Lecture Notes Continuous-Time Finance

Prof. Rüdiger Frey,
ruediger.frey@wu.ac.at

Version from January 11, 2021, Comments welcome

Contents

Introduction

The goal of these notes is to give the reader a formal yet accessible introduction to continuous time financial mathematics. Continuous-time models are admittedly more complicated than their discrete-time counterparts. Nonetheless there are a number of good reasons to deal with them: To begin with, on many markets with very frequent trading the assumption of continuous security trading is closer to reality than assuming that markets are open only at fixed time points such as once a day. Moreover, in continuous-time models we can often get closed form solutions for derivatives prices which are not available in discrete models. Finally continuous-time modelling is the 'state-of-the art' in the modern literature.

The presentation starts with a brief introduction to discrete-time models (Chapter 1). We explain the notion of dynamic hedging and introduce the concept of an equivalent martingale-measure. Moreover, we discuss the fundamental theorems of asset pricing and derive the risk-neutral pricing principle. To illustrate these concepts we briefly discuss the binomial model of Cox, Ross & Rubinstein (1979). The core part of these notes is dedicated to models in continuous time. In Chapter 2 we give some basic facts about stochastic processes and introduce Brownian motion. We discuss sample paths properties and in particular the quadratic variation of Brownian motion. Chapter 3 is devoted to parts of the 'pathwise Itô-calculus' of Föllmer (1981). This approach enables us to derive all the mathematical tools necessary for an analysis of the Black-Scholes model in a rigorous but simple way. In Chapter 4 we present a first analysis of the Black-Scholes model via partial differential equations (PDEs), followed by a brief digression into portfolio optimization via stochastic control methods and the HJB equation. Chapter 5 provides further tools from stochastic calculus, most notably a discussion of the Girsanov theorem. In Chapter 6 these tools are applied to financial issues: we analyze basic principles of derivative pricing in continuous time, discuss the Black-Scholes model from a probabilistic perspective and study generalized Black Scholes models with more than one asset. We give a brief introduction to portfolio optimization and dynamic programming in 7. The text closes with a discussion of interest rate models and gives applications to interest-rate and and currency derivatives (in Chapter 8). Finally, a short appendix contains some background material on conditional expectations and discrete-time martingales.

There are many excellent textbooks on pricing and hedging of derivatives on various levels available. Good elementary texts are Cox & Rubinstein (1985) or Jarrow & Turnbull (1996); Hull (1997) is particularly popular with practitioners. Slightly more advanced texts which give also an introduction to stochastic calculus include Lamberton & Lapeyre (1996), Shreve (2004), Björk (2004) and Bingham & Kiesel (1998). In preparing these notes we relied a lot on the last two texts. Advanced texts on mathematical finance are Musiela & Rutkowski (1997) and Karatzas & Shreve (1998); Cont & Tankov (2004) gives an excellent introduction to financial modelling with jump processes. The necessary tools from probability theory can be found in Williams (1991) or in Jacod & Protter (2004). Good introductions to stochastic calculus in general are (in increasing order of technicality) Oksendal (1998), Karatzas & Shreve (1988), Protter (2005) and Revuz & Yor (1994).

These lecture notes grew out of various lecture courses taught by the author at the Vienna University of Economics and Business, the University of Leipzig and the University of Zürich; the audience consisted of master or PhD students in financial mathematics or in quantitative finance. At this point a warning is in order. This text is **not** a published textbook. Hence some sections are more polished than others, there are (slight) inconsistencies in the notation between chapters and there is 'almost surely' a number of errors and typos in the text. Of course I intend to improve the text over time, and I am grateful for any error which is being pointed out to me (`ruediger.frey@wu.ac.at`).

Vienna, November 2014 Rüdiger Frey

Chapter 1

Discrete-Time Models: a Wrap-Up

In this section we give a brief introduction to the pricing and hedging of derivatives in finite market models, i.e. models with a finite number of trading dates in which all asset prices take on only finitely many different values. In this simple setting we can work out the key financial and mathematical ideas underlying modern derivative asset analysis without having to deal with the technicalities of stochastic calculus. We basically follow the approach of Harrison & Pliska (1981).

1.1 Basic notions

We work on a probability space $(\Omega, \mathcal{F}, P)$ with finite state space $\Omega = \{\omega_1, \ldots, \omega_s\}$. We consider a finite number of trading dates $t = 0, 1, \ldots N$ where $t = N$ often corresponds to the maturity of the derivative contract under consideration. As usual we use a filtration $\{\mathcal{F}_n\}$, $n = 0, 1, \ldots, N$ to model the information-flow over time: an event A belongs to $\mathcal{F}_n$ if the agents in our model can decide from the information available to them at $t = n$ if the event A has occurred or not.

Assets: There are two assets in our model, a riskless money market account with price process S^0 and a risky security S^1 (the stock). None of these assets is paying dividends between $t = 0$ and $t = N$. We work with a deterministic interest rate r per period such that $S_n^0 = (1+r)^n$. The discount factor is given by $D_n = (1+r)^{-n}$. The discounted stock price process is given by $\tilde{S}_n^1 := D_n S_n^1$; the discounted price of the money market account obviously equals $\tilde{S}_n^0 \equiv 1$. We assume that the stock price process is adapted to $\{\mathcal{F}_n\}$. In the sequel we refer to the filtered probability space $(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_n\}$, the set of trading dates and the price processes of S^0 and S^1 together as our security market model $\mathcal{M}$.

Trading strategies: The investors in our model are allowed to form dynamic portfolios in stock and money market account. Formally a trading strategy (or dynamic portfolio strategy) is a stochastic process (a sequence of random variables) $\phi = (\phi_n^0, \phi_n^1)_{n=1,\ldots,N}$ with the following economic interpretation: ϕ_n^0 respectively ϕ_n^1 represent the number of units of the money market account respectively the number of shares of the stock the investor selects for his portfolio at $t = n-1$ and holds up to and including time $t = n$. To capture economic reality a trading strategy should be non-anticipating, i.e. in deciding about ϕ_n at time $t = n-1$ the investor has only the information contained in $\mathcal{F}_{n-1}$ – such as the stock price S_{n-1}^1 – at her disposal and *not* the information contained in $\mathcal{F}_n$. This is formalized in the following definition.

Definition 1.1. Given a security market model $\mathcal{M}$.

(i) A trading strategy $\phi = (\phi_n)_{n=1,\dots,N}$ is called admissible if ϕ_n^0 and ϕ_n^1 are $\mathcal{F}_{n-1}$-measurable for $n = 1, \dots N$, i.e. if ϕ is a predictable process.

(ii) The *value* of the strategy ϕ at time $t = n$ equals $V_n = V_n(\phi) = \phi_n^0 S_n^0 + \phi_n^1 S_n^1$; the discounted value is given by $\tilde{V}_n = D_n V_n = \phi_n^0 + \phi_n^1 \tilde{S}_n^1$.

(iii) An admissible strategy is called *selffinancing* if for all $n = 1, \dots, N$

$$V_n = \phi_n^0 S_n^0 + \phi_n^1 S_n^1 = \phi_{n+1}^0 S_n^0 + \phi_{n+1}^1 S_n^1\,, \tag{1.1}$$

i.e. if no funds are withdrawn from or injected into the strategy.

The following characterization of selffinancing strategies will be very convenient in the future.

Lemma 1.2. *An admissible strategy ϕ is selffinancing if and only if we have for all $n = 1, ..., N$*

$$V_n(\phi) = V_0(\phi) + \sum_{j=1}^{n} \phi_j^0 (S_j^0 - S_{j-1}^0) + \sum_{j=1}^{n} \phi_j^1 (S_j^1 - S_{j-1}^1)\,. \tag{1.2}$$

The value of a selffinancing strategy hence consists of the initial investment V_0 and the gains (or losses) from trade in stock and money market account.

Proof. We get by definition of the value of a portfolio that

$$V_{n+1}(\phi) - V_n(\phi) = \phi_{n+1}^0 S_{n+1}^0 + \phi_{n+1}^1 S_{n+1}^1 - \phi_n^0 S_n^0 - \phi_n^1 S_n^1\,. \tag{1.3}$$

Now ϕ is selffinancing if and only if $\phi_n^0 S_n^0 + \phi_n^1 S_n^1 = \phi_{n+1}^0 S_n^0 + \phi_{n+1}^1 S_n^1$ for all $n = 1, \dots, N$. Plugging this into (1.3) yields

$$V_{n+1}(\phi) - V_n(\phi) = \phi_{n+1}^0 (S_{n+1}^0 - S_n^0) + \phi_{n+1}^1 (S_{n+1}^1 - S_n^1)\,. \tag{1.4}$$

As $V_{n+1}(\phi) = V_0(\phi) + \sum_{i=0}^{n} \big(V_{i+1}(\phi) - V_i(\phi)\big)$ the lemma follows by summing over (1.4). □

We can give a similar characterization of selffinancing strategies in terms of discounted quantities.

Lemma 1.3. *An admissible strategy is selffinancing if and only if we have for all $n = 1, ..., N$*

$$\tilde{V}_n(\phi) = \tilde{V}_0(\phi) + \sum_{j=1}^{n} \phi_j^1 (\tilde{S}_j^1 - \tilde{S}_{j-1}^1)\,. \tag{1.5}$$

Being similar to the proof of Lemma 1.2 the proof will be omitted.

1.2 No-arbitrage and Equivalent Martingale Measures

Roughly speaking an arbitrage opportunity is a trading strategy which allows us to create strictly positive profits without risk i.e. with zero initial investment.

Definition 1.4. (i) A self-financing, admissible strategy ϕ with $V_0(\phi) = 0$ is called an arbitrage opportunity if $V_N(\phi) \geq 0$ and $P(V_N(\phi) > 0) > 0$.

(ii) A security market model $\mathcal{M}$ is arbitrage-free, if there are no arbitrage opportunities.

Remark 1.5. 1) Of course an admissible strategy ϕ such that $V_0(\phi) < 0$ and $V_N(\phi) \geq 0$ also constitutes an arbitrage opportunity as such a strategy allows an investor to consume the positive amount $U_0 = (-V_0(\phi))$ in $t = 0$ without any further obligations. However, it is always possible to turn ϕ into an arbitrage opportunity in the sense of Definition 1.4 by investing U_0 into the riskless asset.

2) There are two different reasons for requiring that a good security market model should be arbitrage-free. To begin with, on real markets arbitrage opportunities do usually not prevail for long as the attempts of rational investors to exploit arbitrage opportunities makes them disappear.[1] More importantly, even if one believes that arbitrage opportunities do exist on real markets, there are still good reasons to insist that a security market model should be arbitrage-free. Otherwise an investor who uses this model for the pricing of derivatives will quote prices for these products which are inconsistent and risks therefore to fall victim to arbitrage-trades himself.

In order to characterize arbitrage-free markets, we use the concept of equivalent martingale measures.

Definition 1.6. Given a security market model $\mathcal{M}$. A probability measure Q on $(\Omega, \mathcal{F})$ such that

(i) Q is equivalent to P, i.e. for all $A \in \mathcal{F}$ we have $Q(A) > 0 \Leftrightarrow P(A) > 0$.

(ii) The discounted stock-price $\tilde{S}$ is a martingale.

is called an equivalent martingale measure or a risk-neutral measure for $\mathcal{M}$.

In a discrete setting condition (i) simply means that $Q(\omega) > 0$ for all ω, i.e.under both measures the same states of the world occur with positive probability. Condition (ii) is equivalent to the requirement that $E\left(\frac{1}{1+r}S_n|\mathcal{F}_{n-1}\right) = S_{n-1}$ for all $n = 1, \ldots N$. The name risk-neutral measure stems from the fact that the existence of a risk-neutral investor whose subjective probability distribution over future stock-prices is given by Q is consistent with our security market model.

Next we want to show that the existence of an equivalent martingale measure excludes arbitrage possibilities. For this we need:

Lemma 1.7. *Let Q be an equivalent martingale-measure for the market $\mathcal{M}$. Consider a selffinancing, admissible trading-strategy ϕ. Then the discounted value process $\tilde{V}_n(\phi)$ is a Q-martingale.*

[1] This does not imply that real markets are always arbitrage-free as institutional constraints and transaction costs can make it difficult to profit from arbitrage opportunities; see for instance Liu & Longstaff (2000) for a discussion.

Proof. As ϕ is selffinancing we get from Lemma 1.3

$$\tilde{V}_{n+1}(\phi) = \tilde{V}_0(\phi) + \sum_{j=1}^{n+1} \phi_j^1 (\tilde{S}_j^1 - \tilde{S}_{j-1}^1) = \tilde{V}_n(\phi) + \phi_{n+1}^1 (\tilde{S}_{n+1}^1 - \tilde{S}_n^1).$$

As ϕ is admissible, ϕ_{n+1} is $\mathcal{F}_n$-measurable. Hence, as $\tilde{S}^1$ is a Q-martingale,

$$E^Q(\tilde{V}_{n+1}(\phi) - \tilde{V}_n(\phi)|\mathcal{F}_n) = E^Q(\phi_{n+1}^1(\tilde{S}_{n+1}^1 - \tilde{S}_n^1)|\mathcal{F}_n) = \phi_{n+1}^1 E^Q(\tilde{S}_{n+1}^1 - \tilde{S}_n^1|\mathcal{F}_n) = 0\,.$$

□

Proposition 1.8. *If an equivalent martingale-measure exists for the security market model $\mathcal{M}$, the model $\mathcal{M}$ is arbitrage-free.*

Proof. Consider a self-financing strategy ϕ with $V_N(\phi) \geq 0$, $P(V_N(\phi) > 0) > 0$. We will show that the existence of an equivalent martingale-measure Q implies $V_0(\phi) > 0$; this shows that no arbitrage opportunities exist.

As $V_N(\phi)$ and $\tilde{V}_N(\phi)$ have the same sign it follows that $\tilde{V}_N(\phi) > 0$ and $P(\tilde{V}_N(\phi) > 0) > 0$. The equivalence of P and Q now implies that $Q(\tilde{V}_N(\phi) > 0) > 0$ and hence $E^Q(\tilde{V}_N(\phi)) > 0$. On the other hand, $(\tilde{V}(\phi))_{n=1,\dots,N}$ being a Q-martingale implies that $\tilde{V}_0(\phi) = E^Q\left(\tilde{V}_N(\phi)\right) > 0$ and hence also $V_0(\phi) > 0$. □

Proposition 1.9. *If the market is arbitrage-free, the class of equivalent martingale-measures is non-empty.*

The proof is based on the separating hyperplane theorem; see e.g. Bingham & Kiesel (1998), Proposition 4.2.3. Summing up, we have the so called first fundamental theorem of asset pricing.

Theorem 1.10. *A security market $\mathcal{M}$ is arbitrage-free if and only if there is a probability measure Q equivalent to P such that discounted asset price processes are Q-martingales.*

Remark: In this very strict form the first fundamental theorem of asset pricing holds only in a discrete-time setting; for a version of this theorem which is valid in more general conditions with continuous trading see Chapter 6.1 of Bingham & Kiesel (1998) and in particular the paper Delbaen & Schachermayer (1994).

1.3 Pricing and hedging of contingent claims

We now turn our attention to the pricing of contingent claims. Formally a contingent claim H with maturity T is an $\mathcal{F}_T$-measurable random variable H; $H(\omega)$ is interpreted as payoff of the claim in state ω. A contingent claim is called a derivative if its payoff depends only on the prices of traded securities; derivatives are obviously the most important class of contingent claims. Contingent claims which are not derivatives are traded in the insurance industry. For instance the payoff of so-called CAT-bonds depends essentially on the value of some aggregated claims index, which is typically not a traded security; for more on these claims see Canter, Cole & Sandor (1996).

The key idea underlying modern approaches to pricing contingent claims is the notion of dynamic replication.

Definition 1.11. Given a security market model $\mathcal{M}$.

(i) A contingent claim H with maturity $T \in \{1, \dots, N\}$ is called attainable, if there is an admissible, selffinancing strategy $\phi_n = (\phi_n^0, \phi_n^1)$ such that $V_T(\phi) = H$; ϕ is called replicating strategy for the derivative.

(ii) A market is called complete, if every contingent claim is attainable.

Definition 1.12. Consider an attainable claim H with replicating strategy ϕ in an arbitrage-free market model $\mathcal{M}$. The fair price of this claim at time $n \le T$ is $V_n(\phi)$.

This definition is motivated by the observation that by investing $V_n(\phi)$ at time $t = n$ and following the strategy the claim can be replicated without any further risk; a price higher (lower) than $V_n(\phi)$ would therefore constitute a riskless profit opportunity for the seller (buyer) of the claim.

The following theorem yields an alternative way to compute the fair price of an attainable claim using the risk-neutral measure.

Theorem 1.13. *Given an arbitrage-free market $\mathcal{M}$ and an attainable contingent claim H with replicating strategy ϕ. Let Q be an equivalent martingale measure for $\mathcal{M}$. Then the fair price of the claim H at time $n \le T$ is given by*

$$V_n(\phi) = E^Q((1+r)^{-(T-n)} H | \mathcal{F}_n); \text{ in particular } V_0(\phi) = E^Q((1+r)^{-T} H) . \tag{1.6}$$

Proof. As the strategy ϕ duplicates the claim, we have $V_T(\phi) = H$ and hence $(1+r)^{-T} H = \tilde{V}_T(\phi)$. As $(\tilde{V}_n(\phi))_{n=0,\dots,T}$ is a Q-martingale (by Lemma 1.7), we have

$$E^Q \left((1+r)^{-T} H | \mathcal{F}_n \right) = E^Q(\tilde{V}_T(\phi) | \mathcal{F}_n) = \tilde{V}_n(\phi) = (1+r)^{-n} V_n(\phi).$$

Hence $V_n(\phi) = E^Q((1+r)^{-(T-n)} H | \mathcal{F}_n)$. □

Relation (1.6) is often referred to as risk-neutral pricing rule. Theorem 1.13 shows in particular that in an arbitrage-free market two different admissible replicating strategies for a claim have the same value such that the definition of the fair price of a claim (Definition 1.12) is logically consistent.

While the *existence* of a risk-neutral measure is related to absence of arbitrage, *uniqueness* of a risk-neutral measure is related to market completeness. This is the content of the so-called second fundamental theorem of asset pricing.

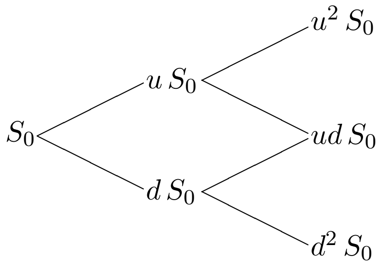
Theorem 1.14. *An arbitrage-free market $\mathcal{M}$ is complete if and only if there exists a* unique *equivalent martingale measure Q.*

For a proof we refer to Section 4.3 of Bingham & Kiesel (1998); generalizations to models with continuous security trading can be found in Harrison & Pliska (1981).

1.4 The binomial Cox-Ross-Rubinstein (CRR)-model

As an example we now present the binomial model of Cox et al. (1979). This simple model is still popular with practitioners as it yields an approximation to the Black-Scholes model under suitable rescaling of the model-parameters, which makes the CRR-model useful as a tool for computing (approximate) prices of derivatives.

