partial^2 v}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2 v}{\partial\sigma^2}\Big)\,\mathrm{d}t\\
&\quad - \varphi_t^2\Big(\frac{\partial v^\sigma}{\partial t} + \sigma_t S_t^1\beta\rho\frac{\partial^2 v^\sigma}{\partial x\partial\sigma} + \frac{1}{2}(\sigma_t S_t^1)^2\frac{\partial^2 v^\sigma}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2 v^\sigma}{\partial\sigma^2}\Big)\mathrm{d}t\\
&\quad - \varphi_t^0 rS_t^0\,\mathrm{d}t\\
&= r\Pi_t\,\mathrm{d}t\\
&= r(v - \varphi_t^1 S_t^1 - \varphi_t^2 v^\sigma - \varphi_t^0 S_t^0)\,\mathrm{d}t,
\end{aligned}$$

where we used the definition (4.6) of Π and the above assumption (i) and (ii) in the last line. Since

$$\varphi_t^2 = \frac{\partial v}{\partial\sigma}\Big(\frac{\partial v^\sigma}{\partial\sigma}\Big)^{-1} \quad \text{and} \quad \varphi_t^1 = \frac{\partial v}{\partial x} - \frac{\partial v}{\partial\sigma}\Big(\frac{\partial v^\sigma}{\partial\sigma}\Big)^{-1}\frac{\partial v^\sigma}{\partial x},$$

we obtain

$$\begin{aligned}
&\frac{\partial v}{\partial t} + \sigma_t S_t^1\beta\rho\frac{\partial^2 v}{\partial x\partial\sigma} + \frac{1}{2}(\sigma_t S_t^1)^2\frac{\partial^2 v}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2 v}{\partial\sigma^2}\\
&\qquad - \frac{\partial v}{\partial\sigma}\Big(\frac{\partial v^\sigma}{\partial\sigma}\Big)^{-1}\Big(\frac{\partial v^\sigma}{\partial t} + \sigma_t S_t^1\beta\rho\frac{\partial^2 v^\sigma}{\partial x\partial\sigma} + \frac{1}{2}(\sigma_t S_t^1)^2\frac{\partial^2 v^\sigma}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2 v^\sigma}{\partial\sigma^2}\Big)\\
&\quad = r\Big(v - \frac{\partial v}{\partial x}S_t^1 + \frac{\partial v}{\partial\sigma}\Big(\frac{\partial v^\sigma}{\partial\sigma}\Big)^{-1}\frac{\partial v^\sigma}{\partial x}S_t^1 - \frac{\partial v}{\partial\sigma}\Big(\frac{\partial v^\sigma}{\partial\sigma}\Big)^{-1}v^\sigma\Big).
\end{aligned}$$

Collecting all terms with v on the left-hand side and all terms with v^σon the right-hand side, we get

$$\begin{aligned}
&\Big(\frac{\partial v}{\partial\sigma}\Big)^{-1}\Big(\frac{\partial v}{\partial t} + \sigma_t S_t^1\beta\rho\frac{\partial^2 v}{\partial x\partial\sigma} + \frac{1}{2}(\sigma_t S_t^1)^2\frac{\partial^2 v}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2 v}{\partial\sigma^2} + r\frac{\partial v}{\partial x}S_t^1 - rv\Big)\\
&\quad = \Big(\frac{\partial v^\sigma}{\partial\sigma}\Big)^{-1}\Big(\frac{\partial v^\sigma}{\partial t} + \sigma_t S_t^1\beta\rho\frac{\partial^2 v^\sigma}{\partial x\partial\sigma} + \frac{1}{2}(\sigma_t S_t^1)^2\frac{\partial^2 v^\sigma}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2 v^\sigma}{\partial\sigma^2} + rS_t^1\frac{\partial v^\sigma}{\partial x} - rv^\sigma\Big).
\end{aligned}$$

Notice that the left-hand side is a function depending only on v and the right-hand side only on v^σ. The only possibility that this can hold true is that both sides are identical to a function f depending only on S_t^1, σ_t and t. For conventional reason we choose

$$f(t, S_t^1, \sigma_t) := -(\alpha(t,\sigma_t) - \beta(t,\sigma_t)\theta(t, S_t^1, \sigma_t))$$

for some function θ. Hence, we deduce that

$$\frac{\partial v}{\partial t} + \sigma_t S_t^1\beta\rho\frac{\partial^2 v}{\partial x\partial\sigma} + \frac{1}{2}(\sigma_t S_t^1)^2\frac{\partial^2 v}{\partial x^2} + \frac{1}{2}\beta^2\frac{\partial^2 v}{\partial\sigma^2} + rS^1\frac{\partial v}{\partial x} - rv = -(\alpha - \beta\theta)\frac{\partial v}{\partial\sigma}. \tag{4.9}$$

The function θ is called the **market price of volatility risk** because it tells us how much of the expected return of V is explained by the risk of σ_t in the Capital Asset Pricing Model. Note that the PDE 4.9 reduces to the Black-Scholes PDE when the volatility σ is constant.

Remark. Alternatively to the above PDE, one can implement an martingale approach to get

$$V_t(X) = \exp(-r(T-t))\mathbb{E}^Q[f(S_T^1)|S_t = x, \sigma_t = y] = v(t, S_t^1, \sigma_t),$$

for an equivalent martingale measure Q, for a deterministic function $v\colon [0,T]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$.

While stochastic volatility models allow to describe rather well real financial markets, two of the main drawbacks are:

- It is extremely difficult to fit the parameters of stochastic volatility models.
- Stochastic volatility leads to a rather high computational complexity, in particular, when pricing exotic options.

Lecture 9

There exists a variety of volatility models, which often belong to the class of stochastic or local volatility models or a mixture of both. They all have their advantages and disadvantages and which one is best to use depends on the purpose in mind (which type of options, which financial market, ...). As generally true when modeling real world phenomena, one faces the trade-off between complex mathematical modeling and operability of the model. We conclude this chapter with two prominent examples of volatility models.

Example 4.3 (Heston Model)**.** In the *Heston model*, proposed by Heston (1993), the risky asset $(S_t^1)_{t\in[0,T]}$ is described by

$$\mathrm{d}S_t^1 = \mu_t S_t^1\,\mathrm{d}t + \sqrt{\sigma_t} S_t^1\,\mathrm{d}W_t^1, \quad t\in[0,T], \quad S_0^1 = s_0^1,$$

and

$$\mathrm{d}\sigma_t = -\lambda(\sigma_t - \overline{\sigma})\,\mathrm{d}t + \eta\sqrt{\sigma_t}\,\mathrm{d}W_t^2, \quad t\in[0,T], \quad \sigma_0 = \sigma_0,$$

where $(W_t^1)_{t\in[0,T]}$ and $(W_t^2)_{t\in[0,T]}$ are two Brownian motions with correlation $\rho\in[-1,1]$, $s_0^1, \sigma_0, \eta, \lambda, \overline{\sigma}\in\mathbb{R}_+$ are constants, and $\mu\colon[0,T]\to\mathbb{R}$ is a deterministic function. Notice that the Heston model is a stochastic volatility model and thus the results derived in this subsection apply to it. In particular, the resulting "pricing" PDE for vanilla options $X = g(S_T^1)$, which directly follows from (4.9) reads as

$$\frac{\partial v}{\partial t} + \sigma_t S_t^1 \eta\rho\frac{\partial^2 v}{\partial x\partial\sigma} + \frac{1}{2}\sigma_t(S_t^1)^2\frac{\partial^2 v}{\partial x^2} + \frac{1}{2}\sigma_t\eta^2\frac{\partial^2 v}{\partial\sigma^2} + rS^1\frac{\partial v}{\partial x} - rv = (\lambda(\sigma_t-\overline{\sigma}) - \theta)\frac{\partial v}{\partial\sigma},$$

Remarkably, this PDE has a closed form solution. This is one reason for the popularity of the Heston model.

Another example of a frequently used volatility model is the so-called SABR model.

Example 4.4 (SABR Model)**.** Hagan, Kumar, Lesniewski and Woodward (2002) introduced the *stochastic alpha-beta-rho (SABR model)*, which models the risky asset $(S_t^1)_{t\in[0,T]}$ by

$$\mathrm{d}S_t^1 = \alpha_t(S_t^1)^\beta\,\mathrm{d}W_t^1, \quad S_0^1 = S_0^1, \quad t\in[0,T],$$

with

$$\mathrm{d}\alpha_t = \nu\alpha_t\,\mathrm{d}W_t^2, \quad \sigma_0 = \sigma_0, \tag{4.10}$$

where $\nu > 0$ and $\beta\in[0,1]$ are constants and W^1, W^2 are two Brownian motions with correlation $\rho\in(-1,1)$. Note the stochastic volatility in the SABR model is $\sigma_t = \alpha_t(S_t^1)^{\beta-1}$ and for $\nu = 0$ the SABR model reduces to the CEV model. The SABR model is very popular in the financial industry, especially in the foreign exchange market and the interest rate markets. This is mainly due to the fact that uses only four parameters α_0, ν, ρ and β to better fit the various types of implied volatility curves observed on real financial markets.

5 Portfolio optimization

The aim of portfolio optimization is to construct investment strategies which are "optimal" in some sense. As a first step we need to discuss and agree on what we mean by "optimal". The classical approach for describing the behavior and preferences of investors is based on utility functions, as already introduced in the course "Mathematical Finance".

Definition 5.1. Let $S \subseteq \mathbb{R}$. A function $U\colon S \to \mathbb{R}$ is called **utility function** if U is strictly increasing, strictly concave and continuous on S.

Remark. For two utility functions U_1 and U_2 and constants $\alpha_1, \alpha_2 > 0$, $U := \alpha_1 U_1 + \alpha_2 U_2$ is again a utility function. In particular, every positive linear transformation of utility functions does not change the underlying preference relation. Hence, the exact values of a utility function are not really important.

Example 5.2. The most commonly used utility functions in mathematical finance are

- $U(x) = -\exp(-\gamma x)$, $x \in S := \mathbb{R}$, with $\gamma > 0$ ("exponential-utility"),
- $U(x) = \frac{x^p}{p}$ with $p < 1$, $p \neq 0$ and $S := (0, \infty)$ ("power-utility"),
- $U(x) = \log(x)$ with $S := (0, \infty)$ ("log-utility").

In the previous section we studied the problem (for an investor) to price and hedge a claim X. In other words, we tried to find the required initial capital/wealth to hedge a short position in the claim X. In this section we change the perspective: Given an initial capital/wealth $x \in \mathbb{R}$, the investor wants to invest in the financial market $(S_t)_{t\in[0,T]}$ with the aim to maximize her expected utility.

Let us consider two examples of portfolio optimization problems. In both examples we assume that an investor may invest in a financial market modeled by the one-dimensional Black-Scholes model $(S_t^0, S_t^1)_{t\in[0,T]}$ with

- a risk-free asset $(S_t^0)_{t\in[0,T]}$ with the dynamics

  $$\mathrm{d}S_t^0 = rS_t^0\,\mathrm{d}t, \quad t \in [0,T], \quad S_0^0 = S_0^0,$$

  where $S_0^0 > 0$, $r \in \mathbb{R}$ is the interest rate,
- a risky asset $(S_t^1)_{t\in[0,T]}$ with the dynamics

  $$\mathrm{d}S_t^1 = \mu S_t^1\,\mathrm{d}t + \sigma S_t^1\,\mathrm{d}W_t, \quad t \in [0,T], \quad S_0^1 = S_0^1,$$

  where $S_0^1 > 0$, $\mu \in \mathbb{R}$ is the drift parameter, $\sigma > 0$ is the volatility parameter, and $(W_t)_{t\in[0,T]}$ is a Brownian motion.

Portfolio allocation

An investor can invest, in self-financing manner, in the one-dimensional Black-Scholes model $(S_t^0, S_t^1)_{t\in[0,T]}$, that means, she has an initial capital $x > 0$ and she can pick a self-financing trading strategy $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$, where φ_t^i stands for the number of the i-th share hold at time t for $i = 0, 1$. Denoting by $V_t(\varphi)$ the wealth at time t and by π_t the proportion of the wealth $V_t(\varphi)$ invested into the risky asset S_t^1 at time t, we see that

$$\varphi_t^1 = \frac{\pi_t V_t(\varphi)}{S_t^1} \quad \text{and} \quad \varphi_t^0 = \frac{(1-\pi_t)V_t(\varphi)}{S_t^0}.$$

Hence, the self-financing wealth process $V_t(\varphi) =: V_t^\pi$ evolves according to following SDE

$$\begin{aligned} \mathrm{d}V_t^\pi &= \frac{(1-\pi_t)V_t^\pi}{S_t^0}\,\mathrm{d}S_t^0 + \frac{\pi_t V_t^\pi}{S_t^1}\mathrm{d}S_t^1 \\ &= ((1-\pi_t)r + \mu\pi_t)V_t^\pi\,\mathrm{d}t + \pi_t V_t^\pi\,\sigma\mathrm{d}W_t \end{aligned}$$

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

Let $\mathcal{A}$ be the set of processes $\pi \in \mathcal{L}(S)$ such that $\pi_t \in A$ for $t \in [0, T]$, for a given set $A \subseteq \mathbb{R}^d$. Given a utility function U modeling the preferences of the investor, the portfolio allocation problem aims to find the optimal investment $\pi^* \in \mathcal{A}$, that is

$$\mathbb{E}[U(V_T^{\pi^*})] = \sup_{\pi\in\mathcal{A}} \mathbb{E}[U(V_T^\pi)].$$

Note that this is the classical setting and portfolio optimization as treated first by Robert C. Merton (1976) assuming that U is a power-utility function

Optimizing wealth and consumption

An investor can again invest in the one-dimensional Black-Scholes model $(S_t^0, S_t^1)_{t\in[0,T]}$. Now the investor wants not only to maximize her profits but also the spending of her wealth on consumption. Let $(\pi_t)_{t\in[0,T]}$ be the proportion of her wealth invested in the risky asset S_t^1 at time t, let c_t be the proportion of her wealth spend for consumption at time t and the resulting wealth process is denoted by $(V_t^{(\pi,c)})_{t\in[0,T]}$ with given initial capital $V_0^{(\pi,c)} = x > 0$. Note that the proportion of her wealth invested in the risk-free asset S_t^0 at time t is $1 - \pi_t - c_t$. Supposing that trading is done in a self-financing manner, the wealth process $(V_t^{(\pi,c)})_{t\in[0,T]}$ should evolve according to following SDE

$$\begin{aligned} \mathrm{d}V_t^{(\pi,c)} &= \frac{(1-\pi_t-c_t)V_t^{(\pi,c)}}{S_t^0}\,\mathrm{d}S_t^0 + \frac{\pi_t V_t^{(\pi,c)}}{S_t^1}\,\mathrm{d}S_t^1 - c_t V_t^{(\pi,c)}\,\mathrm{d}t \\ &= ((1-\pi_t-c_t)r + \mu\pi_t - c_t)V_t^{(\pi,c)}\,\mathrm{d}t + \pi_t V_t^{(\pi,c)}\sigma\,\mathrm{d}W_t \end{aligned}$$

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

Given two utility functions U_1, U_2, the investor wants to find the optimal investment $(\pi_t^*)_{t\in[0,T]}$ and consumption $(c_t^*)_{t\in[0,T]}$, that is

$$\mathbb{E}\Big[\int_0^T U_1(c_t^* V_t^{(\pi^*,c^*)})\,\mathrm{d}t + U_2(V_T^{(\pi^*,c^*)})\Big] = \sup_{(\pi,c)} \mathbb{E}\Big[\int_0^T U_1(c_t V_t^{(\pi,c)})\,\mathrm{d}t + U_2(V_T^{(\pi,c)})\Big].$$

In words, the investor aims to maximize the expected utility of her terminal wealth as well as the expected utility of her consumption over her life time.

These two aforementioned maximization problems can be solved by using *stochastic control theory*, as we will see at the end of this chapter. Although not explicitly stated, it is essential that the underlying financial market satisfies, at least, no bounded profit with bounded risk since otherwise these maximization problems could degenerate in the sense that there are investment strategies leading to an infinite expected utility.

In the next subsection, we shall develop stochastic control theory in a fairly general context as it is an important theory allowing to deal with various types of control problems.

5.1 Stochastic control theory

Let us consider a general stetting for stochastic control problems which covers, in particular, the examples of the portfolio allocation problem and the problem of optimizing wealth and consumption, as discussed in the previous subsection.

Let $A \subseteq \mathbb{R}^m$ be a fixed set and $t \in [0,T]$. We consider a general control model where the state of the system is modeled by the stochastic differential equation

$$\mathrm{d}X_s^{t,x} = b(X_s^{t,x}, \alpha_s)\,\mathrm{d}s + \sigma(X_s^{t,x}, \alpha_s)\,\mathrm{d}W_s, \quad s \in [t,T], \quad X_t^{t,x} = x, \tag{5.1}$$

where

- $x \in \mathbb{R}^n$,
- $(W_t)_{t\in[0,T]}$ is a d-dimensional Brownian motion,
- $b\colon \mathbb{R}^n \times A \to \mathbb{R}^n$ is a measurable function such that there exists a constant $K > 0$:

$$|b(x,a) - b(y,a)| \le K|x-y|, \quad x,y \in \mathbb{R}^n, \quad a \in \mathbb{R}^m,$$

- $\sigma\colon \mathbb{R}^n \times A \to \mathbb{R}^{d\times n}$ is a measurable function such that there exists a constant $K > 0$:

$$|\sigma(x,a) - \sigma(y,a)| \le K|x-y|, \quad x,y \in \mathbb{R}^n, \quad a \in \mathbb{R}^m.$$

Note, due to the Lipschitz condition on the coefficients b and σ, for any initial condition $(t,x) \in [0,T]\times\mathbb{R}^n$, the stochastic differential equation (5.1) has a unique solution $(X_s^{t,x})_{s\in[t,T]}$.

Definition 5.3.

- For $t \in [0,T]$ the set of all stopping times τ with values in $[t,T]$ is denoted by $\mathcal{T}_{t,T}$.
- An adapted process $\alpha = (\alpha_t)_{t\in[0,T]}$ with values in A is called **control**.

Next let us introduce the objective functions

$$f\colon [0,T] \times \mathbb{R}^n \times A \to \mathbb{R} \quad \text{and} \quad g\colon \mathbb{R}^n \to \mathbb{R},$$

which are assumed to be measurable functions. Furthermore, we make the standing assumption:

Assumption 5.4. *For $g\colon \mathbb{R}^n \to \mathbb{R}$ we assume one of the following conditions:*

(i) g is bounded from below.

(ii) g satisfies

$$|g(x)| \leq C(1+|x|^2), \quad x \in \mathbb{R}^n,$$

for some constant $C > 0$.

For $(t,x) \in [0,T] \times \mathbb{R}^n$, we introduce the set $\mathcal{A}(t,x)$ of all controls $\alpha \in \mathcal{A}$ such that

$$\mathbb{E}\Big[\int_t^T |f(s, X_s^{t,x}, \alpha_s)| \,\mathrm{d}s\Big] < \infty,$$

and we assume that $\mathcal{A}(t,x)$ is not empty for all $(t,x) \in [0,T] \times \mathbb{R}^n$.

We define the gain function

$$J(t,x,\alpha) := \mathbb{E}\Big[\int_t^T f(s, X_s^{t,x}, \alpha_s) \,\mathrm{d}s + g(X_T^{t,x})\Big]$$

for all $(t,x) \in [0,T] \times \mathbb{R}^n$ and $\alpha \in \mathcal{A}(t,x)$.

The aim is to maximize the gain function J over all control processes. To this end, we introduce the associated value function

$$v(t,x) := \sup_{\alpha \in \mathcal{A}(t,x)} J(t,x,\alpha). \tag{5.2}$$

In the following we always assume that the value function is a measurable function. However, note that this is a non-trivial statement, which requires usually a measurable selection theorem.

Definition 5.5. Let $(t,x) \in [0,T) \times \mathbb{R}^m$ be a given initial condition.

- A control $\widehat{\alpha} \in \mathcal{A}(t,x)$ is called **optimal control** if $v(t,x) = J(t,x,\widehat{\alpha})$.
- A control process α of the form $\alpha_s = a(s, X_s^{t,x})$ for some measurable function $a\colon [0,T] \times \mathbb{R}^n \to A$ is called **Markov control**.

Dynamic programming principle

Lecture 10

In order to solve the stochastic control problem, we rely on the so-called *dynamic programming principle (DPP)*. The DPP is a fundamental principle in the theory of stochastic control, which is formulated in the next theorem for our case of a controlled stochastic differential equation.

Theorem 5.6 (Dynamic programming principle (DPP))**.** *For $(t,x) \in [0,T] \times \mathbb{R}^n$ one has*

$$\begin{aligned} v(t,x) &= \sup_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(\theta, X_\theta^{t,x})\Big] \\ &= \sup_{\alpha \in \mathcal{A}(t,x)} \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(\theta, X_\theta^{t,x})\Big]. \end{aligned}$$

Remark 5.7 (Interpretation of the DPP)*.* In words, the DPP states that the optimization problem can be split in two parts: an optimal control on the whole time interval $[t,T]$ can be obtained by first searching for an optimal control from the time θ given the value $X_\theta^{t,x}$ (i.e. computing $v(\theta, X_\theta^{t,x})$) and then maximizing over the controls on $[t,\theta]$ the quantity

$$\mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(\theta, X_\theta^{t,x})\Big].$$

Proof of Theorem 5.6. Fix $(t,x) \in [0,T] \times \mathbb{R}^n$. First note that

$$v(t,x) \geq \sup_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(\theta, X_\theta^{t,x}) \Big]$$

and

$$v(t,x) \leq \sup_{\alpha \in \mathcal{A}(t,x)} \sup_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(\theta, X_\theta^{t,x}) \Big]$$

are obvious as taking $\theta := T \in \mathcal{T}_{t,T}$ gives $v(T, X_T^{t,x}) = g(X_T^{t,x})$.

Step 1: Let $\alpha \in \mathcal{A}(t,x)$ be a control. For any stopping time $\theta \in \mathcal{T}_{t,T}$ we observe that

$$X_s^{t,x} = X_s^{\theta, X_\theta^{t,x}}, \quad s \in [\theta, T] \quad (s \geqslant t),$$

which is actually not trivial but can be deduced from the uniqueness and the Markov structure of the SDE (5.1). Hence, we obtain

$$\begin{aligned}
J(t,x,\alpha) &:= \mathbb{E}\Big[\int_t^T f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + g(X_T^{t,x}) \Big] \\
&= \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + \int_\theta^T f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + g(X_T^{t,x}) \Big] \\
&= \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + \int_\theta^T f(s, X_s^{\theta, X_\theta^{t,x}}, \alpha_s)\,\mathrm{d}s + g(X_T^{\theta, X_\theta^{t,x}}) \Big] \\
&= \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + J(\theta, X_\theta^{t,x}, \alpha) \Big].
\end{aligned}$$

Since $\theta \in \mathcal{T}_{t,T}$ was arbitrary and

$$J(t,x,\alpha) \leq v(t,x) := \sup_{\alpha \in \mathcal{A}(t,x)} J(t,x,\alpha),$$

we conclude that

$$\begin{aligned}
J(t,x,\alpha) &\leq \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(\theta, X_\theta^{t,x}) \Big] \\
&\leq \sup_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(\theta, X_\theta^{t,x}) \Big].
\end{aligned}$$

Taking the supremum over α in the left-hand side yields

$$v(t,x) \leq \sup_{\alpha \in \mathcal{A}(t,x)} \inf_{\theta \in \mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(\theta, X_\theta^{t,x}) \Big].$$

Step 2: Let us fix an arbitrary control $\alpha \in \mathcal{A}(t,x)$ and $\theta \in \mathcal{T}_{t,T}$. By definition of the value function v and for every $\omega \in \Omega$ and for every $\varepsilon > 0$ there exists an $\alpha^{\varepsilon,\omega} \in \mathcal{A}(\theta(\omega), X_{\theta(\omega)}^{t,x}(\omega))$ such that

$$v(\theta(\omega), X_{\theta(\omega)}^{t,x}(\omega)) - \varepsilon \leq J(\theta(\omega), X_{\theta(\omega)}^{t,x}(\omega), \alpha^{\varepsilon,\omega}).$$

Now we define

$$\widehat{\alpha}_s(\omega) := \begin{cases} \alpha_s(\omega) & \text{if } s \in [0, \theta(\omega)] \\ \alpha_s^{\varepsilon,\omega}(\omega) & \text{if } s \in [\theta(\omega), T] \end{cases}, \quad s \in [0,T].$$

It can be shown that the process $(\widehat{\alpha}_s)_{t\in[0,T]}$ is adapted and lies in $\mathcal{A}(t,x)$. This technical measurability issue is dropped here. Hence, we deduce that

$$\begin{aligned} v(t,x) &\geq J(t,x,\widehat{\alpha}) \\ &= \mathbb{E}\Big[\int_t^T f(s,X_s^{t,x},\widehat{\alpha}_s)\,\mathrm{d}s + g(X_T^{t,x})\Big] \\ &= \mathbb{E}\Big[\int_t^\theta f(s,X_s^{t,x},\alpha_s)\,\mathrm{d}s + J(\theta,X_\theta^{t,x},\alpha^\varepsilon)\Big] \\ &\geq \mathbb{E}\Big[\int_t^\theta f(s,X_s^{t,x},\alpha_s)\,\mathrm{d}s + v(\theta,X_\theta^{t,x})\Big] - \varepsilon. \end{aligned}$$

Since $\alpha \in \mathcal{A}(t,x)$, $\theta \in \mathcal{T}_{t,T}$ and $\varepsilon > 0$ are arbitrary, we obtain

$$v(t,x) \geq \sup_{\alpha\in\mathcal{A}(t,x)} \sup_{\theta\in\mathcal{T}_{t,T}} \mathbb{E}\Big[\int_t^\theta f(s,X_s^{t,x},\alpha_s)\,\mathrm{d}s + v(\theta,X_\theta^{t,x})\Big].$$

□

Remark 5.8. The DPP (Theorem 5.6) can be equivalently formulated as follows: For $(t,x) \in [0,T]\times\mathbb{R}^n$ we have that

(i) For all $\alpha \in \mathcal{A}(t,x)$ and $\theta \in \mathcal{T}_{t,T}$ it holds:

$$v(t,x) \geq \mathbb{E}\Big[\int_t^\theta f(s,X_s^{t,x},\alpha_s)\,\mathrm{d}s + v(\theta,X_\theta^{t,x})\Big].$$

(ii) For all $\varepsilon > 0$, there exists an $\alpha \in \mathcal{A}(t,x)$ such that for all $\theta \in \mathcal{T}_{t,T}$:

$$v(t,x) - \varepsilon \leq \mathbb{E}\Big[\int_t^\theta f(s,X_s^{t,x},\alpha_s)\,\mathrm{d}s + v(\theta,X_\theta^{t,x})\Big].$$

As a consequence of the DPP (Theorem 5.6) we observe that

$$v(t,x) = \sup_{\alpha\in\mathcal{A}(t,x)} \mathbb{E}\Big[\int_t^\theta f(s,X_s^{t,x},\alpha_s)\,\mathrm{d}s + v(\theta,X_\theta^{t,x})\Big] \tag{5.3}$$

for any stopping time $\theta \in \mathcal{T}_{t,T}$.

Hamilton-Jacobi-Bellman equation

As a next step towards solving the stochastic control problem (5.2), we derive a partial differential equation characterising the value function of the stochastic control problem. This PDE is called Hamilton-Jacobi-Bellman equation (HJB). The HJB equation is an infinitesimal version of the dynamic programming principle: It describes the local behavior of the value function v when sending θ to t in (5.3).