We consider first a simple two-period example. Fix two numbers u and d with $u > 1+r > d$ which model the return of the stock in the up-state and in the down-state and an initial stock-price-level S_0. In a two-period CRR-model the stock-price then evolves as depicted in Figure **??**.



Note that the tree for the evolution of the stock-price is recombining, i.e. we obtain the same value for the stock-price at time $t = 2$ independent of the order in which up- and down movements occur. This property of the model facilitates its numerical implementation.

We now give a formal description of the N-period model. As state space Ω we take the set $\{u, d\}^N$ such that the elements of Ω are N-tupels with entries $\omega_i \in \{u, d\}$, $i = 1, \ldots, N$. Define for $1 \le n \le N$ $j_n(\omega) := \#\{i \le n\,;\, \omega_i = u\}$, such that $j_n(\omega)$ gives the number of up-movements in ω until $t = n$. We define the stock-price process S^1 by

$$S^1_n(\omega) = S_0 u^{j_n(\omega)} d^{(n-j_n(\omega))}\,, \quad 0 \le n \le N\,. \tag{1.7}$$

As filtration $\{\mathcal{F}_n\}$ we take the filtration generated by the stock price process, i.e. we put $\mathcal{F}_n = \sigma(S^1_i,\, i \le n)$. The probability measure P is left unspecified, we only require that $P(\omega) > 0$ for all $\omega \in \Omega$.

Equivalent martingale measure: We start with the case $N = 1$. Here the equivalent martingale measure Q must satisfy $E^Q((1+r)^{-1}S^1_1) = S_0$. If we define $\pi := Q(\omega_1 = u)$ we obtain the following condition for π

$$\frac{1}{1+r}(\pi u S_0 + (1-\pi) d S_0) = S_0 \text{ and hence } \pi = \frac{(1+r)-d}{u-d}\,. \tag{1.8}$$

It is immediate that $\pi \in (0, 1)$ if and only if $u > 1 + r > d$; moreover, in that case π is uniquely determined. If $N > 1$ we use our results from the one-period case to define transition probabilities. We put

$$Q(\omega_{n+1} = u|\mathcal{F}_n) := \pi \text{ and } Q(\omega_{n+1} = d|\mathcal{F}_n) = 1-\pi\,. \tag{1.9}$$

The probability of any $\omega \in \Omega$ is hence given by $Q(\omega) = \pi^{j_N(\omega)}(1-\pi)^{N-j_N(\omega)}$. Relation (1.8) implies that the discounted stock price process is a martingale. The uniqueness of π in the one-period case implies that (1.9) is the only choice of transition probabilities which makes $\tilde{S}^1$ a martingale, such that Q is unique. Note that under the risk-neutral measure Q the projections ω_n on the components of ω form a sequence of two-valued iid random variables.

Replicating strategies and market completeness: Our binomial model is complete in the sense of Definition 1.11. This follows from the uniqueness of the equivalent martingale measure and Theorem 1.14. Alternatively, using the *backward-induction principle* we can give an explicit recursive construction for hedging strategies for arbitrary claims. For simplicity we explain the approach in the two-period model of Figure **??**; the extension to N periods is obvious. Consider some claim which matures in $t = 2$ and has payoff $H(\omega)$. At time $t = 2$ the value of this claim is equal to its payoff. At $t = 1$ we have to distinguish between the up-state ($\omega_1 = u$) and the down-state ($\omega_1 = d$).

In the up-state our replicating portfolio $(\phi_1^0(u), \phi_1^1(u))$ must satisfy the following system of equations.

$$\begin{aligned}\phi_1^0(u)(1+r)^2 + \phi_1^1(u)u^2 S_0 &= H(u,u)\\ \phi_1^0(u)(1+r)^2 + \phi_1^1(u)udS_0 &= H(u,d)\,.\end{aligned}$$

As $u > d$ this linear system of equations has a unique solution given by

$$\phi_1^1(u) = \frac{H(u,u) - H(u,d)}{uS_0(u-d)}\,, \quad \phi_1^0(u) = \frac{-dH(u,u) + uH(u,d)}{(u-d)(1+r)^2}\,. \tag{1.10}$$

The value of the hedge-portfolio equals

$$V_1(u) = \phi_1^0(u)(1+r) + \phi_1^1(u)uS_0 = \frac{1}{1+r}\left(\pi H(u,u) + (1-\pi)H(u,d)\right)\,,$$

which is in line with the risk-neutral pricing rule. In the down-state we can compute a hedging portfolio $(\phi_1^0(d), \phi_1^1(d))$ using a similar argument. The value of the portfolio at $t = 1$ is given by

$$V_1(d) = \frac{1}{1+r}\left(\pi H(d,u) + (1-\pi)H(d,d)\right)\,.$$

We now determine a hedging portfolio at time $t = 0$. To finance our hedge in $t = 1$ the value of our portfolio must be $V_1(u)$ if the up-state occurs and $V_1(d)$ in the down-state. Hence our hedge (ϕ_0^0, ϕ_0^1) must solve the equations

$$\phi_0^0(1+r) + \phi_0^1 uS_0 = V_1(u) \ \text{ and } \ \phi_0^0(1+r) + \phi_0^1 dS_0 = V_1(d)\,,$$

which determines uniquely ϕ_0^0, ϕ_0^1 and V_0.

As the CRR-model is complete we may use the risk-neutral pricing rule to price arbitrary contingent claims. In case of a European call option we obtain the following

Proposition 1.15. *In a binomial CRR model with up-state-return u, down-state return d and interest-rate r such that $u > 1 + r > d$ the arbitrage price C_n at $t = n$ of a European call with strike price K and maturity N equals*

$$C_n = \frac{1}{(1+r)^{N-n}} \sum_{j=0}^{N-n} \binom{N-n}{j} \pi^j (1-\pi)^{N-n-j} \left(S_n(\omega)u^j d^{N-n-j} - K\right)^+\,.$$

Proof. We get from the risk-neutral pricing rule (1.6) that

$$C_n = \frac{1}{(1+r)^{N-n}} \sum_{\omega\in\Omega} Q(\omega|\mathcal{F}_n)(S_N(\omega) - K)^+\,.$$

Now note that for ω with $j_N(\omega) - j_n(\omega) = j$ (exactly j up-movements between now and maturity) we obtain $Q(\omega|\mathcal{F}_n) = \pi^j(1-\pi)^{N-n-j}$. Moreover, for these paths we have $S_N(\omega) = S_n(\omega)u^j d^{N-n-j}$. Hence we obtain

$$C_n = \frac{1}{(1+r)^{N-n}} \sum_{j=0}^{N-n} \#\{\omega,\, j_N(\omega) - j_n(\omega) = j\}\pi^j(1-\pi)^{N-n-j}\left(S_n(\omega)u^j d^{N-n-j} - K\right)^+ .$$

Now $\#\{\omega,\, j_N(\omega) - j_n(\omega) = j\}$ is given by the binomial coefficient $\binom{N-n}{j}$, which yields the result. $\square$

Chapter 2

Stochastic Processes in Continuous Time

2.1 Stochastic Processes, Stopping Times and Martingales

2.1.1 Basic Notions

We work on a probability space $(\Omega, \mathcal{F}, P)$ with filtration $\{\mathcal{F}_t\}$. Recall that a filtration is a family of σ-fields $\{\mathcal{F}_t,\, t \geq 0\}$ such that $\mathcal{F}_t \subseteq \mathcal{F}_s$ for $s > t$. As usual $\mathcal{F}_t$ is interpreted as the set of events which are observable at time t such that the filtration represents information-flow over time.

A stochastic process $X = (X_t)_{t \geq 0}$ is a family of random variables on $(\Omega, \mathcal{F}, P)$. We introduce the following notions:

- The process X is called *adapted*, if for all $t > 0$ the random variable (rv) X_t is $\mathcal{F}_t$-measurable.
- The *marginal distribution* of the process at a given $t \geq 0$ is the distribution $\mu(t)$ of the rv X_t.
- Consider a finite set of time points $(t_1, ..., t_n)$ in $[0, \infty)$. Then $(X_{t_1}, ..., X_{t_n})$ is a random vector with values in $\mathbb{R}^n$ and distribution $\mu(t_1, ..., t_n)$, say. The class of all such distributions is called the set of finite-dimensional distributions of X. The finite-dimensional distributions satisfy a set of obvious consistency requirements, moreover, they determine the stochastic properties of a stochastic process; see for instance Chapter 2 of Karatzas & Shreve (1988).
- Fix some $\omega \in \Omega$. The mapping

  $$X_{\cdot}(\omega) \colon [0, \infty) \to \mathbb{R}, \quad t \to X_t(\omega)$$

  is called *trajectory* or *sample path* of X; a stochastic process can be viewed as random draw of a sample-path. We are only interested in processes whose sample paths have certain regularity properties. Of particular interest will be processes with continuous sample paths like Brownian motion or right continuous with left limits (RCLL) such as the Poisson process (see below).

Equality of stochastic processes. There are two notions of equality for stochastic processes.

Definition 2.1. Given two stochastic processes X and Y. Then X is called *modification of* Y if for all $t \geq 0$ we have

$$P\big(\{\omega \in \Omega \colon X_t(\omega) = Y_t(\omega)\}\big) = 1.$$

The processes are called *indistinguishable* if

$$P\big(\{\omega \in \Omega \colon X_t(\omega) = Y_t(\omega)\, \forall t > 0\}\big) = 1.$$

We obviously have that if X and Y are indistinguishable then X is a modification of Y. For the converse implication extra regular assumptions on the trajectories are needed.

Lemma 2.2. *Suppose that X and Y have right-continuous trajectories and that X is a modification of Y. Then X and Y are indistinguishable.*

Proof. Put $N_t := \{\omega \in \Omega \colon X_t(\omega) \neq Y_t(\omega)\}$ and let $N := \bigcup_{q \in \mathbb{Q} \cap [0,\infty)} N_q$. Since $\mathbb{Q}$ is countable and since X is a modification of Y we have $0 = P(N_q) = P(N)$. We want to show that for $\omega \in \Omega \backslash N$ we have $X_t(\omega) = Y_t(\omega)\ \forall t \geq 0$. This is clear for $t \in \mathbb{Q}$. For $t \in \mathbb{R} \backslash \mathbb{Q}$ there is a sequence $q_n \in \mathbb{Q}$ with $q_n \downarrow t$. By definition of N we have $X_{q_n}(\omega) = Y_{q_n}(\omega)$ for all $\omega \in \Omega \backslash N$. Since X and Y have right-continuous trajectories we moreover get

$$X_t(\omega) = \lim_{n \to \infty} X_{q_n}(\omega)) = \lim_{n \to \infty} Y_{q_n}(\omega) = Y_t(\omega),$$

which proves the claim. □

2.1.2 Classes of Processes

1. Martingales: An adapted stochastic process X with $E(|X_t|) < \infty$ for all $t > 0$ is

- a submartingale if $\forall t, s$ with $t > s$ we have $E(X_t|\mathcal{F}_s) \geq X_s$.
- a supermartingale if $\forall t, s$ with $t > s$ we have $E(X_t|\mathcal{F}_s) \leq X_s$.
- a martingale if X is both a sub- and a supermartingale, i.e. if $E(X_t|\mathcal{F}_s) = X_s$ for all t, s.

Important examples for martingales are the Brownian motion and the compensated Poisson process. Both processes will be introduced below.

2. Semimartingales In financial modelling we often encounter processes which are the sum of a completely unpredictable part – modelled by a martingale – and a systematic predictable component such as the long-term growth rate of an asset. If the systematic component of such a process satisfies certain regularity properties these processes are called semimartingales. A formal definition of semimartingales is given in Definition 3.12 below.

3. Markov-Processes: An adapted stochastic process X is called Markov process, if for all $t, s > 0$ and all bounded functions $f \colon \mathbb{R} \to \mathbb{R}$

$$E(f(X_{t+s}) \mid \mathcal{F}_t) = E(f(X_{t+s}) \mid \sigma(X_t)). \tag{2.1}$$

Here $\sigma(X_t)$ denotes the σ-field generated by the rv X_t, a notation which we will use throughout these notes. Intuitively speaking a process is Markovian if the conditional distribution of future values X_{t+s}, $s \geq 0$, of the process is completely determined by the present value X_t of the process; in particular given the value of X_t, past values X_u, $u < t$ of the process do not contain any additional information which is useful for predicting X_{t+s}.

Remark 2.3. A Markov process X is called a strong Markov-process, if (2.1) holds for all stopping-times τ and not only for deterministic times t.[1] All Markov-processes we will encounter are also strong Markov processes, but there are a few 'pathological' exceptions.

4. Diffusions: A diffusion is a strong Markov process with continuous trajectories such that for all (t, x) the limits:

$$\mu(t,x) = \lim_{h \to 0} \frac{1}{h} E(X_{t+h} - X_t | X_t = x) \quad \text{and} \tag{2.2}$$

$$\sigma^2(t,x) = \lim_{h \to 0} \frac{1}{h} E((X_{t+h} - X_t)^2 | X_t = x) \tag{2.3}$$

exist. Then $\mu(t,x)$ is called the drift, $\sigma^2(t,x)$ the diffusion coefficient. The name diffusion stems from applications in physics; the most important mathematical examples are solutions to stochastic differential equations.

5. Point processes and the Poisson process: Assume that certain relevant 'events' — for instance claims in an insurance context or defaults of counterparties in a financial context — occur at random points in time $\tau_0 < \tau_1 < \dots$. The corresponding point process N_t is then given by $N_t := \sup\{n, \tau_n \leq t\}$, i.e. N_t measures the number of events which have occurred up to time t.

The Poisson process is a special point process. To construct it we take a sequence Y_n of independent exponentially distributed random variables with $P(Y_n \leq x) = 1 - e^{-\lambda x}$ and define $\tau_n := \sum_{j=1}^n Y_j$, such that Y_n is the waiting time between event $n-1$ and event n. The process $N_t = \sup\{n \colon \tau_n \leq t\}$ is then a Poisson process with intensity λ. It has among others the following properties

- $P(N_t = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$, $k = 0, 1, \dots,\ t \geq 0$.
- $N_{t+u} - N_t$ is independent of N_s for $s \leq t$ and Poisson-distributed with parameter (λu).
- The compensated Poisson process $M_t := N_t - \lambda t$ is a martingale; in particular $E(N_t) = \lambda t$.

2.2 Stopping Times and Martingales

Throughout this section we work on a probability space $(\Omega, \mathcal{F}, P)$ with filtration $\{\mathcal{F}_t\}$.

2.2.1 Stopping Times

Definition 2.4. A rv $\tau \colon \Omega \to [0, \infty]$ is called *stopping time wrt.* $\{\mathcal{F}_t\}$ if for all $t \geq 0$ it holds that $\{\tau \leq t\} \in \mathcal{F}_t$.

[1] As in discrete time a random variable τ with values in $[0, \infty]$ will be called a stopping time if for all $t \geq 0$ the set $\{\omega\,, \tau(\omega) \leq t\}$ belongs to the sigma-field $\mathcal{F}_t$; see Section 2.2 below.

Remark 2.5. τ can be interpreted as the time of the occurrence of an observed event. $\{\tau = \infty\}$ means that the event never occurs.

Lemma 2.6. *Let $\{\mathcal{F}_t\}$ be right-continuous, i.e. $\mathcal{F}_t = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}$, for all $t \geq 0$. Then $\tau\colon \Omega \to [0,\infty]$ is a stopping time if and only if $\{\tau < t\} \in \mathcal{F}_t$, $\forall t \geq 0$.*

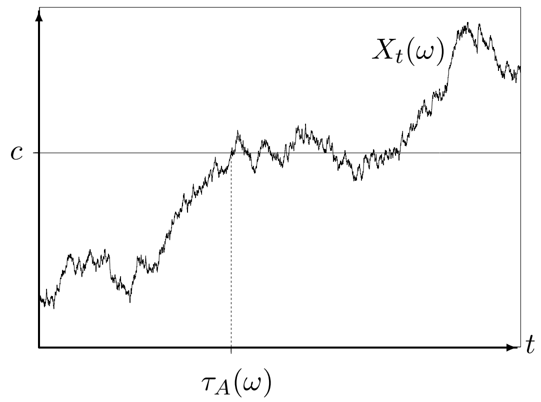
Proof. It holds $\{\tau \leq t\} = \bigcap_{\varepsilon>0}\{\tau < t+\varepsilon\}$. Let $\{\tau < t\} \in \mathcal{F}_t$, $\forall t \geq 0$. Hence $\{\tau < t+\varepsilon\} \in \mathcal{F}_{t+\varepsilon}$ and $\{\tau \leq t\} \in \bigcap\limits_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = F_t$. For the converse statement note that

$$\{\tau < t\} = \bigcup_{\varepsilon\in\mathbb{Q}_+} \underbrace{\{\tau \leq t-\varepsilon\}}_{\in \mathcal{F}_{t-\varepsilon}\subseteq\mathcal{F}_t} \in \mathcal{F}_t\,.$$

□

The most important example for stopping times are first hitting times for Borel sets.

Definition 2.7. Given a stochastic process X and a Borel set A in $\mathbb{R}$. Define $\tau_A := \inf\{t \geq 0\colon X_t \in A\}$. Then the rv τ_A is called *first hitting time* into the set A.



$A := \{x \in \mathbb{R}\colon x > c\}$

Next we address the question if τ_A is a stopping time.

Lemma 2.8. *Let X be $\{\mathcal{F}_t\}$-adapted and right-continuous and let $A \subseteq \mathbb{R}^d$ be open. If the filtration $\{\mathcal{F}_t\}$ is right-continuous, then the hitting time τ_A is a stopping time.*

Proof. Suppose, that $\{\mathcal{F}_t\}$ is right-continuous. We only have to show $\{\tau_A < t\} \in \mathcal{F}_t$ for all $t > 0$. Since X is right-continuous and A is open, it holds that

$$\{\tau_A < t\} = \bigcup_{q\in[0,t)\cap\mathbb{Q}} \{X_q \in A\}\,.$$

As $\{X_q \in A\} \in \mathcal{F}_q \subseteq \mathcal{F}_t$ and as the union of countable sets from $\mathcal{F}_t$ also belongs to $\mathcal{F}_t$, the claim follows. □

Lemma 2.9. *Let X be continuous and $A \subseteq \mathbb{R}$ closed. Then τ_A is a stopping time.*

Proof. Define open sets $A_n \supseteq A$ by $A_n := \{x\colon d(x,A) < 1/n\}$. Since X is continuous and A is closed, it holds that

$$\{\omega\colon \tau_A(\omega) \leq t\} = \{\omega\colon \exists s \in [0,t],\ X_s(\omega) \in A\} = \{\omega\colon \forall n \in \mathbb{N}\ \exists s \in [0,t] \text{ with } X_s(\omega) \in A_n\}.$$

As all A_n are open, the last set equals

$$\{\omega\colon \forall n \in \mathbb{N} \exists q \in \mathbb{Q} \cap [0,t], X_q(\omega) \in A_n\} = \bigcap_{n\in\mathbb{N}} \bigcup_{q\in\mathbb{Q}\cap[0,t]} \underbrace{\{\omega\colon X_q(\omega) \in A_n\}}_{\in \mathcal{F}_q \subseteq \mathcal{F}_t}.$$

The right-hand side consists of countably many operations on sets from $\mathcal{F}_t$, hence it belongs to $\mathcal{F}_t$. ☐

The sigma-field $\mathcal{F}_\tau$. Next, we define the σ-field generated by all observable events up to a stopping time τ.