Let us derive formally the HJB equation. For this purpose we take the stopping time $\theta := t+h$ and a constant control $\alpha_s = a$ for all $s \in [0,T]$ and for some arbitrary $a \in A$. Using Remark 5.8 (i) we deduce that

$$v(t,x) \geq \mathbb{E}\Big[\int_t^{t+h} f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(t+h, X_{t+h}^{t,x})\Big],$$

which gives

$$0 \geq \mathbb{E}\Big[\int_t^{t+h} f(s, X_s^{t,x}, \alpha_s)\,\mathrm{d}s + v(t+h, X_{t+h}^{t,x}) - v(t,x)\Big].$$

Assuming $v\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}$ is sufficiently smooth, we may apply Itô formula to deduce (recalling $X_t^{t,x} = x$) that

$$v(t+h, X_{t+h}^{t,x}) - v(t,x) = \int_t^{t+h} \Big(\frac{\partial v}{\partial t} + \mathcal{L}^a v\Big)(s, X_s^{t,x})\,\mathrm{d}s + \text{ (local) martingale},$$

where

$$\mathcal{L}^a v := b(x,a) \cdot \mathrm{D}_x v + \frac{1}{2}\mathrm{tr}(\sigma(t,a)\sigma^T(t,a)\mathrm{D}_x^2 v)$$

with

$$\mathrm{D}_x v := \begin{pmatrix} \frac{\partial v}{\partial x_1} \\ \vdots \\ \frac{\partial v}{\partial x_n} \end{pmatrix} \quad \text{and} \quad \mathrm{D}_x^2 v := \begin{pmatrix} \frac{\partial^2 v}{\partial x_1^2} & \cdots & \frac{\partial^2 v}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 v}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 v}{\partial x_n^2} \end{pmatrix}.$$

Here M^T denotes the transpose of the matrix M and $\mathrm{tr}(M)$ the trace of the matrix M.

Hence, we get

$$0 \geq \mathbb{E}\Big[\int_t^{t+h} \Big[\Big(\frac{\partial v}{\partial t} + \mathcal{L}^a v\Big)(s, X_s^{t,x}) + f(s, X_s^{t,x}, \alpha_s)\Big]\,\mathrm{d}s\Big].$$

Dividing both sides by $h > 0$, we conclude that

$$0 \geq \mathbb{E}\Big[\frac{1}{h}\int_t^{t+h} \Big[\Big(\frac{\partial v}{\partial t} + \mathcal{L}^a v\Big)(s, X_s^{t,x}) + f(s, X_s^{t,x}, \alpha_s)\Big]\,\mathrm{d}s\Big].$$

Letting h tend to 0 and keeping in mind the mean value theorem and $X_t^{t,x} = x$, reveals

$$0 \geq \frac{\partial v}{\partial t}(t,x) + \mathcal{L}^a v(t,x) + f(t,x,a).$$

Because this inequality holds for every $a \in A$, we obtain the inequality

$$-\frac{\partial v}{\partial t}(t,x) - \sup_{a \in A}\Big(\mathcal{L}^a v(t,x) + f(t,x,a)\Big) \geq 0. \tag{5.4}$$

On the other hand, let us suppose there exists an optimal control $\alpha^* \in \mathcal{A}(t,x)$ and let X^* be the corresponding solution of the SDE (5.1) given α^* and starting at time t from x, i.e. $X_t^* = x$. Then, applying the DPP formulation of (5.3), we see that

$$v(t,x) = \mathbb{E}\Big[\int_t^{t+h} f(s, X_s^*, \alpha_s^*)\,\mathrm{d}s + v(t+h, X_{t+h}^*)\Big].$$

By the same arguments as above, we conclude that

$$-\frac{\partial v}{\partial t}(t,x) - \mathcal{L}^{a_t^*} v(t,x) - f(s,x,a_t^*) = 0. \tag{5.5}$$

Combining (5.4) and (5.5) suggests that v should fulfill

$$-\frac{\partial v}{\partial t}(t,x) - \sup_{a\in A}\Big(\mathcal{L}^a v(t,x) + f(t,x,a)\Big) = 0, \quad (t,x)\in[0,T)\times\mathbb{R}^n,$$

assuming that the supremum in a is finite. Often, this partial differential equation (PDE) is rewritten as Lecture 11

$$-\frac{\partial v}{\partial t}(t,x) - H(t,x,\mathrm{D}_x v(t,x),\mathrm{D}_x^2 v(t,x)) = 0, \quad (t,x)\in[0,T)\times\mathbb{R}^n, \tag{5.6}$$

where

$$H(t,x,p,M) := \sup_{a\in A}\Big(b(x,a)\cdot p + \frac{1}{2}\mathrm{tr}(\sigma\sigma^T(x,a)M) + f(t,x,a)\Big),$$

for $(t,x,p,M)\in[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^{n\times n}$. The PDE (5.6) is called **Hamilton-Jacobi-Bellman (HJB) equation** (or **dynamic programming equation**). The function H is called **Hamiltonian** of the associated stochastic control problem.

Finally, we need to choose a suitable terminal condition for the HJB equation (5.6) so that it really characterizes the value function v of the stochastic control problem. Keeping the definition of the value function v as given in (5.2) in mind, the canonical terminal condition of the HJB equation (5.6) is

$$v(T,x) = g(x), \quad \text{for all } x\in\mathbb{R}^n.$$

Verification theorem

Based on the HJB equation, we found a candidate for the value function v associated to the stochastic control problem. Assuming that there exists a sufficiently smooth w to the HJB equation (5.6), we still need to verify that this solution coincides with the value function of the stochastic control problem. This is the content of the next theorem, the verification theorem. The proof relies essentially on Itô formula and gives as a byproduct that the optimal control is a Markovian control.

Theorem 5.9 (Verification theorem)**.** *Let w be a function in $C^2([0,T)\times\mathbb{R}^n)\cap C([0,T]\times\mathbb{R}^n)$ such that there exists a constant $C>0$ with*

$$|w(t,x)| \le C(1+|x|^2), \quad \textit{for all } (t,x)\in[0,T]\times\mathbb{R}^n.$$

(i) If w satisfies

$$-\frac{\partial w}{\partial t}(t,x) - \sup_{a\in A}\Big(\mathcal{L}^a w(t,x) + f(s,x,a)\Big) \ge 0, \quad (t,x)\in[0,T)\times\mathbb{R}^n, \tag{5.7}$$
$$w(T,x) \ge g(x), \quad x\in\mathbb{R}^n,$$

then $w(t,x)\ge v(t,x)$ for all $(t,x)\in[0,T]\times\mathbb{R}^n$.

(ii) Suppose that

(a) w satisfies (5.7) with $w(T,x) = g(x)$ for all $x \in \mathbb{R}^n$,

(b) there exists a measurable function $\widehat{\alpha}(t,x)$, $(t,x) \in [0,T) \times \mathbb{R}^n$, with values in A such that

$$-\frac{\partial w}{\partial t}(t,x) - \Big(\mathcal{L}^{\widehat{a}(t,x)} w(t,x) + f(s,x,\widehat{a}(t,x))\Big) = -\frac{\partial w}{\partial t}(t,x) - \sup_{a \in A}\Big(\mathcal{L}^a w(t,x) + f(s,x,a)\Big) = 0,$$

(c) the stochastic differential equation

$$\mathrm{d}X_s = b(X_s, \widehat{\alpha}(s,X_s))\,\mathrm{d}s + \sigma(X_s, \widehat{\alpha}(s,X_s))\,\mathrm{d}W_s, \quad s \in [t,T],$$

admits a unique solution, denoted by $\widehat{X}_s^{t,x}$, given the initial condition $X_t = x$,

(d) the process $(\widehat{\alpha}(s,\widehat{X}_s^{t,x}))_{s\in[t,T]}$ lies in $\mathcal{A}(t,x)$.

Then, one has

$$w(t,x) = v(t,x), \quad (t,x) \times [0,T] \times \mathbb{R}^n,$$

and $(\widehat{\alpha}(s,\widehat{X}_s^{t,x}))_{t\in[0,T]}$ is an optimal Markovian control.

Proof. (i) Let $(t,x) \in [0,T) \times \mathbb{R}^n$, $\alpha \in \mathcal{A}(t,x)$ and $\tau \in \mathcal{T}_{t,T}$ be a stopping time with values in $[t,T]$. Since $w \in C^2([0,T) \times \mathbb{R}^n)$, Itô formula reveals that

$$\begin{aligned} w(s\wedge\tau, X_{s\wedge\tau}^{t,x}) = w(t,x) &+ \int_t^{s\wedge\tau} \Big(\frac{\partial w}{\partial t}(u,X_u^{t,x}) + \mathcal{L}^{\alpha_u} w(u,X_u^{t,x})\Big)\,\mathrm{d}u \\ &+ \int_t^{s\wedge\tau} \mathrm{D}_x w(u,X_u^{t,x})^T \sigma(X_u^{t,x},\alpha_u)\,\mathrm{d}W_u \end{aligned} \tag{5.8}$$

for $s \in [t,T)$. Note, we only know that the stochastic integral is a local martingale.

Let us introduce the localizing sequence

$$\tau_n := \inf\Big\{ s \geq t \,:\, \int_t^s |\mathrm{D}_x w(u,X_u^{t,x})^T \sigma(X_u^{t,x},\alpha_u)|^2\,\mathrm{d}u \geq n \Big\} \wedge T, \quad n \in \mathbb{N},$$

and note that $\tau_n \to T$ as $n \to \infty$. Now for every $n \in \mathbb{N}$ the stopped process

$$\Big(\int_t^{s\wedge\tau_n} \mathrm{D}_x w(u,X_u^{t,x})^T \sigma(X_u^{t,x},\alpha_u)\,\mathrm{d}W_u\Big)_{s\in[t,T]} \quad \text{is a martingale.}$$

Hence, using (5.8) with $\tau = \tau_n$ and taking the expectation, we get

$$\mathbb{E}\big[w(s\wedge\tau_n, X_{s\wedge\tau_n}^{t,x})\big] = w(t,x) + \mathbb{E}\Big[\int_t^{s\wedge\tau_n} \Big(\frac{\partial w}{\partial t}(u,X_u^{t,x}) + \mathcal{L}^{\alpha_u} w(u,X_u^{t,x})\Big)\,\mathrm{d}u\Big].$$

Since w satisfies (5.7), we have

$$\frac{\partial w}{\partial t}(u,X_u^{t,x}) + \mathcal{L}^{\alpha_u} w(u,X_u^{t,x}) \leq -f(X_u^{t,x},\alpha_u), \quad \text{for all } \alpha \in \mathcal{A}(t,x),$$

which leads to

$$\mathbb{E}\big[w(s\wedge\tau_n, X_{s\wedge\tau_n}^{t,x})\big] \leq w(t,x) - \mathbb{E}\Big[\int_t^{s\wedge\tau_n} f(X_u^{t,x},\alpha_u)\,\mathrm{d}u\Big], \quad \text{for all } \alpha \in \mathcal{A}(t,x).$$

Notice that

$$\left|\int_t^{s\wedge\tau_n} f(X_u^{t,x},\alpha_u)\,\mathrm{d}u\right| \le \int_t^T |f(X_u^{t,x},\alpha_u)|\,\mathrm{d}u \tag{5.9}$$

and the right-hand-side term is integrable by the integrability condition of elements in $\mathcal{A}(t,x)$. Moreover, since w satisfies

$$|w(s\wedge\tau_n, X_{s\wedge\tau_n}^{t,x})| \le C\Big(1+\sup_{s\in[t,T]}|X_s^{t,x}|^2\Big) \tag{5.10}$$

and the right-hand-side term is integrable due to the integrability properties of the solutions to stochastic differential equations.

Keeping in mind (5.9) and (5.10), sending $n\to\infty$ the dominated convergence theorem implies that

$$\mathbb{E}[w(s,X_s^{t,x})] \le w(t,x) - \mathbb{E}\Big[\int_t^s f(X_u^{t,x},\alpha_u)\,\mathrm{d}u\Big], \quad \text{for all } \alpha\in\mathcal{A}(t,x).$$

Since w is continuous and $w(T,x)\ge g(x)$ for $x\in\mathbb{R}^n$, by sending $s\to T$ leads again by the dominated convergence theorem to

$$\mathbb{E}[g(X_T^{t,x})] \le \mathbb{E}[w(T,X_T^{t,x})] \le w(t,x) - \mathbb{E}\Big[\int_t^T f(X_u^{t,x},\alpha_u)\,\mathrm{d}u\Big], \quad \text{for all } \alpha\in\mathcal{A}(t,x).$$

As $\alpha\in\mathcal{A}(t,x)$ was arbitrary, we deduce that

$$v(t,x) = \sup_{\alpha\in\mathcal{A}(t,x)} \mathbb{E}\Big[\int_t^T f(s,X_s^{t,x},\alpha_s)\,\mathrm{d}s + g(X_T^{t,x})\Big] \le w(t,x), \quad \text{for all } (t,x)\in[0,T]\times\mathbb{R}^n.$$

(ii) Applying Itô formula to $w(s,\widehat{X}_s^{t,x})$ between $t\in[0,T)$ and $s\in[t,T)$, taking the expectation and using (c), leads to

$$\mathbb{E}[w(s,\widehat{X}_s^{t,x})] = w(t,x) + \mathbb{E}\Big[\int_t^s \frac{\partial w}{\partial t}(u,\widehat{X}_u^{t,x}) + \mathcal{L}^{\widehat{\alpha}(u,\widehat{X}_u^{t,x})} w(u,\widehat{X}_u^{t,x})\,\mathrm{d}u\Big]$$

after an eventual localization for removing the stochastic integral term in the expectation. By the definition of $\widehat{a}(t,x)$ (see (b)) we further get

$$-\frac{\partial w}{\partial t} - \mathcal{L}^{\widehat{a}(t,x)} w(t,x) - f(t,x,\widehat{a}(t,x)) = 0$$

and thus

$$\mathbb{E}\Big[w(s,\widehat{X}_s^{t,x})\Big] = w(t,x) - \mathbb{E}\Big[\int_t^s f(\widehat{X}_u^{t,x},\widehat{\alpha}(u,\widehat{X}_u^{t,x}))\,\mathrm{d}u\Big].$$

Sending s to T and using (a) reveals

$$w(t,x) = \mathbb{E}\Big[\int_t^T f(\widehat{X}_u^{t,x},\widehat{\alpha}(u,\widehat{X}_u^{t,x}))\,\mathrm{d}u + g(\widehat{X}_T^{t,x})\Big] =: J(t,x,\widehat{\alpha}).$$

This shows that

$$w(t,s) = J(t,x,\widehat{\alpha}) \le v(t,x)$$

Hence, together with (i) we have $w=v$ and $\widehat{\alpha}$ is an optimal Markovian control due to (d). □

5.2 Portfolio optimization for power utilities

In this subsection we shall treat portfolio optimizations for an investor who can invest in the one-dimensional Black-Scholes model and whose preferences are represented by a power utility function. We shall solve the portfolio allocation problem and allocation problem regarding wealth and consumption using the developed results about stochastic control theory.

An investor may invest in a financial market modeled by the one-dimensional Black-Scholes model. Recall that the one dimensional Black-Scholes model $(S_t^0, S_t^1)_{t\in[0,T]}$ is given by

- a risk-free asset $(S_t^0)_{t\in[0,T]}$ with the dynamics

$$\mathrm{d}S_t^0 = rS_t^0\,\mathrm{d}t, \quad t\in[0,T], \quad S_0^0 = S_0^0,$$

  where $S_0^0 > 0$, $r\in\mathbb{R}$ is the interest rate,

- a risky asset $(S_t^1)_{t\in[0,T]}$ with the dynamics

$$\mathrm{d}S_t^1 = \mu S_t^1\,\mathrm{d}t + \sigma S_t^1\,\mathrm{d}W_t, \quad t\in[0,T], \quad S_0^1 = S_0^1,$$

  where $S_0^1 > 0$, $\mu\in\mathbb{R}$ is the drift parameter, $\sigma > 0$ is the volatility parameter, and $(W_t)_{t\in[0,T]}$ is a Brownian motion.

Portfolio allocation

At each time $t\in[0,T]$ the investor invests a proportion of π_t of her wealth into the risky asset S_t^1 and a proportion of $(1-\pi_t)$ of her wealth into the risk-free asset S_t^0. Hence, given an initial wealth $x>0$, her wealth process $(V_t)_{t\in[0,T]}$ evolves according to

$$\mathrm{d}V_t = \frac{\pi_t V_t}{S_t^1}\,\mathrm{d}S_t^1 + \frac{(1-\pi_t)V_t}{S_t^0}\,\mathrm{d}S_t^0, \quad t\in[0,T], \quad V_0 = x.$$

Inserting the dynamics of $(S_t^0, S_t^1)_{t\in[0,T]}$, this leads to

$$\mathrm{d}V_t = V_t(\pi_t\mu + (1-\pi_t)r)\,\mathrm{d}t + V_t\pi_t\sigma\,\mathrm{d}W_t, \quad t\in[0,T], \quad V_0 = x. \tag{5.11}$$

The set $\mathcal{A}$ of admissible controls is given by

$$\mathcal{A} := \left\{\pi = (\pi_t)_{t\in[0,T]} : \pi \text{ is } \mathbb{R}\text{-valued and adapted s.t. } \int_0^T |\pi_s|^2\,\mathrm{d}s < \infty \text{ a.s.}\right\}.$$

Note that the integrability condition required in the definition of $\mathcal{A}$ ensures that the stochastic differential equation (5.11) has a unique solution. As before, we denote by $(V_s^{t,x})_{s\in[t,T]}$ the solution of the SDE (5.11) starting from x at time t.

The portfolio allocation problem is to find the optimal investment among the set $\mathcal{A}$ which maximizes the expected power utility of the terminal wealth. More precisely, given the utility function

$$U(x) := \frac{x^p}{p}, \quad x\geq 0, \quad \text{with} \quad p\in(0,1),$$

and initial wealth x at time 0, the investor faces the utility optimization problem of finding $\pi^* \in \mathcal{A}$ such that

$$\mathbb{E}[U(V_T^{\pi^*,0,x})] = \sup_{\pi \in \mathcal{A}} \mathbb{E}[U(V_T^{0,x})].$$

Note that this is the classical setting and portfolio optimization as treated first by Robert C. Merton (1976).

In order to solve this utility optimization problem, we write down the corresponding value function

$$v(t,x) = \sup_{\alpha \in \mathcal{A}(t,x)} \mathbb{E}[U(V_T^{t,x})], \quad (t,x) \in [0,T] \times \mathbb{R}_+. \tag{5.12}$$

Remark 5.10. Since the utility function U is increasing and concave on $(0,\infty)$, the value function v is increasing and concave in $x \in \mathbb{R}_+$ for every $t \in [0,T]$. See Problem Sheet 6.

The HJB equation for the stochastic control problem (5.12) is

$$-\frac{\partial w}{\partial t} - \sup_{\pi \in \mathbb{R}} \left(x(\pi\mu + (1-\pi)r)\frac{\partial w}{\partial x} + \frac{1}{2}x^2\pi^2\sigma^2\frac{\partial^2 w}{\partial x^2} \right) = 0 \tag{5.13}$$

with the terminal condition

$$w(T,x) = U(x) = \frac{x^p}{p}, \quad x \geq 0. \tag{5.14}$$

It turns out that one can find an explicitly smooth solution to (5.13) and (5.14).

Lecture 12

To derive the solution w to this HJB equation, we make the ansatz that the solution is of the form

$$w(t,x) = \beta(t)U(x), \quad (t,x) \in [0,T] \times \mathbb{R}_+,$$

for some function $\beta\colon [0,T] \to \mathbb{R}_+$. In this case, by the terminal condition (5.14)

$$U(x) = w(T,x) = \beta(T)U(x),$$

we need to have

$$\beta(T) = 1.$$

Furthermore, we observe that

$$\frac{\partial w}{\partial t}(t,x) = \beta'(t)U(x) = \beta'(t)\frac{x^p}{p}, \quad \frac{\partial w}{\partial x}(t,x) = \beta(t)U'(x) = \beta(t)x^{p-1}$$

and

$$\frac{\partial^2 w}{\partial x^2}(t,x) = \beta(t)U''(x) = \beta(t)(p-1)x^{p-2}.$$

Plugging the ansatz into (5.13), we get

$$-\beta'(t)\frac{x^p}{p} - \sup_{\pi \in \mathbb{R}} \left(x(\pi\mu + (1-\pi)r)\beta(t)x^{p-1} + \frac{1}{2}x^2\pi^2\sigma^2\beta(t)(p-1)x^{p-2} \right) = 0,$$

which leads to

$$-\beta'(t) - \beta(t)p \sup_{\pi \in \mathbb{R}} \left((\pi\mu + (1-\pi)r) + \frac{1}{2}\pi^2\sigma^2(p-1) \right) = 0.$$

Therefore, β has to solve

$$\beta'(t) + \beta(t)\rho = 0, \quad t \in [0,T], \quad \beta(T) = 1, \tag{5.15}$$

where

$$\rho := p \sup_{\pi\in\mathbb{R}} \left((\pi(\mu - r) + r) - \frac{1}{2}\pi^2\sigma^2(1-p) \right).$$

By the differential equation (5.15) we get

$$\beta(t) = \exp(\rho(T-t)).$$

and

$$w(t,x) = \exp(\rho(T-t))U(x), \quad (t,x) \in [0,T] \times \mathbb{R}_+, \tag{5.16}$$

is a smooth solution to the HJB equation (5.13) with terminal condition (5.14).

Now we need to calculate ρ. Since the function

$$\pi \mapsto \pi(\mu - r) + r) - \frac{1}{2}\pi^2\sigma^2(1-p)$$

is strictly concave and twice continuously differentiable, we also explicitly derive its maximum and where it is attained: The maximum is attained at

$$\pi^* := \frac{\mu - r}{\sigma^2(1-p)}$$

and

$$\rho = \frac{(\mu - r)^2}{2\sigma^2}\frac{p}{1-p} + rp.$$

Moreover, the wealth process associated to the constant control π^* is

$$\mathrm{d}V_t = V_t(\pi^*\mu + (1-\pi^*)r)\,\mathrm{d}t + V_t\pi^*\sigma\,\mathrm{d}W_t, \quad t \in [0,T], \quad V_0 = x,$$

which admits a unique solution, and $\pi^* \in \mathcal{A}(0,x)$. All together we can apply the verification theorem (Theorem 5.9), which reveals that w as defined in (5.16) indeed coincides with the value function v as defined by (5.12) and that the optimal control is π^*, which represent the optimal proportion of the wealth to invest in the risky asset.

Optimizing wealth and consumption

Let us come back to our second introductory example: An investor wants to not only maximize her profits but also the spending of her wealth on consumption. Let $(\pi_t)_{t\in[0,T]}$ be the proportion of her wealth invested in the risky asset S_t^1 at time t, let c_t be the proportion of her wealth spend for consumption at time t and let $V_0 = x > 0$ be the initial capital. As discussed before, the corresponding wealth process $(V_t)_{t\in[0,T]}$ evolves according to

$$\mathrm{d}V_t^{(\pi,c)} = \frac{(1-\pi_t - c_t)V_t^{(\pi,c)}}{S_t^0}\,\mathrm{d}S_t^0 + \frac{\pi_t V_t^{(\pi,c)}}{S_t^1}\,\mathrm{d}S_t^1 - c_t V_t^{(\pi,c)}\,\mathrm{d}t, \quad t \in [0,T].$$

Inserting the dynamics of $(S_t^0, S_t^1)_{t\in[0,T]}$, this leads to

$$\mathrm{d}V_t = ((1-\pi_t - c_t)r + \mu\pi_t - c_t)V_t\,\mathrm{d}t + \pi_t V_t\sigma\,\mathrm{d}W_t, \quad t \in [0,T].$$

The set $\mathcal{A}$ of admissible controls is given by

$$\mathcal{A} := \Big\{(\pi, c) = (\pi_t, c_t)_{t\in[0,T]} \;:\; \pi \text{ is } \mathbb{R}^2\text{-valued and adapted s.t. } \int_0^T |\pi_s|^2 + |c_s|^2 \,\mathrm{d}s < \infty \text{ a.s.}\Big\}.$$

Given the initial capital $x > 0$ and the power utility function

$$U(x) := \frac{x^p}{p}, \quad x \geq 0, \quad \text{for } p \in (0,1),$$

the investor wants to find the optimal investment $(\pi_t^*)_{t\in[0,T]}$ and consumption $(c_t^*)_{t\in[0,T]}$, that is

$$\mathbb{E}\Big[\int_0^T U(c_t^* V_t^{(\pi^*,c^*),0,x})\,\mathrm{d}t + U(V_T^{(\pi^*,c^*),0,x})\Big] = \sup_{(\pi,c)\in\mathcal{A}} \mathbb{E}\Big[\int_0^T U(c_t V_t^{0,x})\,\mathrm{d}t + U(V_T^{0,x})\Big].$$

In order to solve this utility optimization problem, we write down the corresponding value function

$$v(t,x) = \sup_{(\pi,c)\in\mathcal{A}(t,x)} \mathbb{E}\Big[\int_t^T U(c_s V_s^{t,x})\,\mathrm{d}s + U(V_T^{t,x})\Big], \quad (t,x) \in [0,T]\times\mathbb{R}_+. \tag{5.17}$$

The HJB equation for the stochastic control problem (5.17) is

$$-\frac{\partial w}{\partial t} - \sup_{(\pi,c)\in\mathbb{R}^2}\Big(((1-\pi-c)r + \mu\pi - c)x\frac{\partial w}{\partial x} + \frac{1}{2}x^2\pi^2\sigma^2\frac{\partial^2 w}{\partial x^2} + \frac{c^p}{p}x^p\Big) = 0 \tag{5.18}$$

with the terminal condition

$$w(T,x) = \frac{x^p}{p}, \quad x \geq 0. \tag{5.19}$$

Again, it turns out that one can find an explicitly smooth solution to (5.18) and (5.19).

To derive the solution w to this HJB equation, we make the ansatz that the solution is of the form

$$w(t,x) = \gamma^{1-p}(t)\frac{x^p}{p}, \quad (t,x)\in[0,T]\times\mathbb{R}_+,$$

for some function $\gamma\colon [0,T]\to\mathbb{R}_+$. In this case, by the terminal condition (5.19)

$$U(x) = w(T,x) = \gamma^{1-p}(T)\frac{x^p}{p},$$

we need to have

$$\gamma(T) = 1.$$

Furthermore, we observe that

$$\frac{\partial w}{\partial t}(t,x) = \gamma'(t)\gamma^{-p}(t)(1-p)\frac{x^p}{p}, \quad \frac{\partial w}{\partial x}(t,x) = \gamma^{1-p}(t)x^{p-1}$$

and

$$\frac{\partial^2 w}{\partial x^2}(t,x) = \gamma(t)^{1-p}(p-1)x^{p-2}.$$

Plugging the ansatz into (5.18), we get

$$-\gamma'(t)\gamma^{-p}(t)(1-p)\frac{x^p}{p}-\sup_{(\pi,c)\in\mathbb{R}^2}\left(((1-\pi-c)r+\mu\pi-c)\gamma^{1-p}(t)x^p+\frac{1}{2}\pi^2\sigma^2\gamma^{1-p}(t)(p-1)x^p+\frac{c^p}{p}x^p\right)=0,$$

which leads to

$$-\gamma'(t)-\frac{p}{1-p}\sup_{(\pi,c)\in\mathbb{R}^2}\left(((1-\pi-c)r+\mu\pi-c)\gamma(t)+\frac{1}{2}\pi^2\sigma^2\gamma(t)(p-1)+\frac{c^p}{p}\gamma^p(t)\right)=0. \quad (5.20)$$

This time we start by calculating the supremum, that is, we are looking for the maximum of the function

$$\begin{aligned}(\pi,c)\mapsto((1-\pi-c)r+\mu\pi-c)\gamma(t)+\frac{1}{2}\pi^2\sigma^2\gamma(t)(p-1)+\frac{c^p}{p}\gamma^p(t)\\ =\frac{1}{2}\pi^2\sigma^2(p-1)\gamma(t)+\pi(\mu-r)\gamma(t)-c(1+r)\gamma(t)+c^p\frac{1}{p}\gamma^p(t)+r\gamma(t),\end{aligned}$$

which is attained at

$$\pi^*:=\frac{r-\mu}{\sigma^2(p-1)} \quad \text{and} \quad c_t^*:=(1+r)^{\frac{1}{p-1}}\gamma^{-1}(t)$$

and given by

$$\frac{1}{2}(\pi^*)^2\sigma^2(p-1)\gamma(t)+\pi^*(\mu-r)\gamma(t)-(1+r)(1+r)^{\frac{1}{p-1}}+(1+r)^{\frac{p}{p-1}}\frac{1}{p}+r\gamma(t).$$

Hence, equation (5.20) can be rewritten to

$$\gamma'(t)+\frac{p}{1-p}\left(\frac{1}{2}(\pi^*)^2\sigma^2(p-1)\gamma(t)+\pi^*(\mu-r)\gamma(t)+r\gamma(t)-(1+r)(1+r)^{\frac{1}{p-1}}+(1+r)^{\frac{p}{p-1}}\frac{1}{p}\right)=0$$

and, setting

$$\rho_1:=\frac{p}{1-p}\left(\frac{1}{2}(\pi^*)^2\sigma^2(p-1)+\pi^*(\mu-r)+r\right)$$

and

$$\rho_2:=\frac{p}{1-p}\left((1+r)^{\frac{p}{p-1}}\frac{1}{p}-(1+r)(1+r)^{\frac{1}{p-1}}\right),$$

we further get

$$\gamma'(t)+\rho_1\gamma(t)+\rho_2=0 \quad (5.21)$$

with terminal condition $\gamma(T)=1$. Solving (5.21) gives

$$\gamma(t)=\left(1-\frac{\rho_2}{\rho_1}\right)\exp(-\rho_1(T-t))-\frac{\rho_2}{\rho_1},$$

which reveals that

$$w(t,x)=\left(\left(1-\frac{\rho_2}{\rho_1}\right)\exp(-\rho_1(T-t))-\frac{\rho_2}{\rho_1}\right)^{1-p}\frac{x^p}{p} \quad (5.22)$$

is a smooth solution to the HJB equation (5.18) with terminal condition (5.19).