Definition 2.10. Given a stopping time τ. Then we call

$$\mathcal{F}_\tau := \big\{A \in \mathcal{F}\colon A \cap \{\tau \le t\} \in \mathcal{F}_t \,\forall t \ge 0\big\} \tag{2.4}$$

the σ-field of the events which are observable up to τ.

$\mathcal{F}_\tau$ is indeed a σ-field as is easily checked; a more intuitive characterization of $\mathcal{F}_\tau$ will be given in Lemma 2.17 below. It is easily seen that the rv $\omega \to \tau(\omega)$ is $\mathcal{F}_\tau$-measurable: For all $t_0 \ge 0$ we have

$$\{\tau \le t_0\} \cap \{\tau \le t\} = \{\tau \le (t_0 \wedge t)\} \in \mathcal{F}_{t_0 \wedge t} \subseteq \mathcal{F}_t\,.$$

Let S and T be stopping times, where $S \le T$. Intuitively, we can say, that if event A is observable up to time S, then event A is also observable up to T, so that one would expect the inclusion $\mathcal{F}_S \subseteq \mathcal{F}_T$. This is indeed true.

Lemma 2.11. *Given two stopping times S and T with $S \le T$. Then $\mathcal{F}_S \subseteq \mathcal{F}_T$.*

Proof. Since $S \le T$, it holds $\{T \le t\} \subseteq \{S \le t\}$. Using this we have for $A \in \mathcal{F}_S$:

$$A \cap \{T \le t\} = \underbrace{A \cap \{S \le t\}}_{\in \mathcal{F}_t,\text{ as } A \in \mathcal{F}_S} \cap \underbrace{\{T \le t\}}_{\substack{\in \mathcal{F}_t,\text{ as } T\\ \text{stopping time}}}\,,$$

and hence, $A \in \mathcal{F}_T$. ☐

Lemma 2.12. *Given two stopping times S and T. Then*

i) $S \wedge T := \min\{S,T\}$ *is a stopping time.*

ii) $S \vee T := \max\{S,T\}$ *is a stopping time.*

iii) $\mathcal{F}_{S\wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$.

Proof. We have $\{S \wedge T \le t\} = \{S \le t\} \cup \{T \le t\}$, $\{S \vee T \le t\} = \{S \le t\} \cap \{T \le t\}$. As S and T are stopping times, that means $\{S \le t\} \in \mathcal{F}_t$ and $\{T \le t\} \in \mathcal{F}_t$, claims i) and ii) are proved.
iii) is proved as follows. From Lemma 2.11 we see, that, $\mathcal{F}_{S\wedge T} \subseteq \mathcal{F}_S$ and $\mathcal{F}_{S\wedge T} \subseteq \mathcal{F}_T$, hence $\mathcal{F}_{S\wedge T} \subseteq \mathcal{F}_S \cap \mathcal{F}_t$. Now, let $A \in \mathcal{F}_S \cap \mathcal{F}_T$. As S and T are stopping times we have $A \cap \{S \le t\} \in \mathcal{F}_t$ and $A \cap \{T \le t\} \in \mathcal{F}_t$. It holds

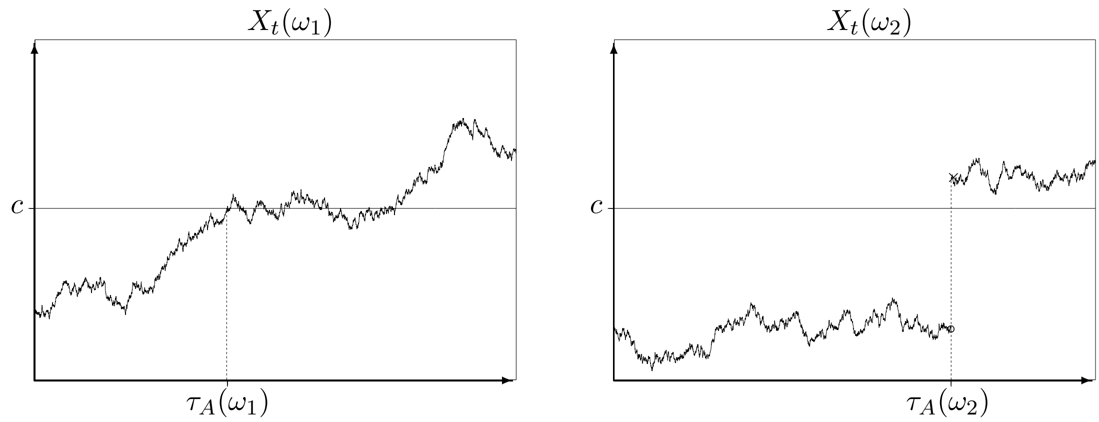
$$(A \cap \{S \le t\}) \cup (A \cap \{T \le t\}) = A \cap (\{S \le t\} \cup \{T \le t\}) = A \cap \{S \wedge T \le t\} \in \mathcal{F}_t,$$

and hence $A \in \mathcal{F}_{S\wedge T}$. ☐

Let $(X_t)_{t\geq 0}$ be a right-continuous stochastic process and let T be a stopping time. We define the stopped rv X_T by

$$X_T(\omega) := X_{T(\omega)}(\omega) \cdot 1_{\{T<\infty\}}(\omega), \quad \omega \in \Omega. \tag{2.5}$$

Example 2.13. Set $A := (c, \infty)$ and $T = \tau_A$.
In the following pictures it holds $X_T(\omega_1) = c$ and $X_T(\omega_2) > c$.

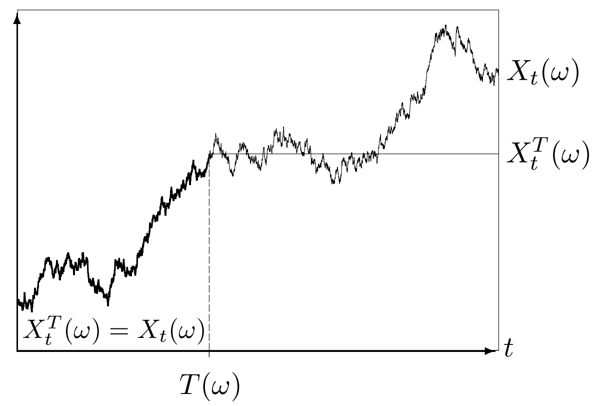


Lemma 2.14. *Let $(X_t)_{t\geq 0}$ be an $\{\mathcal{F}_t\}$-adapted and right-continuous stochastic process and let T be a stopping time. Then the rv X_T is $\mathcal{F}_T$-measurable.*

For a proof we refer to Protter (2005) or to Karatzas & Shreve (1988).

Definition 2.15. Given a right-continuous stochastic process $(X_t)_{t\geq 0}$ and a stopping time T. Then the *in T stopped process* $X^T = (X_t^T)_{t\geq 0}$ is defined by

$$X_t^T(\omega) := X_{t\wedge T(\omega)}(\omega) = \begin{cases} X_{T(\omega)}(\omega), & T(\omega) \leq t. \\ X_t(\omega), & T(\omega) > t. \end{cases} \tag{2.6}$$



Lemma 2.16. *If X is adapted and right-continuous, then the stopped process X^T is adapted.*

Proof. We have $X_t^T = X_t \cdot 1_{\{t<T\}} + X_T \cdot 1_{\{t\geq T\}}$. The first summand is $\mathcal{F}_t$-measurable, since it is a product consisting of two $\mathcal{F}_t$-measurable rvs. For the second summand we conclude as follows. $X_T \cdot 1_{\{t\geq T\}} = X_{T\wedge t} \cdot 1_{\{t\geq T\}}$. The rv $X_{T\wedge t}$ is $\mathcal{F}_{T\wedge t}$-measurable and $\mathcal{F}_{T\wedge t} \subseteq \mathcal{F}_t$. The rv $1_{\{t\geq T\}}$ is $\mathcal{F}_t$-measurable, as T is a stopping time. □

Finally we give a more intuitive interpretation of $\mathcal{F}_T$, which legitimates the description "σ-field of the observable events up to time T".

Lemma 2.17. *Let T be a stopping time. If $P(T < \infty) = 1$, then $\mathcal{F}_T = \sigma(X^T,\ X$ adapted and cadlag).*

Proof. Let X be adapted and cadlag. Then the rv $X_t^T = X_{T\wedge t}$ is $\mathcal{F}_{T\wedge t}$-measurable. Since $\mathcal{F}_{T\wedge t} \subseteq \mathcal{F}_T$ the rv X_t^T is also $\mathcal{F}_T$-measurable, hence $\mathcal{F}_T \supseteq \sigma(X^T,\ X$ adapted and cadlag). Now, let $A \in \mathcal{F}_T$ and define a stochastic process $X = (X_t)_{t\geq 0}$ by $X_t(\omega) = 1_A(\omega) \cdot 1_{\{T\leq t\}}(\omega)$. The process X is cadlag. It holds $\{X_t = 1\} = A \cap \{T \leq t\} \in \mathcal{F}_t$, hence X is adapted. Moreover we have $A = \bigcup_n \{X_n(\omega) = 1\}$, since T is finite. Hence, $A \in \sigma(X^T,\ X$ adapted and cadlag). □

2.2.2 The optional sampling theorem

The following result gives a crucial link between stopping times and martingales.

Theorem 2.18 (Optional sampling theorem)**.** *Consider an adapted stochastic process $X = (X_t)_{t\geq 0}$ with $E(|X_t|) < \infty$, $t \geq 0$. Then the following statements are equivalent.*

(1) X is a martingale.

(2) For all bounded stopping times τ ($\tau(\omega) \leq C$ for some $C > 0$, all $\omega \in \Omega$) one has $E(X_\tau) = E(X_0)$.

(3) Given two stopping times S and T such that $S \leq T \leq C$ for some $C > 0$. Then $E(X_T \mid \mathcal{F}_S) = X_S$.

We omit the proof; see for instance Protter (2005), Section I.2.

Corollary 2.19. *Let X be a martingale with right-continuous trajectories and let τ be a stopping time. Then the stopped process X^τ with $X_t^\tau = X_{t\wedge\tau}$ is also a martingale.*

See again Protter (2005), Section I.2 for a proof.

Corollary 2.20 (Martingale inequality)**.** *Let X be a right-continuous martingale such that $X_t > 0$ a.s. Then we have for $C > 0$*

$$P\left(\sup_{t\geq 0} X_t > C\right) \leq \frac{1}{C} E(X_0).$$

Proof. Put $T_C := \inf\{t \geq 0\colon X_t > C\}$. Since $X_t > 0$, we have for an arbitrary $n \in \mathbb{N}$

$$P\left(\sup_{0\leq t\leq n} X_t > C\right) \leq E\left(\frac{1}{C} X_{T_C\wedge n}\right) = \frac{1}{C} E(X_0),$$

where the last equality is due to Theorem 2.18, (2). For $n \to \infty$ we obtain the result by monotone convergence. □

2.3 Brownian Motion

Brownian motion is the most important building block for continuous-time asset pricing models. It has a long history in the modelling of random events in science. Around 1830 R. Brown, a Scottish botanist, discovered that molecules of water in a suspension perform an erratic movement under the buffeting of other water molecules. While Brown's research had no relation to mathematics this observation gave Brownian motion its name. In 1900 Bachelier introduced Brownian motion as model for stock-prices; see Bachelier (1900). In 1905 Einstein proposed Brownian motion as a mathematical model to describe the movement of particles in a suspension. The first rigorous theory of Brownian motion is due to N. Wiener (1923); therefore Brownian motion is often referred to as Wiener process.

2.3.1 Definition and Construction

Definition 2.21. A stochastic process $X = (X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ is standard one-dimensional *Brownian motion*, if

(i) $X_0 = 0$ a.s.

(ii) X has independent increments: for all $t, u \geq 0$ the increment $X_{t+u} - X_t$ is independent of X_s for all $s \leq t$.

(iii) X has stationary, normally distributed increments: $X_{t+u} - X_t \sim N(0, u)$.

(iv) X has continuous sample paths.

In honor of Brown and Wiener Brownian motion is often denoted by $(B_t)_{t\geq 0}$ or by $(W_t)_{t\geq 0}$.

Definition 2.22. Standard Brownian motion in $\mathbb{R}^d$ is a d-dimensional process $W_t = (W_t^1, ..., W_t^d)$ where $W^1, ..., W^d$ are d independent standard Brownian motions in $\mathbb{R}^1$.

Definition 2.21 has some elementary consequences:

(i) $W_t = W_t - W_0$ is $N(0, t)$-distributed.

(ii) Let $t > s$. Then the covariance of W_t and W_s is given by $\text{cov}(W_t, W_s) = E(W_t W_s) = E\big((W_t - W_s)W_s\big) + E(W_s^2) = E(W_t - W_s)E(W_s) + s = s$.

(iii) The finite-dimensional distributions of W are multivariate normal distributions with mean 0 and covariance matrix given in (ii).

Theorem 2.23. *A stochastic process with the properties of Definition 2.21 exists.*

Remarks: 1) There are various methods to construct Brownian motion and hence to prove Theorem 2.23, which are also useful for simulating Brownian sample paths; see for instance Karatzas & Shreve (1988).

2) Theorem 2.23 is more than a mere exercise in mathematical rigour: if we replace the normal distribution in Definition 2.21 (iii) by a fat-tailed α-stable distribution, the corresponding process – referred to as α-stable motion – necessarily has discontinuous trajectories; see for instance Section 2.4 of Embrechts, Klüppelberg & Mikosch (1997) for details.

2.3.2 Some stochastic properties of Brownian motion:

Proposition 2.24. *Let W_t be standard Brownian motion and define $\mathcal{F}_t := \sigma(W_s, s \le t)$. Then a) $(W_t)_{t\ge 0}$ b) $(W_t^2 - t)_{t\ge 0}$ and c) $\exp(\sigma W_t - 1/2\sigma^2 t)$ are martingales with respect to the filtration $\{\mathcal{F}_t\}$.*

Proof. We start with claim a). Let $t > s$; by point (ii) of Definition 2.21 the increment $W_t - W_s$ is independent of $\mathcal{F}_s$. Hence we get

$$E(W_t|\mathcal{F}_s) = E(W_t - W_s + W_s|\mathcal{F}_s) = E(W_t - W_s) + W_s = W_s\,.$$

To prove claim b) we first show that $E(W_t^2 - W_s^2|\mathcal{F}_s) = E((W_t - W_s)^2|\mathcal{F}_s)$. We have

$$\begin{aligned} E((W_t - W_s)^2|\mathcal{F}_s) &= E(W_t^2 - 2W_tW_s + W_s^2|\mathcal{F}_s) = E(W_t^2|\mathcal{F}_s) - 2W_sE(W_t|\mathcal{F}_s) + W_s^2 \\ &= E(W_t^2|\mathcal{F}_s) - 2W_s^2 + W_s^2 = E(W_t^2 - W_s^2|\mathcal{F}_s). \end{aligned}$$

The claim is proved if we can show that $E(W_t^2 - W_s^2|\mathcal{F}_s) = (t-s)$. By the first step of the proof this is equivalent to $E\left((W_t - W_s)^2|\mathcal{F}_s\right) = (t-s)$. Now $W_t - W_s$ is independent of $\mathcal{F}_s$; hence $E((W_t - W_s)^2|\mathcal{F}_s) = E((W_t - W_s)^2) = t - s$, as $W_t - W_s \sim N(0, t-s)$.

Sketch of c) Let $G_t = e^{(W_t - W_s - \frac{1}{2}(t-s))}$. Then we have, that $E(G_t) = G_sE(e^{(W_t - W_s - \frac{1}{2}(t-s))})$. Moreover, we get, using properties of lognormal distributions and the fact that $W_t - W_s$ is independent of W_s, that

$$E(G_t|\mathcal{F}_s) = G_sE(e^{(W_t - W_s - \frac{1}{2}(t-s))}|\mathcal{F}_s) = G_sE(e^{(W_t - W_s - \frac{1}{2}(t-s))}) = G_s$$

□

2.3.3 Quadratic Variation

Fix some point in time $\overline{T}$, which represents the time-point where our model ends. To define first and quadratic variation we need the notion of a partition of the interval $[0, \overline{T}]$.

Definition 2.25. A *partition* τ of $[0, \overline{T}]$ is a set of time-points $t_0 = 0 < t_1 < \ldots < t_n = \overline{T}$. The *mesh* of this partition is given by $|\tau| := \sup_{1\le i\le n} |t_i - t_{i-1}|$.

Definition 2.26 (First Variation). Consider a function $X : [0, \overline{T}] \to \mathbb{R}$. The first variation of X on $[0, \overline{T}]$ is defined as

$$\mathrm{Var}(X) := \sup \left\{ \sum_{t_i \in \tau} |X(t_i) - X(t_{i-1})|\,,\ \tau \text{ a partition of } [0, \overline{T}] \right\} \in [0,\, \infty]\,. \tag{2.7}$$

If $\mathrm{Var}(X) < \infty$ X is said to be of finite variation.

Remarks on notation: 1) Following standard conventions we denote by $\mathrm{Var}(f)$ the first variation of a function f, whereas $\mathrm{var}(Y)$ stands for the variance of a random variable Y.

2) Whenever a summation formula such as (2.7) contains the index t_{-1} it is understood that the corresponding summand is equal to zero.

Definition 2.27 (Quadratic Variation). Consider again a function $X : [0, \overline{T}] \to \mathbb{R}$ and a sequence $(\tau_n)_{n\in\mathbb{N}}$ of partitions of $[0, \overline{T}]$ such that $|\tau_n| \to 0$ as $n \to \infty$. Define for $t \in [0, \overline{T}]$ the quadratic variation of X along the partition τ_n by

$$V_t^2(X; \tau_n) := \sum_{t_i \in \tau_n;\, t_i < t} (X(t_i) - X(t_{i-1}))^2\,.$$

Assume that for all $t \in [0, \overline{T}]$ the limit $[X]_t := \lim_{n\to\infty} V_t^2(X;\tau_n)$ exists. In that case X is said to admit the quadratic variation $[X]_t$. If the function $t \to [X]_t$ is moreover continuous, we say that X has continuous quadratic variation.

In principle $[X]_t$ might depend on the sequence $(\tau_n)_{n\in\mathbb{N}}$. However, we are mainly interested in the case where X is a sample path of a continuous semimartingale such as Brownian motion. It can be shown that in this case $[X]_t$ is independent of the sequence of partitions used in its definition. Obviously $[X]_t$ is increasing in t and hence in particular of finite variation.

We now discuss the relation between first and quadratic variation.