Moreover, the wealth process associated to the control (π^*, c^*) is

$$\mathrm{d}V_t = ((1-\pi^* - c_t^*)r + \mu\pi_t - c_t^*)V_t\,\mathrm{d}t + \pi^* V_t \sigma\,\mathrm{d}W_t \quad t\in[0,T], \quad V_0 = x,$$

which admits a unique solution, and $(\pi^*, c^*) \in \mathcal{A}(0,x)$. All together we can apply the verification theorem (Theorem 5.9), which implies that w as given in (5.22) indeed coincides with the value function v as defined by (5.17) and that the optimal control is (π^*, c^*), which represents the optimal proportion of the wealth to invest in the risky asset and the optimal proportion of the wealth spend on consumption.

Exercise. With the ideas developed in this subsection you can also solve the following related optimization problem regarding wealth and consumption: As before, assuming the wealth $(V_t)_{t\in[0,T]}$ evolves according to the SDE

$$\mathrm{d}V_t = ((1-\pi_t - c_t)r + \mu\pi_t - c_t)V_t\,\mathrm{d}t + \pi_t V_t \sigma\,\mathrm{d}W_t, \quad t\in[0,T], \quad V_0 = x,$$

and given the power utility function

$$U(x) := \frac{x^p}{p}, \quad x\geq 0, \quad p\in(0,1),$$

the investor wants to find the optimal investment $(\pi_t^*)_{t\in[0,T]}$ and consumption $(c_t^*)_{t\in[0,T]}$, that is

$$\mathbb{E}\Big[\int_0^T U(c_t^* V_t^{(\pi^*,c^*),0,x})\,\mathrm{d}t\Big] = \sup_{(\pi,c)} \mathbb{E}\Big[\int_0^T U(c_t V_t^{0,x})\,\mathrm{d}t\Big].$$

6 Term structure models

Lecture 13

In the previous chapters we typically assumed that there is one risk-free asset modelled with a constant interest rate r, cf. the risk-free asset of the one-dimensional Black-Scholes model:

$$S_t^0 = S_0^0 \exp(rt), \quad t\in[0,T]. \tag{6.1}$$

However, on real financial markets the interest rate is changing over time depending on the market situation and, even more, real financial markets do not have a single risk-free asset or a single interest rate. Instead, there are interest rate products (or fixed income products) like traded bonds with different maturities. In this chapter we study more advanced modeling regarding interest rates and interest rate products. Since these financial quantities have different statistical properties than prices of stocks, their modeling requires a particularly treatment.

We will always make the assumption that the financial models $S = (S_t^0,\dots,S_t^d)_{t\in[0,T]}$ satisfies NFLVR (i.e. the financial market is, in particular, arbitrage-free) and there is no default risk of financial products.

6.1 Zero-coupon bonds and short rate

Let us start by introducing zero-coupon bonds, which constitute a basic object in interest rate theory.

Definition 6.1. A **zero-coupon bond** (also called **T-bond**) is a contract which guarantees the holder the payment of 1 euro (or dollar or another currency) at the maturity T. The price process of a zero-coupon bond with maturity T is denoted by $(B(t,T))_{t\in[0,T]}$. There are no interest payments before maturity.

Remark 6.2. Generally speaking, a **term structure** is a function that relates a certain financial variable or parameter to its maturity. The prototypical example is the term-structure of zero-coupon bond prices $(B(t,T))_{t\in[0,T]}$.

In the following we make the **idealized assumptions**:

- There exists a liquid market for T-bonds for every $T > 0$.
- $B(T,T) = 1$ for all T.
- $B(t,T)$ is differentiable in T.

Note, on real-world markets one could have $B(T,T) < 1$ if the issuer of the bond defaults.

Question. *How are these idealized assumptions represented the bond market by a simple dynamics like*

$$S_t^0 = S_0^0 \exp\left(\int_0^t r_u \,\mathrm{d}u\right), \quad t \in [0,T], \quad \textit{for a suitable process } (r_t)_{t\in[0,T]}, \tag{6.2}$$

as we did throughout the lecture course so far? More specifically, how do these idealized assumptions fit together with our previous assumption in the Black-Scholes model, where we assumed that the risk-free asset evolves according to (6.1)?

Assuming there is a liquid market for T-bonds, through buying and selling zero-coupon bonds, one can create a simple rate for an investment over an interval $[S,T]$. A **simple rate** L is paid for a time period $[S,T]$ meaning that x euros invested at time S leads to

$$y = x(1 + L(T-S))$$

euros at time $T > S$, which implies that

$$L = \frac{1}{(T-S)}\frac{y-x}{x}. \tag{6.3}$$

Thus a simple rate can be achieved by forward rate agreements.

Definition 6.3. A prototypical **forward rate agreement** (FRA) is a contract involving three times $t < S < T$, where t is the current time, $S > t$ the expiry time and $T > S$ the maturity with the following structure:

- At t: sell one S-bond and buy $\frac{B(t,S)}{B(t,T)}$ T-bonds.
- At S: pay one euro.
- At T: receive $\frac{B(t,S)}{B(s,T)}$ euros.

At time t the net cash flow of this investment is equal to

$$B(t,S) - \frac{B(t,S)}{B(t,T)} B(t,T) = 0.$$

Prevailing at t, this creates a simple rate L over the time interval $[S,T]$ given by

$$L = \frac{1}{(T-S)} \frac{\frac{B(t,S)}{B(t,T)} - 1}{1} = \frac{1}{(T-S)} \Big(\frac{B(t,S)}{B(t,T)} - 1 \Big),$$

where we used formula (6.3). This leads to the **simple forward rate** for $[S,T]$ defined as

$$F(t,S,T) := \frac{1}{T-S} \Big(\frac{B(t,S)}{B(t,T)} - 1 \Big),$$

as implied by bond prices at time t. Note, by the law of one price, $F(t,S,T)$ is the only risk-free interest rate that is consistent the assumption of no arbitrage.

The yield from the above forward rate agreement can also be expressed in terms of continuously compounded rates. Following (6.1), a **continuously compounded rate** r paid on an investment of x euros at time S leads to a payoff of

$$y = x \exp(r(T-S)) \quad \text{euros} \tag{6.4}$$

at time $T > S$. Combing (6.3) and (6.4) gives

$$x(1 + L(T-S)) = y = x \exp(r(T-S)),$$

and, hence, we see that the simple rate L from (6.3) corresponds to a continuously compounded rate of

$$r = \frac{1}{T-S} \log\big(1 + L(T-S)\big).$$

Therefore, the simple rate $F(t,S,T)$ corresponds to the **continuously compounded forward rate**

$$f(t,S,T) := \frac{1}{T-S} \log(1 + F(t,S,T)(T-S)) = \frac{1}{T-S} \log\Big(\frac{B(t,S)}{B(t,T)} \Big).$$

Assuming (as we do) that bonds of any maturity T are traded, we can consider the limit $S \to T$. The limit

$$\begin{aligned}
f(t,T) &:= \lim_{\varepsilon \downarrow 0} f(t, T-\varepsilon, T) \\
&= \lim_{\varepsilon \downarrow 0} \frac{1}{T-(T-\varepsilon)} \log\Big(\frac{B(t,T-\varepsilon)}{B(t,T)} \Big) \\
&= \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \big(\log(B(t,T-\varepsilon) - \log(B(t,T)) \\
&= -\frac{\mathrm{d}}{\mathrm{d}T} \log(B(t,T))
\end{aligned}$$

is called **(instantaneous) forward rate**. It corresponds to the interest rate paid for an investment in an infinitesimal time period abound T in the future as implied by the bond prices at time $t < T$. Notice that

$$\begin{aligned}\int_t^T f(t,s)\,\mathrm{d}s &= \int_t^T \Big(-\frac{\mathrm{d}}{\mathrm{d}s}\log(B(t,s))\Big)\,\mathrm{d}s \\ &= -\log(B(t,T)) + \log(B(t,t)).\end{aligned}$$

Therefore, taking the exponential and using $B(t,t) = 1$, we get

$$B(t,T) = \exp\Big(-\int_t^T f(t,s)\,\mathrm{d}s\Big), \quad t \in (0,T], \tag{6.5}$$

where $T > 0$ stands for the maturity.

Now, let us start with 1 euro at time 0 and invest successively into just maturing bonds up to time t. Say we re-invest all the cash at the times $0 = t_0 < t_1 < \cdots < t_n \le t$, that means, at time t_i we invest all the cash in bonds starting at t_i with maturity t_{i+1} for the price $B(t_i, t_{i+1})$. This trading strategy leads to the yield $\frac{1}{B(t_0,t_1)}$ at time t_1, the yield $\frac{1}{B(t_0,t_1)B(t_1,t_2)}$ and, thus, the yield

$$\Pi_{i=0}^{n-1} B(t_i, t_{i+1})^{-1} = \Pi_{i=0}^{n-1} \exp\Big(\int_{t_i}^{t_{i+1}} f(t_i,s)\,\mathrm{d}s\Big) \quad \text{at time } t,$$

where we used the identity (6.5). Heuristically, sending $\max_{i=1,\dots,n-1}|t_{i+1} - t_i| \to 0$, this roll-over portfolio reveals that 1 euro grows to

$$S_t^0 := \exp\Big(\int_0^t f(s,s)\,\mathrm{d}s\Big) = \exp\Big(\int_0^t r_s\,\mathrm{d}s\Big), \tag{6.6}$$

where $r_s := f(s,s)$. The process $(r_s)_{s\in[0,T]}$ is called **short rate** (or, roughly speaking, *interest rate*). Note, the roll-over portfolio coincides with the usual risk-free asset $(S_t^0)_{t\in[0,T]}$ (see e.g. (6.2)) or, when setting $r_t = r$ for all $t \in [0,T]$ and some constant r, with the risk-free asset in the one-dimensional Black-Scholes model.

Let us suppose (from now on) that the underlying market $S = (S_t^0, S_t^1, \dots, S_t^d)_{t\in[0,T]}$ satisfies NFLVR and is complete, where $(S_t^0)_{t\in[0,T]}$ denotes the yields generated by the above roll-over portfolio as obtained in (6.6). To apply the arbitrage theory as developed in Chapter 2, we make the same assumption on the processes $S_t^0, S_t^1, \dots, S_t^d$. In particular, the short rate $r\colon \Omega \times [0,T] \to \mathbb{R}$ is supposed to be an adapted, measurable process satisfying the conditions

$$\sup_{u\in[0,T]} |r_u| < \infty, \qquad \mathbb{P}\text{-a.s.}.$$

As before, we can take $(S_t^0)_{t\in[0,T]}$ as risk-free asset and, thus, the first and second fundamental theorem of asset pricing ensure the existence of a unique equivalent martingale measure Q.

As zero-coupon bound on the financial market $S = (S_t^0, S_t^1, \dots, S_t^d)_{t\in[0,T]}$ can be treated like any other claim, we can derive the price of an zero-coupon bound $B(t,T)$:

Lemma 6.4. *If the short rate $(r_s)_{s\in[0,T]}$ is a stochastic process, then*

$$B(t,T) = \mathbb{E}^Q\Big[\exp\Big(-\int_t^T r_s\,\mathrm{d}s\Big)\Big|\mathcal{F}_t\Big], \quad t\in[0,T], \tag{6.7}$$

where Q is the equivalent martingale measure.

Proof. By Corollary 2.30 we get that

$$B(t,T) = S_t^0\mathbb{E}^Q\Big[\frac{B(T,T)}{S_T^0}\Big|\mathcal{F}_t\Big] = \mathbb{E}^Q\Big[\exp\Big(-\int_t^T r_s\,\mathrm{d}s\Big)\Big|\mathcal{F}_t\Big], \quad t\in[0,T],$$

as $B(T,T)=1$. □

Remark 6.5. If the short rate $(r_s)_{s\in[0,T]}$ is deterministic, then

$$B(t,T) = \exp\Big(-\int_t^T r_s\,\mathrm{d}s\Big), \quad t\in[0,T].$$

To proof this identity, we actually do not need the completeness of the market. Indeed, we can rely on a contrapositive argument, that is, suppose $B(t,T) < \exp\big(-\int_t^T r_s\,\mathrm{d}s\big)$ or $B(t,T) > \exp\big(-\int_t^T r_s\,\mathrm{d}s\big)$ and then construct an arbitrage opportunity.

With respect to the unique equivalent martingale measure Q, we also want that the discounted bond price process is a martingale.

Proposition 6.6. *The discounted bond price process $(\widehat{B}(t,T))_{t\in[0,T]}$, given by*

$$\widehat{B}(t,T) := \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)B(t,T),$$

is a martingale under Q.