Proposition 2.28. *If $X : [0, \overline{T}] \to \mathbb{R}$ is continuous and of finite variation, its quadratic variation $[X]_t$ is zero.*

By negating this result we have

Corollary 2.29. *If X is continuous and if the function $t \to [X]_t$ is strictly increasing, X is of infinite first variation on every subinterval $[a, b]$ of $[0, \overline{T}]$.*

Proof. (of Proposition 2.28) Choose a sequence of partitions τ_n of $[0, \overline{T}]$ such that $\lim_{n\to\infty} |\tau_n| = 0$. Then

$$\begin{aligned}\sum_{t_i\in\tau_n;\, t_i\le t} (X(t_i) - X(t_{i-1}))^2 &\le \sup_{t_i\in\tau_n} |X(t_i) - X(t_{i-1})| \sum_{t_i\in\tau_n} |X(t_i) - X(t_{i-1})| \\ &\le \sup_{t_i\in\tau_n} |X(t_i) - X(t_{i-1})| \operatorname{Var}(X). \qquad (2.8)\end{aligned}$$

Now note that $\operatorname{Var}(X) < \infty$ and that $\sup_{t_i\in\tau_n} |X(t_i) - X(t_{i-1})| \to 0$ for $n \to \infty$ as X is continuous and as $\lim_{n\to\infty} |\tau_n| = 0$. Hence the right side of (2.8) converges to zero which proves the proposition. □

The following result allows us to conclude that the quadratic variation of the sample paths of a continuous semimartingale is determined by the quadratic variation of its martingale part.

Proposition 2.30. *Assume that X is continuous with quadratic variation $[X]_t$ and consider a continuous function $A : [0, \overline{T}] \to \mathbb{R}$ which is of finite first variation. Let $Y_t := X_t + A_t$, $t \ge 0$. Then we have $[Y]_t = [X]_t$.*

Proof. We have

$$\begin{aligned}\sum_{t_i\in\tau_n;\, t_i\le t} (Y_{t_i} - Y_{t_{i-1}})^2 &= \sum_{t_i\in\tau_n;\, t_i\le t} (X_{t_i} - X_{t_{i-1}})^2 + \sum_{t_i\in\tau_n;\, t_i\le t} (A_{t_i} - A_{t_{i-1}})^2 \\ &+ 2 \sum_{t_i\in\tau_n;\, t_i\le t} (X_{t_i} - X_{t_{i-1}})(A_{t_i} - A_{t_{i-1}})\end{aligned}$$

Now $\sum_{t_i\in\tau_n;\, t_i\le t}(X_{t_i} - X_{t_{i-1}})^2$ converges to $[X]_t$ by assumption and $\sum_{t_i\in\tau_n}(A_{t_i} - A_{t_{i-1}})^2$ converges to zero as A is continuous and of finite variation. The last term can be estimated as follows:

$$\sum_{t_i\in\tau_n;\, t_i\le t} (X_{t_i} - X_{t_{i-1}})(A_{t_i} - A_{t_{i-1}}) \le \sup_{t_{i-1}\in\tau_n} |X_{t_i} - X_{t_{i-1}}| \operatorname{Var}(A),$$

which converges to zero as X is continuous. □

Now we deal with quadratic variation of the sample paths $B_\cdot(\omega)$ of Brownian motion. Roughly speaking, for (almost) all $\omega \in \Omega$ we have $[B_\cdot(\omega)]_t = t$. The following theorem makes this relation precise.

Theorem 2.31. *Consider a sequence of partitions τ_n of $[0,\overline{T}]$ such that $\lim_{n\to\infty} |\tau_n| = 0$. Then we have for all $t \in [0,\overline{T}]$ that $E\left(\left(V_t^2(B_\cdot(\omega);\tau_n) - t\right)^2\right) \to 0$ as $n \to \infty$.*

Proof. For a fixed partition τ_n we have

$$\begin{aligned} E\left(\Big(\sum_{t_i\in\tau_n, t_i\le t} (B_{t_i} - B_{t_{i-1}})^2 - t\Big)^2\right) &= E\left(\Big(\sum_{t_i\in\tau_n, t_i\le t} ((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}))\Big)^2\right) \\ &= \sum_{t_i,t_j\in\tau_n, t_i,t_j\le t} E\Big(\big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\big)\big((B_{t_j} - B_{t_{j-1}})^2 - (t_j - t_{j-1})\big)\Big) \\ &= \sum_{t_i\in\tau_n, t_i\le t} E\left((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\right)^2 . \end{aligned}$$

For the last equality we have used that $(B_{t_i} - B_{t_{i-1}})$ and $(B_{t_j} - B_{t_{j-1}})$ are independent for $i \neq j$ and that moreover $E\big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\big) = 0$, so that the mixed terms vanish. Now note that

$$E\left(\left((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i_1})\right)^2\right) = \operatorname{var}\left((B_{t_i} - B_{t_{i-1}})^2\right) .$$

It is well known that for a $N(\mu,\sigma^2)$-distributed rv ζ we have $\operatorname{var}(\zeta^2) = 2\sigma^4$. Hence $\operatorname{var}(B_{t_i} - B_{t_{i-1}})^2 = 2(t_i - t_{i-1})^2$, and we get

$$E\left(\Big(\sum_{t_i\in\tau_n, t_i\le t} (B_{t_i} - B_{t_{i-1}})^2 - t\Big)^2\right) = 2\sum_{t_i\in\tau_n, t_i\le t} (t_i - t_{i-1})^2 \le 2|\tau_n|t \to 0 ,$$

which proves the theorem. □

The type of convergence in Theorem 2.31 is known as $\mathcal{L}^2$-convergence. It implies in particular that $V_t^2(B_\cdot(\omega);\tau_n)$ converges to t in probability as $n \to \infty$. By exploiting the fact that every sequence of random variables which converges in probability has a subsequence which converges almost surely we obtain the following

Corollary 2.32. *There exists a sequence τ_n of partitions of $[0,\overline{T}]$ with $\lim_{n\to\infty} |\tau_n| = 0$ such that almost surely $V_t^2(B_\cdot(\omega);\tau_n) \to t$ for every $t \in [0,\overline{T}]$.*

This corollary is important as it shows that the pathwise Itô-calculus developed in Section 3 below applies to sample paths of Brownian motion.

Combining Theorem 2.31 and Corollary 2.29 yields another surprising property of Brownian sample paths.

Corollary 2.33. *Sample paths of Brownian motion are of infinite first variation.*

Remark 2.34. The sample paths of Brownian motion have many surprising properties. We refer the reader to Karatzas & Shreve (1988) and in particular to Revuz & Yor (1994) for further information.

We have seen that Brownian motion is a martingale with continuous trajectories and quadratic variation $[B_\cdot(\omega)]_t = t$. The following theorem, which is usually referred to as Levy's characterization of Brownian motion, establishes the converse:

Theorem 2.35. *If M is a martingale with continuous trajectories such that $M_0 = 0$ and $[M]_t = t \ \forall t$ then M is Brownian motion.*

Chapter 3

Pathwise Itô-Calculus

Motivation. Consider a function $f : \mathbb{R} \to \mathbb{R}$ which is once continuously differentiable (abbreviated f is a $\mathcal{C}^1$-function) with derivative f' and a $\mathcal{C}^1$-function $X : \mathbb{R}^+ \to \mathbb{R}$ with derivative $\dot{X} := \frac{\partial}{\partial t} X(t)$. The fundamental theorem of calculus yields

$$f(X(t)) - f(X(0)) = \int_0^t f'(X(s))\dot{X}(s)ds =: \int_0^t f'(X_s)dX_s. \tag{3.1}$$

A similar expression for the difference $f(X_t) - f(X_0)$ can be given if X is not $\mathcal{C}^1$ but only continuous and of finite variation:

Proposition 3.1. *Consider a continuous function $X : [0, \overline{T}] \to \mathbb{R}$ which is of finite variation and a $\mathcal{C}^1$-function $f : \mathbb{R} \to \mathbb{R}$ with derivative f'. Let τ_n denote a sequence of partitions of $[0, \overline{T}]$ with $\lim_{n\to\infty} |\tau_n| = 0$. Then we have that*

$$\lim_{n\to\infty} \sum_{t_i \in \tau_n, t_i \le t} f'(X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) =: \int_0^t f'(X_s)dX_s \tag{3.2}$$

exists. Moreover, we have the change of variable rule

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s. \tag{3.3}$$

Proposition 3.1 is a special case of Itô's formula (Theorem 3.2 below), hence we omit the proof.

3.1 Itô's formula

In this section we derive the Itô-formula – in financial texts often referred to as Itô's Lemma – which extends the chain-rule (3.3) to functions with infinite first but finite quadratic variation. Our exposition is based on the so-called pathwise Itô-calculus developed by Föllmer (1981); this approach allows us to give an elementary and relatively simple derivation of most results from stochastic calculus which are needed for the Black-Scholes option pricing model without having to develop the full theory of stochastic integration.

Throughout this section we consider a continuous function $X : [0,\overline{T}] \to \mathbb{R}$ which admits a continuous quadratic variation $[X]_t$ in the sense of Definition 2.27. As shown in Corollary 2.32 this is true for paths of Brownian motion. More generally, it can be shown that the sample paths of every continuous semimartingale admit a continuous quadratic variation.

As $[X]_t$ is increasing in t, the integral $\int_0^t g(s)d[X]_s$ is defined for every continuous function $g : [0,\overline{T}] \to \mathbb{R}$ in the ordinary 'Riemann-sense'; as $[X]_t$ is continuous this integral is moreover a continuous function of the upper bound t. Now we can state

Theorem 3.2 (Itô's formula). *Given a continuous function $X : [0,T] \to \mathbb{R}$ with continuous quadratic variation $[X]_t$. Let $F : \mathbb{R} \to \mathbb{R}$ denote a twice continuously differentiable function. Then we have for $t \le \overline{T}$*

$$F(X_t) = F(X_0) + \int_0^t F'(X_s)dX_s + \frac{1}{2}\int_0^t F''(X_s)d[X]_s \tag{3.4}$$

where

$$\int_0^t F'(X_s)dX_s := \lim_{n\to\infty} \sum_{t_i\in\tau_n;\, t_i\le t} F'(X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}). \tag{3.5}$$

Remarks: 1) The existence of the limit in (3.5) is shown in the proof of the theorem. The integral $\int_0^t F'(X_s)dX_s$ is called Itô-integral; it is a continuous function of the upper boundary t as is immediately apparent from (3.4).

2) The classical case of Proposition 3.1, where X is of finite variation is a special case of Theorem 3.2. If $[X]_t$ is non-zero the additional 'correction-term' $\frac{1}{2}\int_0^t F''(X_s)d[X]_s$ enters our formula for the differential $F(X_t) - F(X_0)$. We will see that this term is of crucial importance for most results in continuous-time finance.

3) Note that the sums used in defining the Itô-integral are non-anticipating, i.e. the integrand $F'(X_s)$ is evaluated at the left boundary of the interval $[t_{i-1}, t_i]$; we will see in Section 4.2 below that this makes the Itô-integral the right tool for the modeling of gains from trade.

4) Often formula (3.4) is expressed in the following short-hand notation: $dF(X_t) = F'(X_t)dX_t + \frac{1}{2}F''(X_t)d[X]_t$.

5) It is possible to give extensions of this theorem to the case where X has discontinuous sample paths; see for instance Chapter II.7 of Protter (2005).

Proof. As a first step we establish the following

Lemma 3.3. *For every piecewise continuous function $g : [0,T] \to \mathbb{R}$ we have*

$$\lim_{n\to\infty} \sum_{t_i\in\tau_n;\, t_i\le t} g(t_{i-1})(X_{t_i} - X_{t_{i-1}})^2 = \int_0^t g(s)d[X]_s. \tag{3.6}$$

Proof of the Lemma. Recall the definition of $V_t^2(X;\tau_n)$ in Definition 2.27. For indicator functions of the form $g(t) = 1_{(a,b]}(t)$ the convergence in (3.6) translates as

$$\lim_{n\to\infty} \left(V_b^2(X;\tau_n) - V_a^2(X;\tau_n)\right) = [X]_b - [X]_a\,,$$

which is satisfied by definition, as X admits the continuous quadratic variation $[X]_t$. For a general piecewise continuous function g the claim of the Lemma follows if we approximate g by piecewise constant functions.

Now we turn to the theorem itself. Consider $t_i, t_{i-1} \in \tau_n$, such that $t_i \leq t$ and denote by $(\Delta X)_{i,n}$ the increment $X_{t_i} - X_{t_{i-1}}$. We get from a Taylor-expansion of F

$$\begin{aligned} F(X_{t_i}) - F(X_{t_{i-1}}) &= F'(X_{t_{i-1}})(\Delta X)_{i,n} + \frac{1}{2}F''(X_{\tilde{t}})(\Delta X)^2_{i,n} \\ &= F'(X_{t_{i-1}})(\Delta X)_{i,n} + \frac{1}{2}F''(X_{t_{i-1}})(\Delta X)^2_{i,n} + R_{i,n}\,, \end{aligned}$$

where $\tilde{t}$ is some point in the interval (t_{i-1}, t_i), and where $R_{i,n} := \frac{1}{2}(F''(X_{\tilde{t}}) - F''(X_{t_{i-1}}))(\Delta X)^2_{i,n}$. Define $\delta_n := \max\{|X_t - X_{t_{i-1}}|\,, t \in [t_{i-1}, t_i], t_i \in \tau_n\}$. As X is continuous and as $|\tau_n| \to 0$ for $n \to \infty$ we have $\delta_n \to 0,\ n \to \infty$. Moreover,

$$|R_{i,n}| \leq \left(\frac{1}{2} \max_{|x-y|<\delta_n} |F''(x) - F''(y)|\right)(\Delta X)^2_{i,n} =: \varepsilon_n (\Delta X)^2_{i,n}\,.$$

Now $\varepsilon_n \to 0$, for $n \to \infty$ as F'' is uniformly continuous and as $\delta_n \to 0$. Hence

$$\left|\sum_{t_i \in \tau_n} R_{i,n}\right| \leq \sum_{t_i \in \tau_n} |R_{i,n}| \leq \varepsilon_n \sum (\Delta X)^2_{i,n} \to 0 \text{ as } n \to \infty\,,$$

as X admits the continuous quadratic variation $[X]_t$. Now

$$\begin{aligned} F(X_t) - F(X_0) &= \lim_{n\to\infty} \sum_{t_i \in \tau_n;\, t_i \leq t} F(X_{t_i}) - F(X_{t_{i-1}}) \\ &= \lim_{n\to\infty} \sum_{t_i \in \tau_n;\, t_i \leq t} F'(X_{t_{i-1}})(\Delta X)_{i,n} + \frac{1}{2} \sum_{t_i \in \tau_n;\, t_i \leq t} F''(X_{t_{i-1}})(\Delta X)^2_{i,n} + \sum_{t_i \in \tau_n;\, t_i \leq t} R_{i,n}\,. \end{aligned}$$

We have just shown that the sum over the $R_{i,n}$ tends to zero. Moreover, by Lemma 3.3 $\sum_{t_i \in \tau_n;\, t_i \leq t} F''(X_{t_{i-1}})(\Delta X)^2_{i,n}$ converges to $\int_0^t F''(X_s)d[X]_s$. Hence the limit

$$\int_0^t F'(X_s)dX_s := \lim_{n\to\infty} \sum_{t_i \in \tau_n;\, t_i \leq t} F'(X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})$$

exists, and we obtain the Itô-formula (3.4). □

Some Examples:

1) Take $F(x) = x^n$. Applying the Itô-formula yields

$$X_t^n = X_0^n + n\int_0^t X_s^{n-1} dX_s + \frac{n(n-1)}{2}\int_0^t X_s^{n-2} d[X]_s\,.$$

In short notation this can be written as $dX_t^n = nX_t^{n-1}dX_t + \frac{n(n-1)}{2}X_t^{n-2}d[X]_t$. In the special case where X is a sample path of a Brownian motion B with $B_0 = 0$ we obtain

$$B_t^2 = 2\int_0^t B_s dB_s + \int_0^t d[B]_s = 2\int_0^t B_s dB_s + t$$

2) Take $F(x) = e^x$. We get $e^{X_t} = e^{X_0} + \int_0^t e^{X_s} dX_s + \frac{1}{2}\int_0^t e^{X_s} d[X]_s$, or in short notation $de^{X_t} = e^{X_t}dX_t + \frac{1}{2}e^{X_t}d[X]_t$.

3.2 Properties of the Itô-Integral

3.2.1 Quadratic Variation

Throughout this section we consider a continuous function $X(t)$ with continuous quadratic variation $[X]_t$.

Proposition 3.4. *Let $F \in \mathcal{C}^1(\mathbb{R})$; then the function $t \to F(X_t)$ has quadratic variation $\int_0^t (F'(X_s))^2 \, d[X]_s$.*

Corollary 3.5. *For $f \in \mathcal{C}^1(\mathbb{R})$ the Itô-integral $I_t := \int_0^t f(X_s) dX_s$ is well-defined; its quadratic variation equals $[I]_t = \int_0^t f^2(X_s) d[X]_s$.*

Proof. Denote by $(\tau_n)_{n\in\mathbb{N}}$ a sequence of partitions of $[0, \overline{T}]$ with $|\tau_n| \to 0$. Then

$$\sum_{t_i\in\tau_n;\, t_i\le t} \big(F(X_{t_i}) - F(X_{t_{i-1}})\big)^2 = \sum_{t_i\in\tau_n;\, t_i\le t} \big(F'(X_{\tilde{t}_i})(\Delta X)_{i,n}\big)^2 \,,\; \tilde{t}_i \in (t_{i-1}, t_i)$$

$$= \sum_{t_i\in\tau_n;\, t_i\le t} F'(X_{t_{i-1}})^2 (\Delta X)^2_{i,n} + \sum_{t_i\in\tau_n;\, t_i\le t} \underbrace{(F'(X_{\tilde{t}_i})^2 - (F'(X_{t_{i-1}}))^2}_{\to 0,\; n\to\infty} (\Delta X)^2_{i,n} \,.$$

The first sum converges to $\int_0^t (F'(X_s))^2 \, d[X]_s$ by Lemma 3.3; a similar argument as in the proof of Theorem 3.2 shows that the second sum converges to zero as $n \to \infty$.

To proof the Corollary we define $F(x) = \int_0^x f(y)dy$, such that $F' = f$. As F is a $\mathcal{C}^2$-function the existence of the integral $I_t = \int_0^t F'(X_s)dX_s$ follows from Theorem 3.2. Moreover, we get from Itô's formula that

$$F(X_t) = F(X_0) + \int_0^t f(X_s)dX_s + \frac{1}{2}\int_0^t f'(X_s)d[X]_s =: F(X_0) + I_t + A_t \,.$$

As the function A is of finite variation we get $[I]_t = [F(X)]_t$. By Proposition (3.4), we know that $[F(X)]_t = \int_0^t f^2(X_s)d[X]_s$, which proves the corollary. □

Example: We compute the quadratic variation of the square of Brownian motion B. We have $B_t^2 = \int_0^t 2B_s dB_s + t$. Define $I_t := \int_0^t 2B_s dB_s$. We get $[B^2]_t = [I]_t = \int_0^t 4B_s^2 ds$.

3.2.2 Martingale-property of the Itô-integral

Up to now we have only used analytic properties of the function X such as the fact that X admits a continuous quadratic variation in our analysis of the Itô-integral. If $X(t)$ is the sample path of a stochastic process such as Brownian motion we may study probabilistic properties of the process $I_t(\omega) = \int_0^t f(X_s(\omega))dX_s(\omega)$. In particular we may consider the case that our integrator is a martingale.

If M is a martingale with trajectories of continuous quadratic variation and f a $\mathcal{C}^1$ function we expect the Itô-integral $I_t := \int_0^t f(M_s)dM_s$ to inherit the martingale property from M, as I_t is defined as limit of non-anticipating sums, $I_t = \lim_{n\to\infty} I_t^n$ with $I_t^n = \sum_{t_i\in\tau_n;\, t_i\le t} f(M_{t_{i-1}})(M_{t_i} - M_{t_{i-1}})$. The martingale property of the I_t^n is just a variation of the 'you can't gain by betting on a martingale' argument used already in our proof that the discounted gains from trade of an admissible selffinancing strategy are a martingale under an equivalent martingale measure in Chapter 1 (Lemma 1.7). Unfortunately some integrability problems arise when we pass from the approximating sums to the limit such that only a slightly weaker result is true. To state this result we need the notion of a local martingale.

Definition 3.6. A stochastic process M is called a local martingale, if there are stopping times $T_1 \leq \ldots \leq T_n \leq \ldots$ such that

(i) $\lim_{n\to\infty} T_n(\omega) = \infty$ a.s.

(ii) $(M_{T_n \wedge t})_{t\geq 0}$ is a martingale for all n.

Obviously every martingale in the sense of Section 2.1.2 (every true martingale) is a local martingale. The opposite assertion is not true; see for instance Remark 3.10 below.

Theorem 3.7. *Consider a local martingale M with continuous trajectories and continuous quadratic variation $[M]_t$ and a function $f \in \mathcal{C}^1(\mathbb{R})$. Then $I_t(\omega) = \int_0^t f(M_s(\omega))dM_s(\omega)$ is a local martingale.*

Partial proof. We restrict ourselves to the case where M is a bounded martingale and where f is bounded; the general case follows by localization (introduction of an increasing sequence of stopping time $(T_n)_{n\in\mathbb{N}}$). The proof goes in two steps.

a) Let $(\tau_n)_{n\in\mathbb{N}}$ be a sequence of partitions with $|\tau_n| \to 0$, and fix n. Then the discrete-time process $I_k^n := \sum_{t_i\in\tau_n, i\leq k} f(M_{t_{i-1}})(M_{t_i} - M_{t_{i-1}})$, $k \leq n$, is a martingale wrt the discrete filtration $\{\mathcal{F}_k^n\}_k$ with $\mathcal{F}_k^n := \mathcal{F}_{t_k}$, as can be seen from the following easy argument.

$$E(I_k^n - I_{k-1}^n|\mathcal{F}_{k-1}^n) = E(f(M_{t_{k-1}})(M_{t_k} - M_{t_{k-1}})|\mathcal{F}_{t_{k-1}}) = f(M_{t_{k-1}})E((M_{t_k} - M_{t_{k-1}})|\mathcal{F}_{t_{k-1}}),$$

and the last term is obviously equal to zero as M is a martingale. Note that here we have used the fact that the Itô-integral is non-anticipating.

b) Let $s < t$. We will show that $E(I_t 1_A) = E(I_s 1_A)$ for all $A \in \mathcal{F}_s$, as this implies that $E(I_t|\mathcal{F}_s) = I_s$. Choose $t_n, s_n \in \tau_n$ with $t_n \searrow t$, $s_n \searrow s$ and $t_n > s_n$. By Step a) we have

$$E(I_{t_n} 1_A) = E(I_{s_n} 1_A);$$

moreover $I_{t_n} \to I_t$, $I_{s_n} \to I_s$, as I has continuous paths. Moreover, one can show that $(I_{t_n})_n$ and $(I_{s_n})_n$ are uniformly integrable (using the boundedness of M and f), so that the claim follows from the theorem of Lebesgue. □

Remark 3.8. If f is defined only on a subset $G \subseteq \mathbb{R}$ the process $I_t = \int_0^t f(M_s)dM_s$ can be defined up to the stopping-time $\tau = \inf\{t > 0,\ M_t \notin G\}$ and it can be shown that I_t is a local martingale until τ.

In applications one often needs to decide if a local martingale M is in fact a true martingale. The following Proposition provides a useful criterion for this

Proposition 3.9. *Let M be a local martingale with continuous trajectories. Then the following two assertions are equivalent.*

(i) M is a true martingale and $E(M_t^2) < \infty$ $\forall t \geq 0$.

(ii) $E([M]_t) < \infty \ \forall t$.

If either (i) or (ii) holds, we have $E\big((M_t - M_0)^2\big) = E([M]_t)$.

For a proof and a generalization to discontinuous martingales see Chapter II.6 of Protter (2005).

Remark 3.10. The following process is an example of a local martingale which is not a true martingale. Consider a three-dimensional Brownian motion $W_t = (W_t^1, W_t^2, W_t^3)$ with $W_0 = (1, 1, 1)$ and define

$$M_t = \frac{1}{||W_t||} = \frac{1}{\sqrt{(W_t^1)^2 + (W_t^2)^2 + (W_t^3)^2}} .$$

Then M is a local martingale, as can be checked using the Itô-formula in higher dimensions, but it is not a full martingale; see again Chapter II.6 of Protter (2005) for details.

The following Proposition shows that interesting martingales with continuous trajectories are necessarily of infinite variation.

Proposition 3.11. *Consider a local martingale* M *with continuous trajectories of finite variation Then the paths of* M *are constant, i.e.* $M_t = M_0$ *almost surely.*

Note that there are martingales with discontinuous non-constant trajectories of finite variation; an example is provided by the compensated Poisson process; see Section 2.1.2.

Proof. By Itô's-formula we get for M_t^2

$$M_t^2 = M_0^2 + 2\int_0^t M_s dM_s + [M]_t = M_0^2 + 2\int_0^t M_s dM_s ,$$

as $[M]_t = 0$ by Proposition 2.28. The martingale-property of the Itô-integral implies that M_t^2 is a local martingale. Assume that both M_t and M_t^2 are a real martingales.[1] Then we have

$$0 \le E\left((M_t - M_0)^2\right) = E\left(M_t^2 - 2M_t M_0 + M_0^2\right) = M_0^2 - 2M_0^2 + M_0^2 = 0 ,$$

which shows that $E\left(M_t - M_0\right)^2 = 0$ so that $M_t = M_0$ a.s. □

We close this section with a formal definition of semimartingales.

Definition 3.12. A stochastic process X with RCLL paths is called a *semimartingale* if X has a decomposition of the form $X_t = X_0 + M_t + A_t$ where M is a local martingale and A is an adapted process with continuous trajectories of finite variation; M is called the martingale part of X, A the finite variation part.

Note that the decomposition of X into a martingale part and a finite variation part is unique by Proposition 3.11.

[1] For an argument how to deal with the case where M_t^2 is only a local martingale we refer the reader to Protter (2005).

3.3 Covariation and d-dimensional Itô-formula

3.3.1 Covariation

Fix a sequence τ_n of partitions of $[0,\overline{T}]$ with $\tau_n \to 0$ and continuous functions X, Y which admit a continuous quadratic variation $[X]_t$ and$[Y]_t$ along the sequence τ_n.

Definition 3.13. Assume that for all $t \in [0,T]$ the following limit exists:

$$\lim_{n\to\infty} \sum_{t_i\in\tau_n;\, t_i\le t} (X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}) =: [X,Y]_t.$$

Then $[X,Y]_t$ is called *covariation* of X and Y.

Theorem 3.14. *$[X,Y]_t$ exists if and only if $[X+Y]_t$ exists; in that case we have the following so-called polarization-identity*

$$[X,Y]_t = \frac{1}{2}\left([X+Y]_t - [X]_t - [Y]_t\right) . \tag{3.7}$$

Proof. Recall the notation $(\Delta X)_{i,n} = X_{t_i} - X_{t_{i-1}}$, for $t_i, t_{i-1} \in \tau_n$. We have

$$\begin{aligned}
[X+Y]_t &= \lim_{n\to\infty} \sum_{t_i\in\tau_n;\, t_i\le t} ((\Delta X)_{i,n} + (\Delta Y)_{i,n})^2 \\
&= \lim_{n\to\infty} \left\{ \sum_{t_i\in\tau_n;\, t_i\le t} (\Delta X)_{i,n}^2 + \sum_{t_i\in\tau_n;\, t_i\le t} (\Delta Y)_{i,n}^2 + 2 \sum_{t_i\in\tau_n;\, t_i\le t} (\Delta X)_{i,n}(\Delta Y)_{i,n} \right\} \\
&= [X]_t + [Y]_t + 2\lim_{n\to\infty} \sum_{t_i\in\tau_n;\, t_i\le t} (\Delta X)_{i,n}(\Delta Y)_{i,n} \,.
\end{aligned} \tag{3.8}$$

Hence the last limit on the right hand side of (3.8) exists iff $[X+Y]_t$ exists. Solving for this limit yields the polarization identity. □

Note that $[X,Y]_t$ is of finite variation as it is the difference of monotone functions. We now use the polarization identity to compute the covariation for a few important examples.

1) If X is a continuous function with continuous quadratic variation $[X]_t$ and A a continuous function of finite variation we have $[X+A]_t = [X]_t$ and hence $[X,A]_t = 0$.

2) Consider two independent Brownian motions B^1, B^2 on our probability space $(\Omega, \mathcal{F}, P)$. Then $[B^1\cdot(\omega), B^2\cdot(\omega)]_t = 0$. To prove this claim we have to compute $[B^1+B^2]_t$. Note that $(B^1_t + B^2_t)/\sqrt{2}$ is again a Brownian motion and has therefore quadratic variation equal to t. Hence

$$\frac{1}{2}([B^1+B^2]_t - [B^1]_t - [B^2]_t) = \frac{1}{2}(2t - t - t) = 0\,.$$

3) Consider a continuous function X with continuous quadratic variation, and $\mathcal{C}^1$-functions f and g. Define $Y_t := \int_0^t f(X_s)dX_s$ and $Z_t := \int_0^t g(X_s)dX_s$. Then $[Y,Z]_t = \int_0^t f(X_s)g(X_s)d[X]_s$. This follows from the polarization identity and the following computation:

$$[Y+Z]_t \;=\; \int_0^t (f+g)^2(X_s)d[X]_s = [Y]_t + [Z]_t + 2\int_0^t f(X_s)g(X_s)d[X]_s.$$

Example 3) is a special case of a more general rule for stochastic Itô-integrals.

3.3.2 The d-dimensional Itô-formula

Theorem 3.15 (d-dimensional Itô-formula)**.** *Given continuous functions* $X = (X^1, \ldots, X^d) : [0, \overline{T}] \to \mathbb{R}$ *with continuous covariation*

$$[X^k, X^l]_t = \begin{cases} [X^k]_t, \ k{=}l, \\ 1/2\left([X^k + X^l]_t - [X^k]_t - [X^l]_t\right), \ k \neq l \end{cases}$$

and a twice continuously differentiable function $F : \mathbb{R}^d \to \mathbb{R}$*. Then*

$$F(X_t) = F(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} F(X_s) dX_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} F(X_s) d[X^i, X^j]_s \,.$$

Remark on notation: For $\frac{\partial}{\partial x_i} F$ we often write F_{x_i}, $\frac{\partial^2}{\partial x_i \partial x_j} F$ is denoted by F_{x_i, x_j}. In short-notation the d-dimensional Itô-formula hence writes as:

$$dF(X_t) = \sum_{i=1}^d F_{x_i}(X_t) dX_t^i + \frac{1}{2} \sum_{i,j=1}^d F_{x_i, x_j}(X_t) d[X^i, X^j]_t \,.$$

Example: Let $W = (W^1, \ldots, W^d)$ be d-dimensional Brownian motion so that that $[W^k, W^l]_t = \delta_{kl} t$ where $\delta_{kl} = 1$ if $k = l$ and $\delta_{kl} = 0$ otherwise. Hence we have

$$F(W_t) = F(W_0) + \sum_{i=1}^d \int_0^t F_{x_i}(W_s) dW_s^i + \frac{1}{2} \sum_{i=1}^d \int_0^t F_{x_i, x_i}(W_s) ds \,. \tag{3.9}$$

Corollary 3.16 (Itô's product formula)**.** *Given* X, Y *with continuous quadratic variation* $[X]_t$, $[Y]_t$ *and covariation* $[X, Y]_t$*. Then*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + [X, Y]_t.$$

Proof. Apply Theorem (3.15) to $F(x, y) = xy$. □

In short notation the product formula can be written as $d(XY)_t = X_t dY_t + Y_t dX_t + d[X, Y]_t$.

Corollary 3.17 (Itô-formula for time-dependent functions)**.** *Given a continuous function* X *with continuous quadratic variation* $[X]_t$ *and a function* $F(t, x)$ *which is once continuously differentiable in* t *and twice continuously differentiable in* x*. Then*

$$F(t, X_t) = F(0, X_0) + \int_0^t F_t(s, X_s) ds + \int_0^t F_x(s, X_s) dX_s + \frac{1}{2} \int_0^t F_{xx}(s, X_s) d[X]_s.$$

We now consider several applications of the d-dimensional Itô-formula.

1) *Geometric Brownian motion:* Given a Brownian motion W, an initial value $S_0 > 0$ and constants μ, σ with $\sigma > 0$ we define geometric Brownian motion S by

$$S_t = S_0 \exp\left(\sigma W_t + (\mu - 1/2\sigma^2) t\right) \,.$$

Geometric Brownian motion will be our main model for the fluctuation of asset prices in Section 4. Using the Itô-formula we now derive a more intuitive expression for the dynamics of S. Define $X_t := \sigma B_t$ and $Y_t := (\mu - 1/2\sigma^2)t$ and note that $[X]_t = \sigma^2 t$ and $[Y]_t = [X,Y]_t = 0$. Let $F(x,y) := S_0 \exp(x+y)$ such that $F_x = F_y = F_{xx} = F$. By definition $S_t = F(X_t, Y_t)$, and we get

$$
\begin{aligned}
S_t &= S_0 + \int_0^t F(X_s,Y_s)dX_s + \int_0^t F(X_s,Y_s)dY_s + \frac{1}{2}\int_0^t F(X_s,Y_s)d[X]_s \\
&= S_0 + \int_0^t F(X_s,Y_s)\sigma dB_s + \int_0^t F(X_s,Y_s)(\mu - \frac{1}{2}\sigma^2)ds + \frac{1}{2}\int_0^t F(X_s,Y_s)\sigma^2 ds \\
&= S_0 + \int_0^t \sigma S_s dB_s + \int_0^t \mu S_s ds\,. \qquad (3.10)
\end{aligned}
$$

In our short-notation the equation solved by S can be written as $dS_t = \mu S_t dt + \sigma S_t dB_t$. In the special case where $\mu = 0$ we get that $S_t = S_0 + \int_0^t \sigma S_s dB_s$ is a local martingale.[2]

2) Brownian motion and the reverse heat-equation. Consider a function $F(t,x)$ that solves the reverse heat-equation $F_t(t,x) + 1/2F_{xx}(t,x) = 0$ and a Brownian motion B. Then $F(t,B_t)$ is a local martingale. The proof is again based on Itô's formula. We get

$$
\begin{aligned}
F(t,B_t) &= F(0,B_0) + \int_0^t F_x(s,B_s)dB_s + \int_0^t F_t(s,B_s)ds + \frac{1}{2}\int_0^t F_{xx}(s,B_s)d[B]_s \\
&= F(0,B_0) + \int_0^t F_x(s,B_s)dB_s + \int_0^t (F_t + \frac{1}{2}F_{xx})(s,B_s)ds \\
&= F(0,B_0) + \int_0^t F_x(s,B_s)dB_s\,.
\end{aligned}
$$

In case that $F_x(t,B_t)$ is sufficiently integrable $F(t,B_t)$ is even a real martingale. In that case one easily obtains a probabilistic representation of the solution of the reverse heat equation. For more on the interplay between solutions of partial differential equations and stochastic processes we refer to Chapter 4 and 5 of Karatzas & Shreve (1988).

[2] It can be shown that in that case S is even a true martingale.

Chapter 4

The Black-Scholes Model: a PDE-Approach

We now have all the mathematical tools at hand we need to study pricing and hedging of derivatives in the classical Black-Scholes model.

4.1 Asset Price Dynamics

As in the classical paper Black & Scholes (1973) we consider a market with two traded assets, a risky non-dividend-paying stock and a riskless money market account. The price of the stock at time t is denoted by S_t^1, the price of the money market account by S_t^0. For simplicity we work with a deterministic continuously compounded interest rate r such that $S_t^0 = \exp(rt)$. We now look for appropriate models for the dynamics of the stock-price. As usual we work on a filtered probability space $(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_t\}$ supporting a standard Brownian motion W_t representing the uncertainty in our market.

In his now famous PhD-thesis Bachelier (1900) proposed to model asset prices by an arithmetic Brownian motion, i.e. he suggested the model $S_t^1 = S_0 + \sigma W_t + \mu t$ for constants $\mu, \sigma > 0$. While this was a good first approximation to the dynamics of stock prices, arithmetic Brownian motion has one serious drawback: as S_t^1 is $N(S_0 + \mu t, \sigma^2 t)$ distributed, the asset price can become negative with positive probability, which is at odds with the fact that real-world stock-prices are always nonnegative because of limited liability of the shareholders.

Samuelson (1965) therefore suggested replacing arithmetic Brownian motion by geometric Brownian motion

$$S_t^1 = S_0^1 \exp\left(\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t\right). \tag{4.1}$$

We know from (3.10) that this model solves the linear stochastic differential equation (SDE)

$$dS_t^1 = \mu S_t^1 dt + \sigma S_t^1 dW_t \,.$$

Geometric Brownian motion - often referred to as Black-Scholes model - is nowadays widely used as reference model both in option pricing theory and in the theory of portfolio-optimization; we therefore adopt it as our model for the stock price dynamics in this section. Model (4.1) implies that log-returns

$$\ln S^1_{t+h} - \ln S^1_t = \sigma(W_{t+h} - W_t) + (\mu - \frac{1}{2}\sigma^2)h$$

are $N((\mu - \frac{1}{2}\sigma^2)h, \sigma^2 h)$-distributed; in particular the volatility σ is the instantaneous standard deviation of the log-returns. Moreover, under (4.1) log-returns over non-overlapping time periods are stochastically independent.

There are a number of practical and theoretical considerations which make geometric Brownian motion attractive as a model for stock price dynamics.

- Geometric Brownian motion fits asset prices data reasonably well, even if the fit is far from perfect. For an overview the empirical deficiencies of the Black-Scholes model we refer to Chapter I of Cont & Tankov (2004) or Section 4.1 of McNeil, Frey & Embrechts (2005).
- Geometric Brownian motion allows for explicit pricing formulae for a relatively large class of derivatives.
- The Black-Scholes model is quite robust as a model for hedging of derivatives: if real asset-price dynamics are 'not too different from geometric Brownian motion' hedging strategies computed using the Black-Scholes model perform reasonably well. There is now a large literature on *model risk* in derivative pricing; see for instance the collection of papers in Gibson (2000) or Cont (2006); we present a brief discussion of the model-risk related to volatility-misspecification in Subsection 4.4.2 below.

There are also a number of theoretical considerations in favour of the Black-Scholes model.

- The model is in line with the efficient markets hypothesis. Moreover, there are a number of economic models which show that the Black-Scholes model can be sustained as a model for economic equilibrium; see for instance He & Leland (1993) for a rational expectations model and Föllmer & Schweizer (1993) for a model based on the temporary equilibrium concept.
- The Black-Scholes model is an arbitrage-free and complete model, making derivative pricing straightforward from a conceptual point of view.