Proof. By (6.7) we have

$$\begin{aligned}\widehat{B}(t,T) &:= \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)B(t,T)\\ &= \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)\mathbb{E}^Q\Big[\exp\Big(-\int_t^T r_s\,\mathrm{d}s\Big)\Big|\mathcal{F}_t\Big]\\ &= \mathbb{E}^Q\Big[\exp\Big(-\int_0^T r_s\,\mathrm{d}s\Big)\Big|\mathcal{F}_t\Big]\end{aligned}$$

for $t\in[0,T]$. Hence, $(\widehat{B}(t,T))_{t\in[0,T]}$ is a martingale under Q. □

6.2 Bond pricing PDE

Let us suppose that the underlying short rate process $(r_t)_{t\in[0,T]}$ is given by a general stochastic differential equation

$$\mathrm{d}r_t = \mu(t,r_t)\,\mathrm{d}t + \sigma(t,r_t)\,\mathrm{d}W_t, \quad t\in[0,T],$$

where $(W_t)_{t\in[0,T]}$ is a Brownian motion under the EMM Q and μ,σ are deterministic (Lipschitz continuous in r) functions.

Remark. Like in the case of stochastic volatility models, introducing an additionally source of randomness, as we do here by modeling the short rate by a stochastic process, the financial market $S = (S_t^0, S_t^1, \dots, S_t^d)_{t\in[0,T]}$ might become incomplete, i.e. not every claim is replicable.

Due to the Markov structure of the shot rate $(r_t)_{t\in[0,T]}$ as a solution to a stochastic differential equation, we observe that

$$B(t,T) = \mathbb{E}^Q\Big[\exp\Big(-\int_t^T r_s\,\mathrm{d}s\Big)\Big|\mathcal{F}_t\Big] = F(t,r_t), \quad t\in[0,T],$$

for some deterministic function F. To find F leads us to the bond pricing PDE:

Proposition 6.7. *The bond pricing PDE for $B(t,T) = F(t,r_t)$ is given by*

$$xF(t,x) = \frac{\partial F}{\partial t}(t,x) + \mu(t,x)\frac{\partial F}{\partial x}(t,x) + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 F}{\partial x^2}(t,x),$$

for $t\in[0,T)$, $x\in\mathbb{R}$, with the terminal condition

$$F(T,x) = 1, \quad x\in\mathbb{R}.$$

Proof. For fixed $T>0$ and $t\in[0,T]$, applying Itô formula gives

$$\begin{aligned}
&\mathrm{d}\Big(\exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)B(t,T)\Big)\\
&= -r_t\exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)B(t,T)\,\mathrm{d}t + \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)\mathrm{d}B(t,T)\\
&= -r_t\exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)F(t,r_t)\,\mathrm{d}t + \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)\mathrm{d}F(t,r_t)\\
&= -r_t\exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)F(t,r_t)\,\mathrm{d}t\\
&\quad + \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)\frac{\partial F}{\partial x}(t,r_t)(\mu(t,r_t)\,\mathrm{d}t + \sigma(t,r_t)\,\mathrm{d}W_t)\\
&\quad + \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)\Big(\frac{1}{2}\sigma^2(t,r_r)\frac{\partial^2 F}{\partial x^2}(t,r_t)\,\mathrm{d}t + \frac{\partial F}{\partial t}(t,r_t)\,\mathrm{d}t\Big)\\
&= \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)\frac{\partial F}{\partial x}(t,r_t)\sigma(t,r_t)\,\mathrm{d}W_t + \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)\\
&\qquad\times\Big(-r_t F(t,r_t) + \frac{\partial F}{\partial x}(t,r_t)\mu(t,r_t) + \frac{1}{2}\sigma^2(t,r_t)\frac{\partial^2 F}{\partial x^2}(t,r_t) + \frac{\partial F}{\partial t}(t,r_t)\Big)\,\mathrm{d}t.
\end{aligned}$$

Since $\big(\exp\big(-\int_0^t r_s\,\mathrm{d}s\big)B(t,T)\big)_{t\in[0,T]}$ is a martingale under Q (see Proposition 6.6), we have

$$\mathrm{d}\Big(\exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)B(t,T)\Big) = \exp\Big(-\int_0^t r_s\,\mathrm{d}s\Big)\frac{\partial F}{\partial x}(t,r_t)\sigma(t,r_t)\,\mathrm{d}W_t$$

and

$$-r_t F(t,r_t) + \frac{\partial F}{\partial x}(t,r_t)\mu(t,r_t) + \frac{1}{2}\sigma^2(t,r_t)\frac{\partial^2 F}{\partial x^2}(t,r_t) + \frac{\partial F}{\partial t}(t,r_t) = 0,$$

which implies the claimed bond pricing PDE. The terminal condition follows by the assumption $B(T,T)=1$. □

6.3 Interest rate products

Lecture 14

There is a wide range of different interest rate products (also called fixed income products) traded on financial markets. In the following some examples of them are discussed.

Fixed coupon bonds

Definition 6.8. A **fixed coupon bond** is a contract which pays the nominal of 1 euro at the maturity T_N and additionally coupon payments $c_1, \dots, c_N$ are made at the dates $T_1, \dots, T_N$, respectively.

Under the assumption that zero-coupon bonds are liquidly traded on the financial market, the cash flow of fixed coupon bonds can be replicated by investing in a portfolio consisting of zero-coupon bonds. For simplicity, let us consider a fixed coupon bond which pays at the times

$$T_n = T_0 + n\delta, \quad n = 1, \dots, N,$$

with $T_0 \geq 0$ and $\delta > 0$, the coupon

$$c_n = l\delta, \quad n = 1, \dots, N,$$

where the coupon rate $l > 0$ is the same for all periods. To replicate this fixed coupon bond, we buy c_n-times T_n-bonds and one additional T_N-bond. Indeed, the T_n-bonds yield a payoff of c_n at the times $T_1, \dots, T_{N-1}$ and $(1 + c_N)$ at time T_N, which coincides with the payments of the fixed coupon bond. The value of the replicating portofolio at time $t < T_1$ equals

$$V_t = \sum_{n=1}^{N} c_n B(t, T_n) + B(t, T_N).$$

By the law of one price, this is also the unique price of the fixed coupon bond, which is consistent with the absence of arbitrage. The price of the bond for $t \geq T_1$ can be analogously obtained by summing over the bonds that have not yet expired.

Floating coupon bonds

Definition 6.9. A **floating coupon bond** is a contract which pays the nominal of 1 euro at the maturity T_N and additionally coupon payments of the form

$$c_n = F(T_{n-1}, T_{n-1}, T_n)(T_n - T_{n-1}), \quad n = 1, \dots, N,$$

are made at the dates $T_1, \dots, T_N$, respectively. That means, the coupon rate is determined by the simple forward rate at the beginning of each corresponding period.

Assuming that zero-coupon bonds with maturities $T_0, \dots, T_N$ are traded, also the floating coupon bond can be replicated by a dynamic portfolio consisting of zero-coupon bonds. This is left as an exercise.

Swaps

Swaps allow the holder to trade a fixed interest rate for a floating one. For example, let us consider the *forward swap settle in arrears*. At any of the times $T_1, \dots, T_N$ the holder of the swap pays the fixed interest $l(T_n - T_{n-1})$ ("fixed leg") and receives a variable interest $F(T_{n-1}, T_{n-1}, T_n)(T_n - T_{n-1})$ ("floating leg") in exchange. The swap rate l is chosen such that the contract value is 0 at initiation, i.e. no initial payment must be made.

Bond options

Bond options are financial derivatives acting on some type of bonds. The usual example are Call and Put options on bonds, e.g., the payoff at time T_0 of a call option with strike K on an T_1-bond is

$$(B(T_0, T_1) - K)^+.$$

Whether or not bond options can be replicated by a portfolio of bonds depends on the particular underlying model.

6.4 Term structure modeling

There are various approaches to model term structures on financial markets. Before looking at concrete models, let us discuss six general aspects of this modeling.

- *Primary processes*: Which underlying processes related to interest rates or prices are modeled right from the start and which are derived as a consequence of the model? For example, do we model the short rate $(r_t)_{t\in[0,T]}$ or the whole forward rate curve $(f(t,T))_{t\le T,\, T\ge 0}$?

- *Necessary input*: In the end we would like to have a model for the whole interest rate market. What are the necessary ingredients/assumptions?

- *Discrete vs. continuous tenor*: Do we make the assumption that bonds of all maturities are traded (continuous tenor) or do we assume that bonds with only finitely many maturities are traded (discrete tenor) on the market?

- *Risk-neutral modeling*: Do we start the modeling from a physical measure $\mathbb{P}$ or do we work directly with a martingale measure Q?

- *Stationarity*: Do we want to work with stationary processes? Does that the current time matter or does only the time to maturity matter, for e.g. price formula or hedging strategies?

- *Tractability*: Estimation and calibration of a model is required to work with it. How feasible is the model?

Of course, there are further aspects which contribute to the choice of a specific model, e.g., numbers of sources of randomness.

Short rate and factor model approach

Short rate or factor models represent the money market account $(S_t^0)_{t\in[0,T]}$ by the dynamics

$$\mathrm{d}S_t^0 = r_t S_t^0 \,\mathrm{d}t, \quad t \in [0,T], \quad S_0^0 = S_0^0,$$

where $S_0^0 \in \mathbb{R}$ and $(r_t)_{t\in[0,T]}$ is a stochastic process, usually the solution of a stochastic differential equations.

Example 6.10. In the **Ho-Lee model** the short rate is of the form

$$\mathrm{d}r_t = \kappa \,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \quad t \in [0,T],$$

with $\kappa, \sigma > 0$ and $(W_t)_{t\in[0,T]}$ is a Brownian motion.

In the **Vasiček model** the short rate is of the form

$$\mathrm{d}r_t = (\kappa - \lambda r_t)\,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \quad t \in [0,T],$$

with constants $\kappa \in \mathbb{R}$, $\lambda, \sigma > 0$ and $(W_t)_{t\in[0,T]}$ is a Brownian motion.

In the **Hull-White model** the short rate is of the form

$$\mathrm{d}r_t = (\kappa_t - \lambda r_t)\,\mathrm{d}t + \sigma \,\mathrm{d}W_t, \quad t \in [0,T],$$

for a function $\kappa\colon [0,T] \to \mathbb{R}$, constants $\lambda, \sigma > 0$ and $(W_t)_{t\in[0,T]}$ is a Brownian motion.

Profile: Only the short rate serves as primary process. Short rate models often allow to derive explicit formulas for bond prices, consequently, for forward rates and other interest rate products as well.

Heath-Jarrow-Morton approach

Short-rate models are not always sufficiently flexible to calibrating them to real-world data of term structures. The Heath-Jarrow-Morton (HJM) approach is at the other extreme and models the entire forward rate curve $(f(t,T))_{t\in[0,T],\,T\in[0,T^*]}$ for some $T^* > 0$. For each $T \in [0,T^*]$, the forward rate process is supposed to have the dynamics

$$\mathrm{d}f(t,T) = \alpha(t,T)\,\mathrm{d}t + \sigma(t,T)\,\mathrm{d}W_t, \quad f(0,T) = f(0,T) \quad t \in [0,T], \tag{6.8}$$

where $(W_t)_{t\in[0,T]}$ is Brownian motion and $(f(T,0))_{T\in[0,T^*]}$ is given. Here the coefficients α and σ suitable functions such that the stochastic differential equation (6.8) has a unique solution. The HJM approach does not specify a particular model but rather a general framework. Therefore, most term structure models can be viewed from that perspective.

Recall that we assume that the financial market $S = (S_t^0, S_{t,}^1, \ldots, S_t^d)_{t\in[0,T]}$ satisfies NFLVR and is complete, i.e., there exists a unique equivalent martingale measure Q. To ensure that the forward rate processes $(f(t,T))_{t\in[0,T],\,T\in[0,T^*]}$ do not create arbitrage on the financial market $S = (S_t^0, S_{t,}^1, \ldots, S_t^d)_{t\in[0,T]}$, the coefficients α and σ need to be chosen carefully.

Proposition 6.11. *The the forward rate model (6.8) does not allow arbitrage if and only if*

$$\alpha(t,T) = \sigma(t,T) \int_t^T \sigma(t,s)\,\mathrm{d}s, \quad \textit{for } t \in [0,T], \quad T \in [0,T^*].$$

Proof. Recall, by (6.5) we have

$$B(t,T) = \exp\Big(- \int_t^T f(t,s)\,\mathrm{d}s \Big).$$

Hence, using the model (6.8) and $r_t = f(t,t)$, we get

$$\begin{aligned}\mathrm{d}\Big(-\int_t^T f(t,s)\,\mathrm{d}s\Big) &= f(t,t)\,\mathrm{d}t - \int_t^T \mathrm{d}f(t,s)\,\mathrm{d}s \\ &= r_t\,\mathrm{d}t - \int_t^T \big(\alpha(t,s)\,\mathrm{d}t + \sigma(t,s)\,\mathrm{d}W_t\big)\,\mathrm{d}s \\ &= r_t\,\mathrm{d}t - \alpha^*(t,T)\,\mathrm{d}t - \sigma^*(t,T)\,\mathrm{d}W_t,\end{aligned}$$

where we used in the last line the Fubini and the stochastic Fubini theorem and set

$$\alpha^*(t,T) := \int_t^T \alpha(t,s)\,\mathrm{d}s \quad \text{and} \quad \sigma^*(t,T) := \int_t^T \sigma(t,s)\,\mathrm{d}s.$$

Applying Itô formula to $\widehat{B}(t,T) = \exp\Big(-\int_t^T f(t,s)\,\mathrm{d}s - \int_0^t r_s\,\mathrm{d}s\Big)$ leads to

$$\mathrm{d}\widehat{B}(t,T) = \widehat{B}(t,T)\Big(-\alpha^*(t,T) + \frac{1}{2}\sigma^*(t,T)^2\Big)\,\mathrm{d}t - \sigma^*(t,T)\,\mathrm{d}W_t.$$

Note that $(\widehat{B}(t,T))_{t\in[0,T]}$ is a martingale under Q if and only if

$$-\alpha^*(t,T) + \frac{1}{2}\sigma^*(t,T)^2 = 0,$$

which by definition is equal to

$$\int_t^T \alpha(t,s)\,\mathrm{d}s = \frac{1}{2}\Big(\int_t^T \sigma(t,s)\,\mathrm{d}s\Big)^2$$

which is equivalent to

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)\,\mathrm{d}s.$$

□

Example 6.12 (Hull-White model revisited)**.** Consider the volatility structure

$$\sigma(t,T) = \sigma\exp(-\lambda(T-t)), \quad t \in [0,T],$$

with constants $\sigma > 0$ and $\lambda \geq 0$.