4.2 Pricing and Hedging of Terminal Value Claims

Consider now a contingent claim with maturity date T and payoff H. As in the discrete-time setup of Chapter 1 we want to find a dynamic trading strategy replicating the claim; such a strategy can be used for pricing and hedging purposes. It can be shown that in the framework of the Black-Scholes model such a strategy exists for every claim whose payoff is measurable with respect to the information generated by the asset price. However, such a result requires the notion of the stochastic Itô-integral $\int_0^t \xi_s dS^1_s$ for general predictable processes ξ which we do not have at our disposal. We therefore restrict our analysis to so-called terminal value claims whose payoff is of the form $H = h(S^1_T)$. For these claims one can find Markov hedging strategies which are functions of time and the current stock-price.

This includes most examples which are relevant from a practical viewpoint; as shown in Section 4.4.1 extensions to path-dependent derivatives are also possible. For the general theory we refer the reader to Bingham & Kiesel (1998) or to the advanced text Karatzas & Shreve (1998).

4.2.1 Basic Notions

TRADING STRATEGY: A *Markov trading strategy* is given by a pair of smooth functions $(\phi(t,S),\ \eta(t,S))$, where $\phi(t,S_t^1)$ and $\eta(t,S_t^1)$ give the number of stocks respectively of units of the money market account in the portfolio at time t. The value at time t of this strategy is given by $V(t,S_t^1)=S_t^1\phi(t,S_t^1)+\eta(t,S_t^1)S^0(t)$. Note that the strategy can alternatively be described by specifying the functions $\phi(t,S)$ and $V(t,S)$; the position in the money market account is then given by the function $\eta(t,S):=(V(t,S)-S\phi(t,S))/S^0(t)$.

GAINS FROM TRADE: To motivate the subsequent definitions we introduce piecewise constant approximations to our trading strategies. Consider a sequence τ_n of partitions with $|\tau_n|\to 0$ and define

$$\phi_t^n(\omega)=\sum_{t_i\in\tau_n}\phi(t_{i-1},S^1_{t_{i-1}}(\omega))\,1_{(t_{i-1},t_i]}(t) \tag{4.2}$$

$$\eta_t^n(\omega)=\sum_{t_i\in\tau_n}\eta(t_{i-1},S^1_{t_{i-1}}(\omega))\,1_{(t_{i-1},t_i]}(t) \tag{4.3}$$

and $V_t^n=\phi_t^nS_t^1+\eta_t^nS_t^0$. A well-known argument from discrete-time finance now yields that this piecewise constant strategy is selffinancing if and only if we have for all $t_i\in\tau_n$

$$V_{t_i}^n=V_0+G_t^n,\ \text{where } G_t^n=\sum_{j=1}^{i}\left(\phi_{t_j}^n(S^1_{t_j}-S^1_{t_{j-1}})+\eta_{t_j}^n(S^0_{t_j}-S^0_{t_{j-1}})\right).$$

Now recall that by definition of the Itô-integral G_t^n converges to $\int_0^t\phi(s,S_s^1)dS_s^1+\int_0^t\eta(s,S_s^1)dS_s^0$. Hence the following definition is natural.

Definition 4.1. Given a Markov trading strategy $(\phi(t,S_t^1),\eta(t,S_t^1))$ induced by smooth functions $\phi,\ \eta:\ [0,T]\times\mathbb{R}^+\to\mathbb{R}$.

(i) The *gains from trade* of this strategy are given by

$$G_t=\int_0^t\phi(s,S_s^1)dS_s^1+\int_0^t\eta(s,S_s^1)dS_s^0\,.$$

(ii) The strategy is *selffinancing*, if $V(t,S_t^1)=V(0,S_0)+G_t$ for all $t\le T$.

(iii) Consider a terminal value claim with payoff $h(S_T^1)$. A selffinancing strategy is a *replicating strategy* for the claim if $V(T,S)=h(S)$ for all $S>0$; in that case $V(t,S_t^1)$ is the fair price of the claim at time t.

4.2.2 The pricing-equation for terminal-value-claims

We now derive a partial differential equation (PDE) for the value of the replicating strategy. We have the following

Theorem 4.2. *Let $V : [0,T] \times \mathbb{R}^+ \to \mathbb{R}$ be a continuous function which solves the PDE*

$$V_t(t,S) + \frac{1}{2}\sigma^2 S^2 V_{SS}(t,S) + rSV_S(t,S) = rV(t,S)\,, \quad (t,S) \in [0,T) \times \mathbb{R}^+\,. \tag{4.4}$$

Then the hedging strategy with stock-position $\phi(t,S) = V_S(t,S)$ and value $V(t,S)$ is self-financing. If V satisfies moreover the terminal condition $V(T,S) = h(S)$, the strategy replicates the terminal value claim with payoff $h(S_T^1)$ and the fair price at time t of the claim equals $V(t,S_t^1)$.

Proof. As a first step we compute the quadratic variation of geometric Brownian motion. Recall that

$$S_t^1 = S_0 + \int_0^t \sigma S_s^1 dW_s + \int_0^t \mu S_s^1 ds =: M_t + A_t\,.$$

As A is of finite variation we get $[S^1]_t = [M]_t = \int_0^t \sigma^2 (S_s^1)^2 ds$.

Now we turn to the first claim. We get from Itô's formula

$$\begin{aligned} V(t,S_t^1) &= V(0,S_0^1) + \int_0^t V_S(s,S_s^1)dS_s^1 + \int_0^t V_t(s,S_s^1)ds + \frac{1}{2}\int_0^t V_{SS}(s,S_s^1)d[S^1]_s \\ &= V(0,S_0^1) + \int_0^t V_S(s,S_s^1)dS_s^1 + \int_0^t \left(V_t(s,S_s^1) + \frac{1}{2}\sigma^2 (S_s^1)^2 V_{SS}(s,S_s^1) \right) ds\,. \end{aligned}$$

Using the PDE (4.4) and the definition of ϕ this equals

$$\begin{aligned} &= V(0,S_0^1) + \int_0^t \phi(s,S_s^1)dS_s^1 + \int_0^t r(V(s,S_s^1) - \phi(s,S_s^1)S_s^1)ds \\ &= V(0,S_0^1) + \int_0^t \phi(s,S_s^1)dS_s^1 + \int_0^t \eta(s,S_s^1)dS_s^0, \end{aligned}$$

where $\eta(t,S_t^1) = (V(t,S_t^1) - \phi(t,S_t^1)S_t^1)/S^0(t)$ is the position in the money-market account which corresponds to our strategy. Hence our strategy is selffinancing. The remaining claims are obvious. ☐

4.3 The Black-Scholes formula

4.3.1 The formula

To price a European call option we have to solve the PDE (4.4) with terminal condition $h(S) = (S-K)^+$. To solve this problem analytically one usually reduces the PDE (4.4) to the heat equation by a proper change of variables. This technique is useful also for the implementation of numerical schemes to solve the pricing PDE; see for instance Wilmott, Dewynne & Howison (1993). Details are given in the following Lemma.

Lemma 4.3. *Define $\tau(t) = \sigma^2(T-t)$ and $z(t,S) = \ln S + (r - \frac{1}{2}\sigma^2)(T-t)$. Denote by $u(t,z) : [0,T/\sigma^2] \times \mathbb{R} \to \mathbb{R}$ the solution of the heat-equation $u_t = \frac{1}{2}u_{zz}$ with initial condition $u(0,z) = (e^z - K)^+$. Then $C(t,S) := e^{-r(T-t)}u(\tau(t), z(t,S))$ solves the terminal value problem for the price of a European call.*

Proof. We have $C(T,S) = u(\tau(T), z(T,S)) = u(0, \ln S) = (S-K)^+$, so that the function C has the right value at maturity. Moreover,

$$\frac{\partial C}{\partial t} = e^{-r(T-t)} \left(ru - \sigma^2 u_\tau + (1/2\sigma^2 - r) u_z \right)$$
$$\frac{\partial C}{\partial S} = e^{-r(T-t)} u_z 1/S\,, \; \frac{\partial^2 C}{\partial S^2} = e^{-r(T-t)} \left(u_{zz} (1/S)^2 - u_z (1/S)^2 \right)$$

Next we plug these expressions into the PDE (4.4). We get (omitting the arguments (t,S) respectively $(\tau(t), z(t,S))$)

$$\begin{aligned} &\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} - rC \\ &= e^{-r(T-t)} \left(ru - \sigma^2 u_\tau + (1/2\sigma^2 - r) u_z + r u_z + \frac{1}{2}\sigma^2 (u_{zz} - u_z) - ru \right) \\ &= e^{-r(T-t)} \sigma^2 \left(-u_\tau + \frac{1}{2} u_{zz} \right), \end{aligned}$$

and the last term is obviously equal to zero as u solves the heat equation. ☐

It is well-known that the solution u of the heat-equation with initial condition $u(0,z) = u_0(z)$ equals

$$u(\tau, z) = \frac{1}{\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} u_0(x) e^{-\frac{(z-x)^2}{2\tau}} dx.$$

From this follows after tedious but straightforward computations (see for instance Sandmann (1999) or Wilmott et al. (1993))

Theorem 4.4. *Denote by $N(\cdot)$ the standard normal distribution function. The no-arbitrage price of a European call with strike K and time to maturity T in the Black-Scholes model with volatility σ and interest rate r is given by*

$$C_{BS}(t, S; \sigma, r, K, T) := SN(d_1) - e^{-r(T-t)} K N(d_2)\,, \tag{4.5}$$

with

$$d_1 = \frac{\ln S/K + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \text{ and } d_2 = d_1 - \sigma\sqrt{T-t}\,. \tag{4.6}$$

The corresponding hedge-portfolio consists of $\frac{\partial}{\partial S} C_{BS} = N(d_1)$ units of the risky asset and $(C_{BS}(t,S) - N(d_1)S)/e^{rt} = -e^{-rT} K N(d_2)$ units of the money market account.

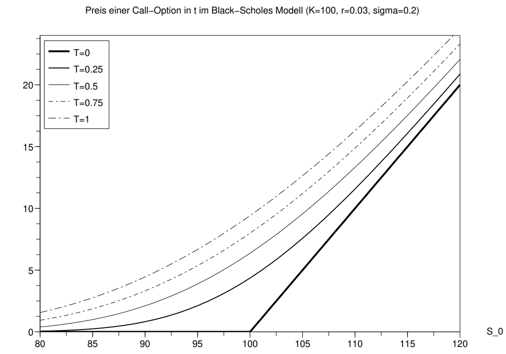
A probabilistic derivation of the Black-Scholes formula is given in Section 6.2 below.

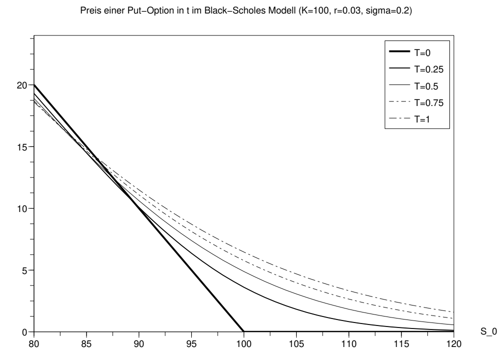
4.3.2 Properties of option prices and the Greeks

Option prices. According to the Put-Call parity there is the following relation between the price in t of a European call (denoted C_t) and the price of a European put with the same characteristics K, T (denoted P_t): $C_t + e^{-r(T-t)} K = S_t + P_t$. This gives for the Black-Scholes price of a European Put

$$P_{BS}(t, S; \sigma, r, K, T) = -S_t N(d_1) + K e^{-r(T-t)} N(-d_2)\,,$$

where d_1 and d_2 are as in (4.6). The next two pictures give the call and put price as a function of the current stock price:





The hedge ratio or Δ of an option. The *delta* of an option is the derivative wrt the price of the underlying. In the Black Scholes model we have

$$\Delta_C = \frac{\partial C}{\partial S} = N(d_1) \text{ and } \Delta_P = \frac{\partial P}{\partial S} = \Delta_C - 1 = -N(-d_1)$$

The Delta is relevant for so-called delta-hedging:

- The hedge-portfolio for a call consists of $\frac{\partial}{\partial S} C_{BS} = N(d_1)$ units of S^1 and $(C_{BS}(t,S) - N(d_1)S)/e^{rt} = -e^{-rT} K N(d_2)$ units of S^0.
- The hedge-portfolio for a put consists of $\frac{\partial}{\partial S} P_{BS} = -N(-d_1)$ units of S^1 and of

$$(P_{BS}(t,S) + (1 - N(d_1))S)/e^{rt} = e^{-rT} K N(-d_2)$$

units of S^0. Note that *Delta* is increasing and that $0 < \Delta_C < 1$ and $-1 < \Delta_P < 0$.

The Gamma of an option. The Gamma of an option is the second derivative wrt the underlying: $\Gamma_C = \frac{\partial^2 C}{\partial S^2}$. It holds that

$$\Gamma_C = \Gamma_P = \frac{\varphi(d_1)}{S_t \sigma \sqrt{T-t}},$$

where φ denotes the density of the standard normal distribution. The Gamma measures how fast the Delta changes and hence how often a hedge needs rebalancing. A large Gamma means that small changes in the price of the underlying lead to large changes in the hedge portfolio; options with a large Gamma are therefore difficult to hedge in practice.

Further Greeks. The other partial derivatives of the option price with respect to the input parameters have (pseudo) Greek names as well. Most relevant is the so-called *Vega* (not really a Greek letter). Vega is the derivative wrt volatility: $\text{Vega}_C = \frac{\partial C}{\partial \sigma}$. It holds that $\text{Vega}_C = \text{Vega}_P = S_t \varphi(d_1)\sqrt{T-t}$. Vega is always positive, as a higher volatility makes hedging more expensive, see also Section 4.4.2 below.

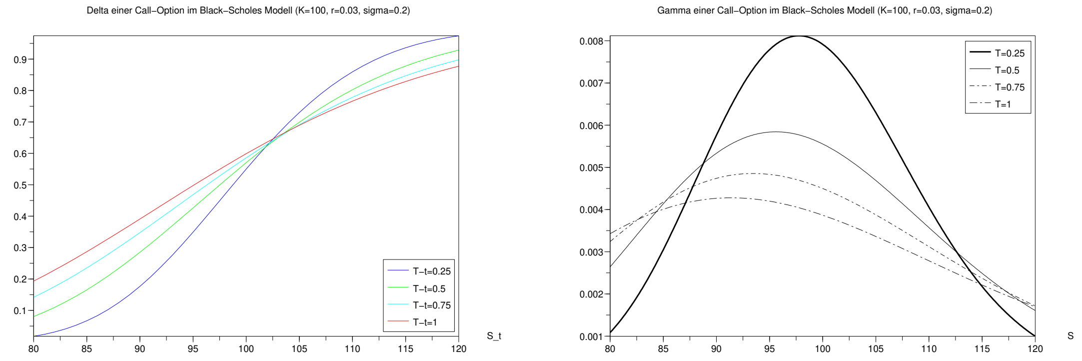


Figure 4.1: Delta (left) and Gamma (right) for a Call as a function of current price S for sevaral values of $T-t$

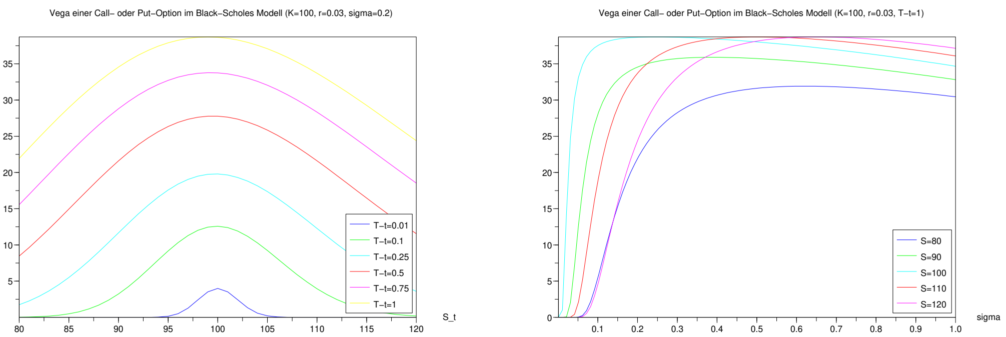


Figure 4.2: Vega of a call/put as a function of S for different $T-t$ (left) and as a function of σ for different S (right)

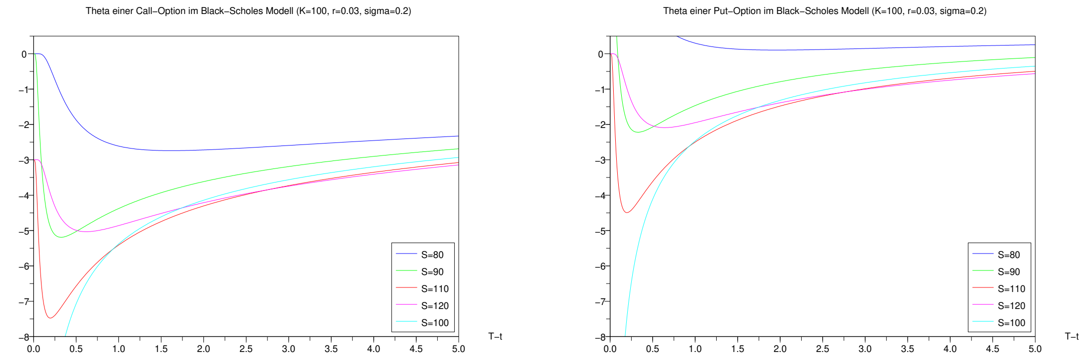


Figure 4.3: Theta of a call (left) and of a put (right). Note that the Theta of a put can become positive for small S.

Finally, we give the other Greeks.

$$\begin{aligned}
\Theta_C &= \frac{\partial C}{\partial t} = -\frac{S_t \sigma \varphi(d_1)}{2\sqrt{T-t}} - rKe^{-r(T-t)}N(d_2) (\text{sensitivity wrt calender time}) \\
\Theta_P &= \frac{\partial P}{\partial t} = -\frac{S_t \sigma \varphi(d_1)}{2\sqrt{T-t}} + rKe^{-r(T-t)}N(-d_2) \\
\rho_C &= \frac{\partial C}{\partial r} = K(T-t)e^{-r(T-t)}N(d_2) \text{ Rho (interest-rate sensitivity)} \\
\rho_P &= \frac{\partial P}{\partial r} = -K(T-t)e^{-r(T-t)}N(-d_2)
\end{aligned}$$

4.3.3 Volatility estimation

For an extensive discussion how the Black-Scholes formula can be applied in practice we refer to Cox & Rubinstein (1985) and Hull (1997). Here we content ourselves with a few remarks about possible approaches to determine the volatility σ. As volatility is not directly observable – in contrast to the other parameters in the Black-Scholes formula – finding a 'good' value for σ is by for the most problematic part in applying the Black-Scholes formula. The fact that in real markets volatility is rarely constant but tends to fluctuate in a rather unpredictable manner makes matters even worse.[1] There are two common approaches to determining σ.

1) Historical volatility: This approach is based on statistical considerations. Recall that under (3.10) log-returns over non-overlapping periods of length Δ are independent and $N((\mu - \frac{1}{2}\sigma^2)\Delta, \sigma^2\Delta)$ distributed. Given asset price data at times t_i, $i = 1, \ldots, N$ with $t_i - t_{i-1} = \Delta$ (e.g. daily returns) define $Y_i = \ln S_{t_i} - \ln S_{t_{i-1}}$. The standard estimator from elementary statistics for σ_Δ, the volatility of the log-returns over the time-period Δ is given by

$$\hat{\sigma}_\Delta = \left(\frac{1}{N-1} \sum_{i=1}^{N} \left(Y_i - \overline{Y}\right)^2 \right)^{\frac{1}{2}}, \quad \text{where } \overline{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i.$$

[1]The stochastic nature of volatility has given rise to the development of the stochastic volatility models; see for instance Frey (1997) for an overview.

Estimated historical volatility is then given by $\hat{\sigma}_{hist} = \hat{\sigma}_\Delta/\sqrt{\Delta}$.

2) Implied volatility The idea underlying the implied volatility concept is the use of observed prices of traded derivatives to find the 'prediction of the market' for the volatility of the stock. To explain the concept we consider the following example:

Assume that a call option with strike K and maturity T is traded at time t and at a given stock-price $(S_t^1)^*$ for a price of C_t^*. The implied volatility $\hat{\sigma}_{impl}$ is then given by the solution to the equation

$$C_{BS}(t,(S_t^1)^*;\hat{\sigma}_{impl},K,T) = C_t^*\,.$$

As C_{BS} is strictly increasing in σ a unique solution to this equation exists; it is usually determined by numerical procedures.

In practice traders tend to use a combination of both approaches, implied volatility being the slightly more popular concept.

4.4 Further applications

4.4.1 Path-dependent derivatives – the case of barrier options

The approach of Section 4.2 can be extended to many exotic options with path-dependent payoff; see for instance Wilmott et al. (1993). Here we content ourselves with a simple example.

Consider a so-called down-and-out call with strike price K and barrier M. The down-and out call is a particular barrier option; the payoff of this contract equals

$$H = \begin{cases} (S_T-K)^+, \text{ if } S_t^1 > M \text{ for all } t\in[0,T], \\ 0 \text{ if } S_t^1 < M \text{ for some } t\in[0,T]. \end{cases}$$

Define the stopping-time $\tau = \inf\{t>0\,,\ S_t^1 < M\}$ and denote by $V(t,S_t^1)$ the value of the down-and-out option on the set $\tau > t$, i.e. provided that the stock-price has not yet crossed the barrier. Again, we are looking for a selffinancing strategy which replicates this payoff. We have the following

Proposition 4.5. *Assume that $V(t,S)$ solves the following boundary value problem*

$$V_t(t,S) + \frac{1}{2}\sigma^2 S^2 V_{SS}(t,S) + rSV_S(t,S) = rV(t,S) \text{ for } (t,S)\in[0,T)\times(M,\infty) \tag{4.7}$$

with terminal condition $V(T,S) = (S-K)^+$ and boundary condition $V(t,M) = 0$. Then the fair price of the down-and-out call equals $V(t,S_t^1)$ if $\tau > t$ and 0 if $\tau \le t$; if $\tau > t$ the stock-position of the replicating strategy consists of $\phi(t,S_t^1) := V_S(t,S_t^1)$ shares of stock.

The proof is similar to the proof of Theorem 4.2.

4.4.2 Model Risk

We finally study the implications of volatility misspecification and stochastic volatility for the performance of hedging strategies. For more on this issue we refer to El Karoui, Jeanblanc-Picqué & Shreve (1998) and to the papers collected in Gibson (2000).

We assume that the asset price follows the SDE

$$dS_t^1 = \mu S_t^1 dt + \sigma_t S_t^1 dW_t$$

for some – possibly stochastic – volatility σ_t. For simplicity we assume that $r = 0$. We consider a trader who uses the Black-Scholes model with volatility σ^* in pricing and hedging a terminal value claim and who maintains a self-financing portfolio. Denote by h^{BS} the solution of the PDE terminal value problem from Theorem 4.2 for $r = 0$, i.e.

$$h_t^{BS}(t,S) + \frac{1}{2}(\sigma^*)^2 S^2 h_{SS}^{BS} = 0\,, \quad h^{BS}(T,S) = h(S)\,. \tag{4.8}$$

We assume that the trader follows the Black-Scholes model and holds $h_S^{BS}(t, S_t^1)$ shares of stock at time t. If he maintains a selffinancing portfolio the actual value at T of his portfolio equals

$$V_T = V_0 + \int_0^T h_S^{BS}(t, S_t^1) dS_t^1\,.$$

Definition 4.6. The tracking error of the hedge is given by $e_T = h(S_T) - V_T$.

Note that the hedge produces a loss if $e_T > 0$ and a gain if $e_T < 0$. We have the following expression for the tracking error.

Proposition 4.7. *The tracking error equals*

$$e_T = \frac{1}{2}\int_0^T (S_t^1)^2 \left(\sigma_t^2 - (\sigma^*)^2\right) h_{SS}^{BS}(t, S_t^1) dt.$$

The proposition shows that the tracking error is proportional to $(\sigma_t^2 - (\sigma^*)^2)$, the estimation error for volatility, and to the average of the 'Gamma' $h_{SS}^{BS}(t, S_t^1)$ over the future path of the stock-price process. If $h_{SS}^{BS}(t, S_t^1) > 0$ the hedge looses (gains) money if $\sigma_t > \sigma^*$ ($\sigma_t < \sigma^*$); if $h_{SS}^{BS}(t, S_t^1) < 0$ the hedge looses (gains) money if $\sigma < \sigma^*$ ($\sigma_t > \sigma^*$).

Proof. As $h^{BS}(T, S) = h(S)$ we get from Itô's formula:

$$h(S_T) = h^{BS}(0, S_0) + \int_0^T h_S^{BS}(t, S_t^1) dS_t^1 + \int_0^T \left(h_t^{BS}(t, S_t^1) + \frac{1}{2}\sigma_t^2 (S_t^1)^2 h_{SS}^{BS}(t, S_t^1)\right) dt\,.$$

This implies that

$$e_T = \int_0^T \left(h_t^{BS}(t, S_t^1) + \frac{1}{2}\sigma_t^2 (S_t^1)^2 h_{SS}^{BS}(t, S_t^1)\right) dt\,.$$

By the Black-Scholes PDE (4.8) we have $h_t^{BS}(t,S) = -\frac{1}{2}(\sigma^*)^2 S^2 h_{SS}^{BS}(t,S)$; hence

$$e_T = \frac{1}{2}\int_0^T (S_t^1)^2 (\sigma_t^2 - (\sigma^*)^2) h_{SS}^{BS}(t, S_t^1) dt\,.$$

□

Chapter 5

Further Tools from Stochastic Calculus

5.1 Stochastic Integration for Continuous Martingales

In this section we want to define the stochastic integral $\int_0^t H_s dX_s$ for semimartingales of the form $X = X_0 + M + A$, where M is a continuous martingale and A is a continuous FV-process, and where H is a suitable limit of integrands of the form

$$H_t^n = \sum_{i=0}^{n-1} h_{t_i}(\omega) 1_{(t_i, t_{i+1}]}(t), \quad t_i \in \tau_n,$$

where τ_n is a fixed partition of $(0, T]$, and where each of the h_{t_i} is $\mathcal{F}_{t_i}$-measurable. This extends the pathwise Itô calculus to a larger class of integrands. This is reasonable from a financial viewpoint, as we want to work with a larger class of trading strategies than just Markov strategies of the form $\Phi(t, S_t)$. The key point is to define the integral $\int_0^t H_s dM_s$, as M will typically have paths of infinite first variation, so that standard Stieltjes integration does not apply. To overcome these problems we use special properties of the space $\mathcal{M}^2$ of square integrable martingales.

5.1.1 The Spaces $\mathcal{M}^2$ and $\mathcal{M}^{2,c}$

Throughout we work on a given a filtered probability space $(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_t\}$ with fixed horizon T.

Definition 5.1. The space of all martingales $M = (M_t)_{0 \le t \le T}$ with $M_0 = 0$ and $E(M_T^2) < \infty$ is denoted $\mathcal{M}^2$; $\mathcal{M}^{2,c} \subseteq \mathcal{M}^2$ denotes the subspace of all martingales with continuous trajectories.

Note that the Jensen inequality gives for $M \in \mathcal{M}^2$, $t \le T$ that

$$E(M_t^2) = E\left((E(M_T|\mathcal{F}_t))^2\right) \le E\left(E(M_T^2|\mathcal{F}_t)\right) = E(M_T^2),$$

so that M_t is square integrable for all $0 \le t \le T$.

Recall that a Hilbert space H is a linear space with scalar product $\langle \cdot, \cdot \rangle_H$ that is complete in the associated norm $\|x\|_H = \langle x, x \rangle_H^{1/2}$.

Lemma 5.2. *$\mathcal{M}^2$ is a Hilbert-space with scalar product*

$$\langle M, N\rangle_{\mathcal{M}^2} = E(M_T N_T) = \langle M_T, N_T\rangle_{\mathcal{L}^2}.$$

Proof. $\langle M, N\rangle_{\mathcal{M}^2}$ is a scalar product. In particular, if $\langle M, M\rangle_{\mathcal{M}^2} = 0$, we have $M_T = 0$ a.s. which yields $M_t = E(M_T|\mathcal{F}_t) = 0$ for all t. The Hilbert-space property is obvious, as $\mathcal{M}^2$ can be identified with $\mathcal{L}^2(\Omega, \mathcal{F}_T, P)$ by identifying a martingale $(M_t)_{t\leq T}$ with its terminal value M_T via $M_t = E(M_T|\mathcal{F}_t)$. □

The following result establishes a link between the pathwise supremum the and the terminal value of a martingale.

Theorem 5.3 (Doob inequality)**.** *Let $p > 1$ and $(M_t)_{0\leq t\leq T}$ be a martingale. Define q by $\frac{1}{p} + \frac{1}{q} = 1$. Then we have for $t \leq T$*

$$\|\sup_{s\leq t} |M_s|\|_{\mathcal{L}^p} \leq q\|M_t\|_{\mathcal{L}^p}. \tag{5.1}$$

In particular we get with $M_T^* := \sup_{t\leq T} |M_t|$ that

$$E\left((M_T^*)^2\right) \leq 4E\left(M_T^2\right). \tag{5.2}$$

We omit the proof.

Lemma 5.4. *$\mathcal{M}^{2,c}$ is a closed subspace of $\mathcal{M}^2$ and hence a Hilbert-space.*

Proof. Consider a sequence $M^n \in \mathcal{M}^{2,c}$ with $M^n \to M$, i.e. with $\lim_{n\to\infty} E\left((M_T^n - M_T)^2\right) = 0$. We want to show that the limit $(M_t)_{t\geq 0}$ with $M_t = E\left(M_T|\mathcal{F}_t\right)$ has continuous sample paths. By Doob's inequality we get

$$E\left(((M^n - M)_T^*)^2\right) \leq 4E\left((M_T^n - M_T)^2\right) \to 0,$$

i.e. $(M^n - M)_T^* = \sup_{t\leq T} |M_t^n - M_t| \to 0$ in $\mathcal{L}^2(\Omega, \mathcal{F}_T, P)$. It follows that $(M^{n'} - M)_T^*$ converges to zero a.s. for a subsequence n', which implies the result, as the limit of continuous functions in the supremum-norm is continuous. □

Quadratic variation. Recall the definition

$$V_t^2(M; \tau_n) := \sum_{t_i\in\tau_n;\, t_i<t} (M_{t_i} - M_{t_{i-1}})^2 = \sum_{t_i\in\tau_n;\, t_i<t} (\Delta M_{i,n})^2, \tag{5.3}$$

for the quadratic variation of M along a partition τ_n of $[0, T]$.

Theorem 5.5. *Consider a sequence $(\tau_n)_n$ of partitions of $[0, T]$ such that $|\tau_n| \to 0$ and let $M \in \mathcal{M}^2$. Then*

1. *$V_t^2(M, \tau_n)$ converges in probability to an increasing process $[M]_t$ with $[M]_0 = 0$ and with $\Delta[M]_t = (\Delta M)_t^2$.*
2. *$M_t^2 - [M]_t$ is a martingale, in particular we have $E\left(M_t^2\right) = E\left([M]_t\right)$, $t \leq T$.*

3. $[M]_t$ is uniquely defined by the requirements

(i) $[M]_t$ is increasing, $\Delta[M]_t = (\Delta M)_t^2$, and $[M]_0 = 0$.

(ii) $M_t^2 - [M]_t$ is a martingale.

Remark 5.6. 1. The hard parts of the theorem are 1 and 2; the uniqueness of $[M]_t$ follows immediately from the fact that a continuous martingale with trajectories of finite variation is a.s. constant.

2. The convergence of $V_t^2(M, \tau_n) \to [M]_t$ in probability implies that $V_t^2(M, \tau_{n'}) \to [M]_t$ almost surely for a subsequence n'; hence the pathwise Itô-calculus of Chapter 3 applies to martingales from $\mathcal{M}^{2,c}$

3. For $M \in \mathcal{M}^{2,c}$ (continuous trajectories) it holds that $[M]_t$ is continuous.

Example 5.7. For a standard Brownian motion W we obviously have $[W]_t = t$; for a compensated Poisson-process with parameter λ given by $M_t = N_t - \lambda t$ for N a Poisson process with parameter λ we have $[M]_t = N$.

Recall that for two functions $X, Y : (0, T] \to \mathbb{R}$ such that $[X]_t$ and $[Y]_t$ exist, the quadratic covariation

$$[X, Y]_t = \lim_{n\to\infty} \sum_{t_i \in \tau_n;\, t_i \le t} \Delta X_{i,n} \Delta Y_{i,n}$$

exists if and only if $[X + Y]_t$ exists and that $[X, Y]_t = \frac{1}{2}\left([X + Y]_t - [X]_t - [Y]_t\right)$, the so-called polarization identity. Now for $M, N \in \mathcal{M}^{2,c}$, $M + N \in \mathcal{M}^{2,c}$ and $[M + N]_t$ exists by Theorem 5.5. Hence we get

Corollary 5.8. *For $M, N \in \mathcal{M}^2$ the covariation*

$$[M, N]_t = \lim_{n\to\infty} \sum_{t_i \in \tau_n;\, t_i \le t} \Delta M_{i,n} \Delta N_{i,n}$$

exists. $[M, N]$ is characterized by the following properties

(i) $[M, N]_t$ is a FV-process with $[M, N]_0 = 0$ and with $\Delta[M, N]_t = \Delta M_t \Delta N_t$.

(ii) $M_t N_t - [M, N]_t$ is a martingale.

The result follows from Theorem 5.5 and the polarization identity.

Note that $[\cdot, \cdot]$ is bilinear and symmetric on $\mathcal{M}^2$, i.e. $[\alpha M_1 + \beta M_2, N] = \alpha[M_1, N] + \beta[M_2, N]$. Note that for $M, N \in \mathcal{M}^{2,c}$, $[M, N]$ gives the conditional covariance of the increments of M and N, i.e. we have for $0 \le s \le t \le T$ that

$$E\left((M_t - M_s)(N_t - N_s)\,|\mathcal{F}_s\right) = E\left([M, N]_t - [M, N]_s|\mathcal{F}_s\right). \tag{5.4}$$

The proof of (5.4) is easy. We have

$$\begin{aligned}
E\left((M_t - M_s)(N_t - N_s)|\mathcal{F}_s\right) &= E\left(M_t N_t - M_t N_s - M_s N_t + M_s N_s|\mathcal{F}_s\right) \\
&= E\left(M_t N_t|\mathcal{F}_s\right) - \underbrace{\left(E\left(M_t N_s|\mathcal{F}_s\right) + E\left(M_s N_t|\mathcal{F}_s\right)\right)}_{=2M_s N_s} + \underbrace{E\left(M_s N_s|\mathcal{F}_s\right)}_{=M_s N_s} \\
&= E\left(M_t N_t|\mathcal{F}_s\right) - M_s N_s \\
&= E\left([M, N]_t - [M, N]_s|\mathcal{F}_s\right),
\end{aligned}$$

where we have used the fact that $M_t N_t - [M, N]_t$ is a martingale.

Example 5.9 (Constructing correlated Brownian motions.)**.** Let W_1 and W_2 be independent Brownian motions and recall from Section 3.3.1 that $[W_1, W_2]_t = 0$. Put $B_1 = W_1$, $B_2 = \rho W_1 + \sqrt{1-\rho^2} W_2$. Then B_1 and B_2 are standard one-dimensional Brownian motions, and $[B_1, B_2]_t = \rho[W_1, W_1]_t + \sqrt{1-\rho^2}[W_2, W_1]_t = \rho t$.

5.1.2 Stochastic Integrals for elementary processes

Consider a martingale $M \in \mathcal{M}^{2,c}$. In this section we sketch the construction of the integral $\int_0^t H_s dM_s$. In principle the same arguments go through for $M \in \mathcal{M}^2$; we omit the details. We begin by taking H from the class of elementary processes, defined as follows:

$$\xi \quad := \quad \Big\{ H : H_t(\omega) = \sum_{i=0}^{n-1} h_{t_i}(\omega) 1_{(t_i, t_{i+1}]}(t) \Big\}, \tag{5.5}$$

for deterministic time points $0 = t_0 < t_1 < ... < t_{n-1} < T = t_n$, $n \in \mathbb{N}$, and bounded, $\mathcal{F}_{t_i}$-measurable random variables h_{t_i}, $0 \le i \le n-1$. Note that the elements of ξ are bounded and left-continuous. From a financial point of view, elements of ξ correspond to simple, piecewise constant trading strategies.

Definition 5.10. For $H \in \xi$ and $M \in \mathcal{M}^{2,c}$ we define the stochastic integral $(H.M)$ by

$$(H.M)_t := \int_0^t H_s dM_s := \sum_{0 \le i \le n} h_{t_i} \left(M_{t_{i+1} \wedge t} - M_{t_i \wedge t} \right), \ t \le T;$$

H is called integrand, M integrator.

Lemma 5.11 (Properties of $(H.M)$ for $H \in \xi$)**.** *We have* $(H.M) \in \mathcal{M}^2$ *and the quadratic variation is given by* $[\,(H.M)\,]_t = \int_0^t H_s^2 \, d[M]_s$.