Profile: The HJM approach starts with modeling the whole forward rate curve $(f(t,T))_{t\leq T,\,T\in[0,T^*]}$*, which corresponds to a continuous tenor. While the modeling framework is rather general, the price to pay is a more complicated tractability.*

A Mathematical Foundation

The appendix collects some material which is used during the course but is actually content of other lecture course, as Mathematical Finance, Stochastic processes (WT1) and Stochastic Calculus.

A.1 Conditional expectation

Definition A.1. Let $X \in L^1$ and $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra. A random variable Y is called **conditional expectation** of X given $\mathcal{G}$, denoted by $\mathbb{E}[X|\mathcal{G}] := Y$, if

(i) Y is $\mathcal{G}$-measurable;

(ii) for every $A \in \mathcal{G}$ one has $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$.

If $X, Y \in L^1$, we set $\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)]$.

Next, we summarize some properties of the conditional expectation.

Theorem A.2 (Properties of the conditional expectation)**.** *Let $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ be σ-algebras and $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Then:*

(i) Linearity: *For $\lambda \in \mathbb{R}$ we have $\mathbb{E}[\lambda X + Y|\mathcal{G}] = \lambda\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}]$.*

(ii) Monotonicity: *If $X \geq Y$, then $\mathbb{E}[X|\mathcal{G}] \geq \mathbb{E}[Y|\mathcal{G}]$.*

(iii) If $\mathbb{E}[|XY|] < \infty$ and Y is measurable w.r.t. $\mathcal{G}$, then

$$\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}] \quad \textit{and} \quad \mathbb{E}[Y|\mathcal{G}] = \mathbb{E}[Y|\sigma(Y)] = Y.$$

(iv) Tower property: $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}]$.

(v) Triangle inequality: $|\mathbb{E}[X\,|\mathcal{G}]| \leq \mathbb{E}[|X|\,|\mathcal{G}]$.

(vi) Independence: *If $\sigma(X)$ and $\mathcal{G}$ are independent, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.*

(vii) Fatou's lemma: *If the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ such that $X_n \geq c$, then*

$$\mathbb{E}[\liminf_{n\to\infty} X_n|\mathcal{G}] \leq \liminf_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] \quad \mathbb{P}\text{-}a.s.$$

(viii) Dominated Convergence: *If the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ such that $|X_n| \leq Y$, then*

$$\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] \quad \mathbb{P}\text{-}a.s. \quad \textit{and in } L^1(\mathbb{P}).$$

Proposition A.3 (Conditional Jensen's inequality)**.** *Let $I \subseteq \mathbb{R}$ be an interval, let $\varphi\colon I \to \mathbb{R}$ be convex and let X be an I-valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. If $\mathbb{E}[|X]] < \infty$ and $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra, then*

$$\varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}] \leq \infty.$$

A.2 Filtration, stochastic processes and stopping times

Let us fix an arbitrary set $I \subseteq \mathbb{R}$. We mostly care about $I = \{t_0, \dots, t_N\}$ or $I = \mathbb{N}$ but $I = [0, \infty)$ or $I = [0, T]$ are also allowed.

Definition A.4.

- A family of random variables $(X_t)_{t \in I}$ (with values in $\mathbb{R}^d$) is called **stochastic process** with index set I and range $\mathbb{R}^d$.
- A family of σ-algebras $(\mathcal{F}_t)_{t \in I} \subseteq \mathcal{F}$ is called **filtration** if $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s, t \in I$ with $s \le t$.
- A stochastic process $(X_t)_{t \in I}$ is called **adapted** to $(\mathcal{F}_t)_{t \in I}$ if X_t is $\mathcal{F}_t$-measurable.

Remark A.5. A stochastic process $(X_t)_{t \in I}$ is always adapted to the filtration $\mathcal{F}_t := \sigma(X_s : s \in I,\ s \le t)$, i.e., this is the smallest filtration to which the process $(X_t)_{t \in I}$ is adapted.

Let us denote the $\mathbb{P}$-null sets by

$$\mathcal{N} := \{A \in \mathcal{F} \,:\, \mathbb{P}(A) = 0\}.$$

Definition A.6. The filtration $(\mathcal{F}_t)_{t \in [0,T]}$ satisfies the **usual conditions** if

- $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets $\mathcal{N}$ ("*completeness*"),
- $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s > t} \mathcal{F}_s$ for $t \in [0, T)$ ("*right-continuity*").

As a probabilistic base we fixed a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in I}, \mathbb{P})$, i.e., a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t \in I}$.

Definition A.7. A random variable τ with values in $I \cup \{\infty\}$ is called a **stopping time** (with respect to $(\mathcal{F}_t)_{t \in I}$) if

$$\{\tau \le t\} \in \mathcal{F}_t \quad \text{for any } t \in I.$$

Recall that the filtration reflects the information of each market participants at time t. Hence, whether or not $\{\tau \le t\} \in \mathcal{F}_t$ is true can be determined based on the information available at time t.

Lemma A.8. *Let σ and τ be stopping times. Then:*

(i) $\sigma \vee \tau := \max\{\sigma, \tau\}$ and $\sigma \wedge \tau := \min\{\sigma, \tau\}$ are stopping times.

(ii) If $\sigma, \tau \ge 0$ and $I \subseteq [0, \infty)$ is closed under addition, then $\sigma + \tau$ is a stopping time.

However, in general, $\tau - s$ and $\tau - \sigma$ are <u>not</u> stopping times for $s \in \mathbb{R}_+$.

Definition A.9. Let τ be a stopping time. The **σ-algebra of τ-past** is defined as

$$\mathcal{F}_\tau := \{A \in \mathcal{F} \,:\, A \cap \{\tau \le t\} \in \mathcal{F}_t \text{ for any } t \in I\}.$$

Lemma A.10. *If σ and τ are stopping times with $\sigma \le \tau$, then $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.*

A.3 Martingales and local martingales

Definition A.11. Let $(X_t)_{t\in[0,T]}$ be a real-valued $(\mathcal{F}_t)$-adapted stochastic process with $\mathbb{E}[|X_t|] < \infty$ for all $t \in I$. X is called a

- **martingale** if $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$,
- **sub-martingale** if $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$,
- **super-martingale** if $\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s$,

for all $s, t \in I$ with $s \leq t$.

Intuitively, a martingale $(X_t)_{t\in[0,T]}$ models the total gain process of a fair game with possibly several rounds. In each round the (conditional) expectation that we make a gain/loss is 0:

$$\mathbb{E}[X_t - X_s|\mathcal{F}_s] = 0, \quad s, t \in I \text{ with } s \leq t.$$

Remark A.12. Every martingale is also a sub- and a super-martingale. For a martingale $(X_t)_{t\in I}$, the map $t \mapsto \mathbb{E}[X_t]$ is constant as

$$\mathbb{E}[X_t] = \mathbb{E}[\mathbb{E}[X_t|\mathcal{F}_0]] = \mathbb{E}[X_0], \quad t \in I.$$

Theorem A.13.

(i) Let $(X_t)_{t\in I}$ and $(Y_t)_{t\in I}$ be martingales, then $(aX_t + bY_t)_{t\in I}$ is a martingale, for $a, b \in \mathbb{R}$

(ii) $(X_t)_{t\in I}$ is a super-martingale if and only if $(-X_t)_{t\in I}$ is a sub-martingale.

(iii) If $(X_t)_{t\in I}$ and $(Y_t)_{t\in I}$ be super-martingales, then $(Z_t)_{t\in I} := (X_t \wedge Y_t)_{t\in I}$ is a super-martingale.

Proposition A.14. *Let $(X_t)_{t\in I}$ be a martingale and $\varphi\colon \mathbb{R} \to \mathbb{R}$ be a convex function. If $\mathbb{E}[|\varphi(X_t)|] < \infty$ for all $t \in I$, then $(\varphi(X_t))_{t\in I}$ is a sub-martingale.*

Theorem A.15 (Doob's optional sampling theorem)**.** *Let σ, τ be bounded stopping times with $\sigma \leq \tau$.*

(i) If (X_t) is a continuous martingale, then

$$\mathbb{E}[X_\tau|\mathcal{F}_\sigma] = X_\sigma \quad \textit{and thus} \quad \mathbb{E}[X_\tau] = X_0.$$

(ii) If (X_t) be a continuous sub-martingale (super-martingale), then

$$\mathbb{E}[X_\tau|\mathcal{F}_\sigma] \geq X_\sigma \quad (\mathbb{E}[X_\tau|\mathcal{F}_\sigma] \leq X_\sigma).$$

As an immediate consequence of Doob's optional sampling theorem (Theorem A.15), we obtain the following corollary.

Corollary A.16 (Doob's stopping theorem)**.** *Let τ be a bounded stopping time and $(X_t)_{t\in I}$ be a martingale. Then, one has*

$$\mathbb{E}[|X_\tau|] < \infty \quad \textit{and} \quad \mathbb{E}[X_\tau] = X_0.$$

Definition A.17. An $(\mathcal{F}_t)$-adapted process $(X_t)_{t\in[0,T]}$ is called **local martingale** if there is an increasing sequence $(\tau_n)_{n\in\mathbb{N}}$ of $(\mathcal{F}_t)$-stopping times with $\tau_n \uparrow T$ $\mathbb{P}$-a.s. and $(X_t^n)_{t\in[0,T]} := (X_{t\wedge\tau_n})_{t\in[0,T]}$ is an $(\mathcal{F}_t)$-martingale for every $n \in \mathbb{N}$. The sequence $(\tau_n)_{n\in\mathbb{N}}$ is called **localizing sequence** for $(X_t)_{t\in[0,T]}$. **Local sub-martingales** and **local super-martingales** are defined analogously.

Theorem A.18. *Let $X = (X_t)_{t\in[0,T]}$ be a continuous local martingale. Then there exists a unique continuous process $\langle X\rangle = (\langle X\rangle_t)_{t\in[0,T]}$ with the following properties*

(i) $\langle X\rangle_0 = 0$ and $\langle X\rangle$ is non-decreasing and

(ii) $(X_t^2 - \langle X\rangle_t)_{t\in[0,T]}$ is a continuous local martingale.

Furthermore, $(\langle X\rangle_t)_{t\in[0,T]}$ satisfies

$$\langle X\rangle_t = \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]}X)^2 \quad \text{in probability for any } t \in [0,T],$$

where the limit is taken along any zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$.

A.4 Brownian motion and Itô integration

The Brownian motion is one of the most important building blocks for asset pricing models in continuous time. Its underlying normal distribution appears as "universal" limit distribution as observed in the central limit theorem.

Definition A.19. A stochastic process $W = (W_t)_{t\in[0,T]}$ is called a (standard one-dimensional) **Brownian motion** if:

(i) $W_0 = 0$.

(ii) W has independent increments, i.e., for all $n \in \mathbb{N}$ and $t_0 < t_1 < \cdots < t_n \subseteq [0,T]$,

$$\left(W_{t_{i+1}} - W_{t_i}\right)_{i=0,\dots,n-1} \quad \text{are independent.}$$

(iii) The increments are stationary and normally distributed:

$$W_t - W_s \sim \mathcal{N}(0, |t-s|), \quad s,t \in [0,T].$$

(iv) W has almost surely continuous sample paths.

A stochastic process $(W_t^1, \dots, W_t^d)_{t\in[0,T]}$ is called **d-dimensional Brownian motion** if $W^1, \dots, W^d$ are mutually independent one-dimensional Brownian motion.

Remark A.20. A Brownian motion $(W_t)_{t\in[0,T]}$ exists on a suitable probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as first shown by Nobert Wiener in 1923. Therefore, the Brownian motion is often called **Wiener process**. It can be considered as a "functional version" of the normal distribution and can be constructed as a scaling limit of a normalized random walk, as stated by Donsker's theorem.

Proposition A.21. *Let $(W_t)_{t\in[0,T]}$ be a Brownian motion. The completed natural filtration $(\mathcal{F}_t)_{t\in[0,T]}$ of a Brownian motion $(W_t)_{t\in[0,T]}$, defined by*

$$\mathcal{F}_t := \sigma(\mathcal{F}_t^W, \mathcal{N}) = \sigma(\{W_s^{-1}(A) : A \in \mathcal{B}(\mathbb{R}),\, s \in [0,t]\}, \mathcal{N}), \quad \textit{for} \quad t \in [0,T],$$

is right-continuous, i.e. $\mathcal{F}_t = \mathcal{F}_{t+}$ for $t \in [0,T)$. Hence, $(\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions and is called Brownian standard filtration.

Let us recall some properties of the Brownian motion, which are proven in the course "Stochastic processes (WT1)". Let $(W_t)_{t\in[0,T]}$ be a Brownian motion.

- **Strong Markov property**: Let $(\mathcal{F}_t)_{t\in[0,T]}$ be the Brownian standard filtration of $(W_t)_{t\in[0,T]}$ and τ be a $(\mathcal{F}_t)$-stopping time. Then, $(W_{\tau+t} - W_\tau)_{t\in[0,T-\tau]}$ is a Brownian motion, which is independent of Brownian standard filtration $(\mathcal{F}_t)_{t\in[0,T]}$.

- **Martingale**: $(W_t)_{t\in[0,T]}$ is a martingale w.r.t. $(\mathcal{F}_t)_{t\in[0,T]}$, i.e. W_t is $\mathcal{F}_t$-measurable, $\mathbb{E}[|W_t|] < \infty$ for all t, and

$$\mathbb{E}[W_t|\mathcal{F}_s] = W_s, \quad s,t \in [0,T] \quad \text{with } s < t.$$

- **Quadratic variation:** $\langle W\rangle_t = t$ for $t \in [0,T]$.

As we have seen in the course "Stochastic Calculus", one can develop an stochastic integration theory with respect to an Brownian motion. The spaces of corresponding integrands is given in the next definition.

Definition A.22. For a fixed $T \in (0,\infty)$ we introduce the space

$$\mathscr{H}^2 := \Big\{f\colon \Omega\times[0,T]\to\mathbb{R} \,:\, f \text{ is measurable, adapted and } \mathbb{E}\Big[\int_0^T f(\cdot,s)^2\,\mathrm{d}s\Big] < \infty\Big\},$$
$$\mathscr{H}^2_{loc} := \Big\{f\colon \Omega\times[0,T]\to\mathbb{R} : f \text{ is measurable, adapted and } \int_0^T f^2(\cdot,s)\,\mathrm{d}s < \infty\,\mathbb{P}\text{-a.s.}\Big\}.$$

The next theorem provides the Itô's integration with respect to a Brownian motion and some of its properties.

Theorem A.23 (Itô's integration w.r.t. Brownian motion)**.**

- *For any $f \in \mathscr{H}^2$ **Itô integral process** $(\int_0^t f(\cdot,s)\,\mathrm{d}B_s)_{t\in[0,T]}$ is well-defined, has a continuous modification, is a martingale and satisfies **Itô isometry***

$$\mathbb{E}\Big[\int_0^t f^2(\cdot,s)\,\mathrm{d}\langle B\rangle_s\Big] = \mathbb{E}\Big[\Big(\int_0^t f(\cdot,s)\,\mathrm{d}B_s\Big)^2\Big], \quad t\in[0,T].$$

- *For $f \in \mathscr{H}^2_{loc}$ **Itô integral process** $(\int_0^t f(\cdot,s)\,\mathrm{d}B_s)_{t\in[0,T]}$ is well-defined, has a continuous modification and is a local martingale.*

A.5 Stochastic integration for Itô processes

The stochastic Itô's integration with respect to a Brownian motion can, in a fairly easy way, be extended to the larger class of so-called Itô process.

Definition A.24. A stochastic process $X = (X_t)_{t\in[0,T]}$ is called **Itô process** if there is a ($\mathbb{P}$-a.s.) representation

$$X_t = X_0 + \int_0^t a(\cdot, s)\,\mathrm{d}s + \int_0^t b(\cdot, s)\,\mathrm{d}B_s, \qquad \text{for } t \in [0,T],$$

where $X_0 \in \mathbb{R}$ and $a, b\colon \Omega \times [0,T] \to \mathbb{R}$ are adapted, measurable processes satisfying the integrability conditions

$$\mathbb{P}\Big(\int_0^T |a(\omega, s)|\,\mathrm{d}s < \infty\Big) = 1 \qquad \text{and} \qquad \mathbb{P}\Big(\int_0^T |b(\omega, s)|^2\,\mathrm{d}s < \infty\Big) = 1.$$

A crucial tool to study

Proposition A.25. *For any Itô process $X = (X_t)_{t\in[0,T]}$ with representation*

$$X_t = X_0 + \int_0^t a(\cdot, s)\,\mathrm{d}s + \int_0^t b(\cdot, s)\,\mathrm{d}B_s, \quad t \in [0,T],$$

*the **quadratic variation** of X is given by*

$$\langle X\rangle_t := \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]}X)^2 = \int_0^t b^2(\cdot, s)\,\mathrm{d}s, \quad t \in [0,T].$$

Definition A.26. Let $X = (X_t)_{t\in[0,T]}$ be an Itô process with representation $X_t = X_0 + \int_0^t a(\cdot, s)\,\mathrm{d}s + \int_0^t b(\cdot, s)\,\mathrm{d}B_s$ for $t \in [0,T]$. We write $\mathcal{L}(X)$ for all adapted, measurable functions $f\colon \Omega \times [0,T] \to \mathbb{R}$ satisfying

$$\int_0^T |f(\cdot, s)a(\cdot, s)|\,\mathrm{d}s < \infty \qquad \text{and} \qquad \int_0^T |f(\cdot, s)b(\cdot, s)|^2\,\mathrm{d}s < \infty \qquad \mathbb{P}\text{-a.s.}$$

For $f \in \mathcal{L}(X)$ we define the **stochastic Itô integral** by

$$\int_0^t f(\cdot, s)\,\mathrm{d}X_s := \int_0^t f(\cdot, s)a(\cdot, s)\,\mathrm{d}s + \int_0^t f(\cdot, s)b(\cdot, s)\,\mathrm{d}B_s, \quad t \in [0,T].$$

Note that the stochastic Itô integral is well-defined as the $f(\cdot, s)b(\cdot, s) \in \mathscr{H}^2_{loc}$ and the representation of the Itô process X is unique.

Lemma A.27. *Let $X = (X_t)_{t\in[0,T]}$ be an Itô process with representation $X_t = X_0 + \int_0^t b(\cdot, s)\,\mathrm{d}B_s$ for $t \in [0,T]$, that is $a = 0$, and $f \in \mathcal{L}(X)$.*

(i) The integral process $(\int_0^t f(\cdot, s)\,\mathrm{d}X_s)_{t\in[0,T]}$ is a continuous local martingale.

(ii) If $\mathbb{E}[\int_0^T f^2(\cdot, s)b^2(\cdot, s)\,\mathrm{d}s] < \infty$, then Itô isometry holds true:

$$\mathbb{E}\Big[\Big(\int_0^t f(\cdot, s)\,\mathrm{d}X_s\Big)^2\Big] = \mathbb{E}\Big[\int_0^t f^2(\cdot, s)\,\mathrm{d}\langle X\rangle_s\Big], \quad t \in [0,T].$$

While fundamental theorem of calculus does not apply to stochastic Itô integration, its role can be recovered by Itô's formula.

Theorem A.28 (Itô's formula for Itô processes)**.** *Let $f \in C^2(\mathbb{R}^2)$ and X, Y be two Itô processes with representations*

$$X_t = X_0 + \int_0^t a(\cdot, s)\,\mathrm{d}s + \int_0^t b(\cdot, s)\,\mathrm{d}B_s \quad \text{and} \quad Y_t = Y_0 + \int_0^t \alpha(\cdot, s)\,\mathrm{d}s + \int_0^t \beta(\cdot, s)\,\mathrm{d}B_s,$$

for $t \in [0,T]$, $\mathbb{P}$-a.s. Then, we have $\mathbb{P}$-a.s for $t \in [0,T]$ that

$$\begin{aligned} f(X_t, Y_t) = f(X_0, Y_0) &+ \int_0^t f_x(X_s, Y_s)\,\mathrm{d}X_s + \int_0^t f_y(X_s, Y_s)\,\mathrm{d}Y_s \\ &+ \frac{1}{2}\int_0^t f_{xx}(X_s, Y_s)\,\mathrm{d}\langle X\rangle_s + \frac{1}{2}\int_0^t f_{yy}(X_s, Y_s)\,\mathrm{d}\langle Y\rangle_s \\ &+ \int_0^t f_{xy}(X_s, Y_s)\,\mathrm{d}\langle X, Y\rangle_s, \end{aligned}$$

where

$$\langle X, Y\rangle_t := \frac{1}{4}(\langle X+Y\rangle_t - \langle X-Y\rangle_t), \qquad t \in [0,T],$$

is the ***co-variation process****.*

A.6 Martingale representation and Girsanov's theorem

This subsection of the appendix collects two fundamental results: the martingale representation theorem and Girsanov's theorem.

Theorem A.29 (Martingale representation theorem)**.** *Let $(X_t)_{t\in[0,T]}$ be an $(\mathcal{F}_t)$-martingale with $\mathbb{E}[X_T^2] < \infty$. Then, there exists some $\varphi \in \mathscr{H}^2$ such that*

$$X_t = \mathbb{E}[X_0] + \int_0^t \varphi(s)\,\mathrm{d}B_s, \quad t \in [0,T].$$

This representation is $\mathbb{P} \otimes \lambda$-a.s. unique.

The fundamental theorem regarding change of measures is the so-called *Girsanov's theorem*, which is here presented in case of Brownian motion.

Theorem A.30 (Girsanov's theorem)**.** *Let $X \in \mathcal{L}(B)$ for a Brownian motion $(B_t)_{t\in[0,T]}$. If $L = (L_t)_{t\in[0,T]}$ where*

$$L_t = \exp\left(\int_0^t X_s\,\mathrm{d}B_s - \frac{1}{2}\int_0^t X_s^2\,\mathrm{d}s\right), \quad t \in [0,T],$$

is a martingale on $(\Omega, \mathcal{F}_T, \mathbb{P})$ with respect to the Brownian standard filtration $(\mathcal{F}_t)_{t\in[0,T]}$, then

$$\widetilde{B}_t := B_t - \int_0^t X_s\,\mathrm{d}s, \quad t \in [0,T],$$

defines a Brownian motion $\widetilde{B} = (\widetilde{B}_t)_{t\in[0,T]}$ with respect to $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, Q)$ where Q is defined by the Radon-Nikodym density $\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} := L_T$.

A.7 Stochastic differential equations

A stochastic differential equation (SDE) is a differential equation of the form

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t, \quad t \in [0, T], \qquad X_0 = x_0, \tag{A.1}$$

which is a short writing for the integral equation

$$X_t = x_0 + \int_0^t \mu(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s, \quad t \in [0, T].$$

where $\mu\colon [0, T] \times \mathbb{R} \to \mathbb{R}$ is the so-called **drift function** and $\sigma\colon [0, T] \times \mathbb{R} \to \mathbb{R}$ is the so-called **volatility** or **diffusion function** and $x_0 \in \mathbb{R}$ is the **initial value**. A stochastic process $X = (X_t)_{t\in[0,T]}$ is called **solution** of the SDE (A.1) if $X \in \mathcal{H}^2_{loc}$ is a continuous process satisfying (A.1).

Theorem A.31 (Existence and uniqueness of solution of SDEs)**.** *Let $\mu, \sigma\colon [0, T] \times \mathbb{R} \to \mathbb{R}$ be two measurable functions satisfying for some constant $C > 0$*

$$\begin{aligned} |\mu(t, x) - \mu(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 &\leqslant C|x - y|^2, \qquad t \in [0, T],\ x, y \in \mathbb{R}, \\ |\mu(t, x)|^2 + |\sigma(t, x)|^2 &\leqslant C(1 + |x|^2), \qquad t \in [0, T],\ x \in \mathbb{R}, \end{aligned}$$

and let $x_0 \in \mathbb{R}$. Then, we have:

(i) There exists a solution $(X_t)_{t\in[0,T]}$ to stochastic differential equation

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t, \quad t \in [0, T], \quad X_0 = x_0. \tag{A.2}$$

(ii) Every solution $(X_t)_{t\in[0,T]}$ of (A.2) is uniformly bounded in L^2, i.e.,

$$\sup_{t\in[0,T]} \mathbb{E}\big[|X_t|^2\big] < \infty.$$

(iii) The solution $(X_t)_{t\in[0,T]}$ of (A.2) is pathwise unique, i.e., if there is another solution $(Y_t)_{t\in[0,T]}$ of (A.2), we have $\mathbb{P}(\forall t \in [0, T] : X_t = Y_t) = 1$.

B Dictionary and abbreviations

B.1 Dictionary English-German

English	**German**
absolutely continuous	absolut stetig
adapted	adaptierte
almost sure convergence	fast sichere Konvergenz
bounded	beschränkt
contingent claim	Zahlungsanspruch
continuous	stetig
countable	abzählbar
conditional expectation	bedingte Erwartungswert
density	Dichte
density process	Dichteprozess
derivative	Ableitung
differentiable	differenzierbar
difference quotient	Differenzenquotient
dominated convergence theorem	Satz von der majorisierten Konvergenz
expectation	Erwartungswert
equivalent	äquivalent
Fatou's lemma	Lemma von Fatou
filtered probability space	gefilterter Wahrscheinlichkeitsraum
filtration	Filtration, Filtrierung
(financial) derivative	Derivat
identically distributed	identisch verteilt
independent	unabhängig
inequality	Ungleichung
integers	ganze Zahlen
integrable	integrierbar
integration	Integration
integral	Integral
interior	Inneres
intermediate value theorem	Zwischenwertsatz
Markov chain	Markow-Kette
martingale	Martingal
martingale measure	Martingalmaß
maturity	Fälligkeit
measure	Maß
measurable space	Messraum, messbarer Raum
monotone convergence theorem	Satz von der monotonen Konvergenz
natural numbers	natürliche Zahlen
$\mathbb{P}$-almost surely	$\mathbb{P}$-fast sicher
power set	Potenzmenge
predictable	vorhersehbar
probability measure	Wahrscheinlichkeitsmaß
probability space	Wahrscheinlichkeitsraum

English	German
quadratic variation	quadratische Variation
Radon-Nikodym density	Radon-Nikodym-Dichte
random variable	Zufallsgröße, Zufallsvariable
rational numbers	rationale Zahlen
real numbers	reelle Zahlen
replicable	replizierbar
risk-neutral probability measure	Martingalmaß
Riesz representation theorem	Darstellungssatz von Riesz
σ-algebra	σ-Algebra
sample path	Pfade
sample space	Ergebnisraum
semimartingale	Semimartingal
separation theorem	Trennungssatz
set	Menge
stochastic process	stochastischer Prozess
strike price	Ausübungspreis, Basispreis
trading strategy	Handelsstrategie
trajectory	Trajektorie
stopping time	Stoppzeit
term structure	Terminstruktur
triangle inequality	Dreiecksungleichung
tower property	Turmeigenschaft
uniformly integrable	gleichgradige integrierbar
variance	Varianz
volatility	Volatilität, Schwankungsanfälligkeit

B.2 English abbreviations

Abbreviation	**Meaning**
ad.	adapted
a.e.	almost everywhere
a.s.	almost surely
bdd	bounded
BM	Brownian motion
cts	continuous
EMM	equivalent martingale measure
eq.	equation
DPP	dynamic programming principle
fct.	function
HJB eq.	Hamilton-Jacobi-Bellman equation
iid	independent and identically distributed
iff	if and only if
loc.	local
mart.	martingale
MM	martingale measure
mb.	measurable
$\mathbb{P}$-a.s.	$\mathbb{P}$-almost surely
pred.	predictable
prob.	probability
RN density	Radon-Nikodym density
r.v.	random variable
SDE	stochastic differential equation
s.t.	such that
stoch.	stochastic
trad.	trading
u.i.	uniformly integrable
w.l.o.g.	without lost of generality
w.r.t.	with respect to

References

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathmatischen Annalen*, 300:463–520.

Gatheral, J. (2006). *The Volatility Surface: A Practitioner's Guide.* John Wiley and Sons, Ltd.

Jarrow, R. A. ([2021] ©2021). *Continuous-time asset pricing theory — a martingale-based approach.* Springer Finance. Springer, Cham.

Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion.* Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg.

Shreve, S. E. (2004). *Stochastic calculus for finance. II.* Springer Finance. Springer-Verlag, New York. Continuous-time models.