Corollary 5.12 (Itô-isometry)**.** *We have*

$$E\left((H.M)_t^2\right) = E\left((\int_0^t H_s dM_s)^2 \right) = E\left(\int_0^t H_s^2 d[M]_s \right). \tag{5.6}$$

Proof of Lemma 5.11. The continuity of $(H.M)$ is obvious, as M is continuous. In order to show that $(H.M)$ is a martingale, by the optional sampling theorem we have to show that for all bounded stopping times τ with $\tau \le T$ a.s. we have $E\left((H.M)_\tau\right) = (H.M)_0$. Now $(H.M)_0$ is obviously equal to zero. Moreover,

$$\begin{aligned} E\big((H.M)_\tau\big) &= E\left(\sum_{0 \le i \le n} h_{t_i}(M_{t_{i+1}\wedge\tau} - M_{t_i \wedge \tau}) \right) \\ &= \sum_{0 \le i \le n} E\left(h_{t_i} E\left(M_{t_{i+1}\wedge\tau} - M_{t_i\wedge\tau} | \mathcal{F}_{t_i} \right) \right). \end{aligned}$$

By the optional sampling theorem $E\left(M_{t_{i+1}\wedge\tau} - M_{t_i\wedge\tau}|\mathcal{F}_{t_i}\right) = 0$, and the martingale property of $(H.M)$ follows. In order to show that $[(H.M)]_t = \int_0^t H_s^2 d[M]_s$, we show that $(H.M)_t^2 - \int_0^t H_s^2 d[M]_s$ is a martingale, which gives the result by Statement 3 of Theorem 5.5. Again we use the optional sampling theorem. Consider a stopping time $\tau \le T$. We have

$$E\left((H.M)_\tau^2\right) = \sum_{i,j=1}^{n+1} E\left(h_{t_{i-1}} h_{t_{j-1}} \Delta M_i^\tau \Delta M_j^\tau \right), \tag{5.7}$$

where $\Delta M_i^\tau = M_{t_i \wedge \tau} - M_{t_{i-1} \wedge \tau}$. Suppose that $i < j$. By conditioning on $\mathcal{F}_{t_{j-1}}$ we see that $E\left[h_{t_{i-1}} h_{t_{j-1}} \Delta M_i^\tau \Delta M_j^\tau\right] = 0$. Hence (5.7) equals

$$\sum_{i=1}^{n+1} E\left(h_{t_{i-1}}^2 (\Delta M_i^\tau)^2\right) = \sum_{i=1}^{n+1} E\left(h_{t_{i-1}}^2 \left([M^\tau]_{t_i} - [M^\tau]_{t_{i-1}}\right)\right),$$

using (5.4). Now we obviously have $[M^\tau]_t = [M]_{\tau \wedge t}$, so that

$$\begin{aligned} \sum_{i=1}^{n+1} E\left(h_{t_{i-1}}^2 \left([M^\tau]_{t_i} - [M^\tau]_{t_{i-1}}\right)\right) &= \sum_{i=1}^{n+1} E\left(h_{t_{i-1}}^2 \left([M]_{t_i \wedge \tau} - [M]_{t_{i-1} \wedge \tau}\right)\right) \\ &= E\left(\int_0^\tau H_t^2 d[M]_t\right). \end{aligned}$$

The relation $E\left((H.M)_T^2\right) = E\left(\int_0^T H_s^2\, d[M]_s\right)$, which we obtain by putting $\tau \equiv T$, shows that $(H.M)_T \in \mathcal{L}^2(\Omega, \mathcal{F}_T, P)$, as H is bounded and as $E([M]_T) = E\left(M_T^2\right) < \infty$ since $M \in \mathcal{M}^2$. $\square$

5.1.3 Extension to General Integrands

The Itô-isometry to extend the integral $(H.M)$ from ξ to a larger class of integrands that is defined next.

Definition 5.13. Fix some martingale $M \in \mathcal{M}^{2,c}$. Then $\overline{\xi}$ denotes the set of all adapted, left-continuous processes H such that $E\left(\int_0^T H_s^2\, d[M]_s\right) < \infty$.

The following theorem shows that the stochastic integral $(H.M)$ can be defined for $H \in \overline{\xi}$:

Theorem 5.14. *Fix $M \in \mathcal{M}^{2,c}$ and consider an adapted, left-continuous processes $H \in \overline{\xi}$, i.e. with $E\left(\int_0^T H_s^2\, d[M]_s\right) < \infty$.*

i) There exists a sequence of simple predictable strategies H^n with

$$\lim_{n \to \infty} E\left(\int_0^T (H_s^n - H_s)^2\, d[M]_s\right) = 0.$$

ii) There is a process $(H.M) \in \mathcal{M}^{2,c}$ with $\lim_{n\to\infty} \|(H^n.M) - (H.M)\|_{\mathcal{M}^2} = 0$, and $(H.M)$ is independent of the sequence H^n.

Definition 5.15. $(H.M)$ is called stochastic integral of H with respect to M; H is called integrand, M is the integrator.

Proof. The proof of (i) is quite technical, see for instance Karatzas & Shreve (1988), Chapter III. In order to establish (ii) we begin with an abstract interpretation of $E\left(\int_0^T H_s^2\, d[M]_s\right)$ for $M \in \mathcal{M}^{2,c}$ fixed.

Put $\bar{\Omega} = \Omega \times [0,T]$, $\bar{\mathcal{F}} = \mathcal{F}_T \otimes \mathcal{B}([0,T])$ and define a measure $\bar{P}_M$ on $\bar{\Omega}$, $\bar{\mathcal{F}}$ by

$$\bar{P}_M(A) = E\left(\int_0^T 1_A(\omega, t)\, d[M]_t\right), \quad A \in \bar{\mathcal{F}}. \tag{5.8}$$

With this definition we have for $H \in \xi$

$$\|H\|_{\mathcal{L}^2(\bar{\Omega},\bar{\mathcal{F}},\bar{P}_M)} = \bar{E}_M\left(H^2\right) = E\left(\int_0^T H_s^2\, d[M]_s\right)$$

i.e. $E\left(\int_0^T H_s^2\, d[M]_s\right)$ is the $\mathcal{L}^2$-norm of H (regarded as a random variable on $\bar{\Omega}$, $\bar{\mathcal{F}}$) with respect to the measure $\bar{P}_M$. By Corollary 5.8 the mapping

$$I : \xi \to \mathcal{M}^{2,c}, \quad H \mapsto (H.M)$$

is therefore an isometry from $(\xi, \|\cdot\|_{\mathcal{L}^2(\bar{\Omega},\bar{\mathcal{F}},\bar{P}_M)})$ to $(\mathcal{M}^{2,c}, \|\cdot\|_{\mathcal{M}^2})$.

Consider now $H \in \overline{\xi}$. By Statement (i) there exists a sequence $H^n \in \xi$ with $H^n \to H$ in $\mathcal{L}^2(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P}_M)$; in particular the sequence H^n is Cauchy. By the Itô-isometry the sequence $I(H^n) = (H^n.M)$ is therefore Cauchy in $\mathcal{M}^{2,c}$. Since $\mathcal{M}^{2,c}$ is a Hilbert space (Lemma 5.2), the limit $\lim_{n\to\infty}(H^n.M) =: (H.M)$ exists in $\mathcal{M}^{2,c}$ □

Corollary 5.16. *Let $M \in \mathcal{M}^{2,c}$, $H_1, H_2 \in \overline{\xi}$ and $\alpha_1, \alpha_2 \in \mathbb{R}$. Then we have*

- *Linearity of $(H.M)$: $(\alpha_1 H^1 + \alpha_2 H^2 \,.\, M) = \alpha_1(H^1 \,.\, M) + \alpha_2(H^2 \,.\, M)$ and*
- *Itô-isometry: $E\left[(H.M)_t^2\right] = E\left[\int_0^t H_s^2\, d[M]_s\right]$, for $t \le T$.*

Extensions. The definition of the stochastic integral $(H.M)$ can be extended in a number of ways.

a) Localisation. If H is adapted, leftcontinuous and M is a continuous local martingale with $\int_0^T H_s^2\, d[M]_s < \infty$ P-a.s. we may find an increasing sequence of stopping times $\tau_n \nearrow T$ such that $M^{\tau_n} \in \mathcal{M}^{2,c}$, $E\left[\int_0^T H_s^2\, d[M^{\tau_n}]_s\right] < \infty$. Then, for $t < \tau_n$ we put

$$(H.M)_t = (H.M^{\tau_n})_t, \tag{5.9}$$

where the right hand side is defined by Definition **??**. It is easily shown that (5.9) gives a consistent definition of $(H.M)_t$, and that $(H.M)_t$ is a continuous local martingale.

b) Semimartingales as integrators. Let $X = X_0 + M + A$, where M is a continuous local martingale and A a continuous FV-process. Then one defines for H adapted and left continuous such that $\int_0^T H_s^2\, d[M]_s + \int_0^T |H_s|\, d|A|s < \infty$,

$$\int_0^t H_s\, dX_s = \int_0^t H_s\, dM_s + \int_0^t H_s\, dA_s,$$

where the integral with respect to M is constructed as before and $\int_0^t H_s\, dA_s$ is an ordinary Stieltjes integral.

5.1.4 Kunita-Watanabe characterization

Theorem 5.17 (Kunita-Watanabe characterization)**.** *Let $H \in \overline{\xi}$, $M \in \mathcal{M}^{2,c}$. Then we have for all $N \in \mathcal{M}^{2,c}$*

$$[(H.M), N]_t = \int_0^t H_s\, d[M,N]_s. \tag{5.10}$$

Conversely if $L \in \mathcal{M}^{2,c}$ satisfies $[L,N]_t = \int_0^t H_s\, d[M,N]_s$, $t \le T$ for all $N \in \mathcal{M}^{2,c}$, we have $L = (H.M)$.

Partial proof. Establishing (5.10) for $H \in \overline{\xi}$ is a technical approximation argument and will be omitted. Uniqueness of a process L that satisfies (5.10) is easy to show. (5.10) implies that

$$E\left[L_T N_T\right] = E\left[[L,N]_T\right] = E\left[\int_0^T H_s\, d[M,N]_s\right] \quad \text{for all } N \in \mathcal{M}^{2,c},$$

i.e. the scalar product $[L,N]_{\mathcal{M}^2}$ is determined from (5.10) for all $N \in \mathcal{M}^{2,c}$, , so that L is uniquely determined. □

Further Properties of $(H.M)$. Theorem 5.17 allows us to establish a number of very useful properties of the stochastic integral $(H.M)$.

Quadratic covariation. If we put $N = M$ we get

$$[(H.M), M]_t = \int_0^t H_s\, d[M,M]_s = \int_0^t H_s\, d[M]_s,$$

and hence

$$\begin{aligned} [(H.M)]_t &= [(H.M),(H.M)]_t = \int_0^t H_s\, d[H.M, M] \\ &= \int_0^t H_s\, d\big(\int_0^s H_u\, d[M]_u\big) \qquad (5.11) \\ &= \int_0^t H_s^2\, d[M]_s, \end{aligned}$$

by the chain-rule of Stieltjes-calculus, i.e. we have shown that $[(H.M)]_t = \int_0^t H_s^2\, d[M]_s$ for $H \in \overline{\xi}$ (and not just in ξ).

Chain rule of stochastic integration. Consider two integrands $H_1, H_2 \in \overline{\xi}$ such that $E\left[\int_0^T (H_{1,s}H_{2,s})^2\, d[M]_s\right] < \infty$. Then we have $(H_1\,.\,(H_2\,.\,M)) = (H_1 H_2\,.\,M)$ or, in the long version,

$$\int_0^t H_{1,s}\, d(H_2\,.\,M)_s = \int_0^t H_{1,s} H_{2,s}\, dM_s. \qquad (5.12)$$

Relation (5.12) is the *chain rule* of stochastic integration.

Proof. Let $L = H_1\,.\,(H_2\,.\,M)$. By Theorem 5.17 we get

$$\begin{aligned} [L,N]_t &= \int_0^t H_{1,s}\, d[(H_2\,.\,M), N]_s = \int_0^t H_{1,s}\, d\int_0^s H_{2,u}\, d[M,N]_u \\ &= \int_0^t H_{1,s}\, H_{2,s}\, d[M,N]_s = [(H_1 H_2\,.\,M), N]_t. \end{aligned}$$

Hence $L_t = \int_0^t H_{1,s}\, H_{2,s}\, dM_s$ as claimed. □

The next result is important in the context of incomplete markets.

Theorem 5.18 (Kunita Watanabe decomposition). *Consider martingales $N, M_1, \dots, M_n \in \mathcal{M}^{2,c}$. Then there is a unique decomposition of the form*

$$N_t = N_0 + \sum_{i=1}^{n} \int_0^t H_{s,i} dM_{s,i} + L_t, \quad t \le T,$$

with $H_1, \dots, H_n \in \overline{\xi}$ and $L \in \mathcal{M}^{2,c}$ strongly orthogonal to $M_1, \dots, M_n$, that is with $[L, M_i] \equiv 0$ for all $1 \le i \le n$.

Note that L is also weakly orthogonal to $M_1, \dots, M_n$, that is $E(L_T M_{T,i}) = L_0 M_{0,i} = 0$ as LM_i is a martingale.

Sketch of the proof. We let $n = 1$. By the Itô-isometry, the space of stochastic integrals $\mathcal{M}^H = \{(H.M) \colon H \in \overline{\xi}\}$ is a closed subspace of $\mathcal{M}^{2,c}$. Hence we can project the rv $N_T - N_0$ on that space, that is there is a representation

$$N_T = N_0 + \int_0^T H_s dM_s + L_T$$

and $E(L \int_0^T H_s dM_s) = 0$ for all $H \in \overline{\xi}$. Define the martingale L via $L_t = E(L_T \mid \mathcal{F}_t)$. Then L is the desired martingale. It remains to show that $[L, M] \equiv 0$. Here one uses that $\mathcal{M}^H$ is *stable under stopping*, i.e. $\tilde{M} \in \mathcal{M}^H$ implies that also the stopped martingale $\tilde{M}^\tau \in \mathcal{M}^H$ for an arbitrary stopping time τ; we omit the details.

□

5.2 Itô Processes and the Feynman-Kac formula

Itô processes. Itô processes are solutions of stochastic differential equations driven by Brownian motion; they will be our basic model for asset price dynamics.

Definition 5.19. Given a d-dimensional Brownian motion $W = (W_{t,1}, ..., W_{t,d})_{t \ge 0}$, a time point $t_0 \ge 0$, some vector $x \in \mathbb{R}^n$ and functions $\mu : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^{n \times d}$. Then the n-dimensional process $X = (X_{t,1}, ..., X_{t,n})_{t \ge 0}$ is called an *Itô process* with initial value t_0, x, drift μ and dispersion matrix σ if X satisfies the SDE

$$X_{t,i} = x_i + \int_{t_0}^t \mu_i(s, X_s)\, ds + \sum_{j=1}^{d} \int_{t_0}^t \sigma_{ij}(s, X_s)\, dW_{sj}\,, \quad t \ge t_0\,. \tag{5.13}$$

In short notation (5.13) is often written in the form $dX_t = \mu(t, X_t)\, dt + \sigma(t, X_t)\, dW_t$. Moreover, one frequently takes $t_0 = 0$.

An intuitive way to understand the SDE (5.13) is the *Euler approximation*. For a small time step Δt one has with $t_n = n\Delta t$

$$X_{t_{n+1}} = X_{t_n} + \mu(t_n, X_{t_n})\Delta t + \sum_{j=1}^{d} \sigma_j(t_n, X_{t_n})\epsilon_{n,j} \tag{5.14}$$

where $\epsilon_{n,j}$ are iid $\sim N(0, \Delta t)$. (5.14) can be used to generate (approximations of) the trajectories of X on a computer.

Define the $n \times n$ matrix $C(t,X) = \sigma(t,X)\sigma'(t,X)$. Then we have

$$\begin{aligned}
[X_i, X_j]_t &= [\sum_{k=1}^{d} \int_0^{\cdot} \sigma_{ik}\, dW_{s,k}\,, \sum_{l=1}^{d} \int_0^{\cdot} \sigma_{jl}\, dW_{s,l}\,]_t \\
&= \sum_{k,l=1}^{d} \int_0^t \sigma_{ik}\sigma_{jl}\, d\underbrace{[W_k, W_l]_s}_{=\delta_{kl}\, s} \\
&= \sum_{k=1}^{d} \int_0^t \sigma_{ik}\sigma_{jk}\, ds \\
&= \int_0^t C_{ij}(s, X_s)\, ds.
\end{aligned}$$

Hence C is sometimes called instantaneous covariance matrix of X.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function and define the differential operator A by

$$\mathcal{A}f(t,x) = \sum_{i=1}^{n} \mu_i(t,x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j=1}^{n} C_{ij}(t,x) \frac{\partial^2 f}{\partial x_i \partial x_j}(t,x). \tag{5.15}$$

$\mathcal{A}$ is known as *generator* of the Ito process X. We have

Lemma 5.20. *For $f : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$ smooth such that f its derivatives are bounded the process $M_t^f = f(t, X_t) - \int_0^t \frac{\partial f}{\partial t}(s, X_s) + \mathcal{A}f(s, X_s)\, ds$ is a martingale.*

Proof. For simplicity we consider only the time-independent case. Applying Itô's formula in n dimensions we get

$$\begin{aligned}
f(X_t) &= f(X_0) + \sum_{i=1}^{n} \int_0^t \frac{\partial f}{\partial x_i}(X_s)\, dX_{s,i} + \frac{1}{2} \sum_{i,j=1}^{n} \int_0^t \frac{\partial^2 f}{\partial x_i x_j}(X_s)\, d[X_i, X_j]_s \\
&= f(X_0) + \sum_{i=1}^{n} \int_0^t \frac{\partial f}{\partial x_i}(X_s) \mu_i(s, X_s)\, ds + \sum_{i=1}^{n} \sum_{j=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_s) \sigma_{ij}(s, X_s)\, dW_{s,j} \\
&\quad + \frac{1}{2} \sum_{i,j=1}^{n} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j} C_{ij}(s, X_s)\, ds \\
&= f(X_0) + \sum_{i=1}^{n} \sum_{j=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_s) \sigma_i(s, X_s)\, dW_{s,j} + \int_0^t Af(s, X_s) ds.
\end{aligned}$$

Hence $f(X_t) - \int_0^t Af(X_s)\, ds$ can be represented as a sum of stochastic integrals wrt Brownian assumed motion, and is therefore a local martingale. Since f and its derivatives are bounded by assumption, this expression is a true martingale and the claim follows. ☐

In the following theorem we give conditions ensuring that a solution to the SDE (5.13) exists.

Theorem 5.21. *Suppose that μ, σ satisfy Lipschitz and growth conditions of the form*

$$\begin{aligned}
\|\mu(t,X) - \mu(t,Y)\| &\leq K\|X - Y\|, \quad t \geq 0, \\
\|\sigma(t,X) - \sigma(t,Y)\| &\leq K\|X - Y\|, \quad t \geq 0, \\
\|\mu(t,X) + \sigma(t,X)\| &\leq K(1 + \|X\|), \quad t \geq 0.
\end{aligned}$$

Then for all initial values $(t_0, x) \in [0,\infty)\times\mathbb{R}^n$ a unique solution to (5.13) exists. Moreover, this solution is $(\mathcal{F}_t^W)$-adapted, where $\mathcal{F}_t^W = \sigma(W_{s,i} : 1 \le i \le d,\ s \le t)$.

The Feynman Kac formula. Lemma 5.20 forms the basis for a close interplay between stochastic processes and solution of parabolic PDEs, which is extremely fruitful in financial mathematics. The basic result is the celebrated Feynman-Kac formula.

Consider an n dimensional Ito process X with drift vector $\mu(t,x)$, dispersion matrix $\sigma(t,x)$, instantaneous covariance matrix $C(t,x) = \sigma\sigma'(t,x)$ and recall that for $F\colon [0,T]\times\mathbb{R}^n \to \mathbb{R}$ smooth the generator of X is given by

$$\mathcal{A}F(t,x) = \sum_{i=1}^n \mu_i(t,x)\frac{\partial F}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^n C_{ij}(t,x)\frac{\partial^2 F}{\partial x_i \partial x_j}(t,x).$$

Given functions $\phi\colon \mathbb{R}^n \to \mathbb{R}$ (the payoff) and $r\colon [0,T]\times\mathbb{R}^n \to \mathbb{R}$ (the interest rate), suppose that F is a solution to the terminal value problem

$$\frac{\partial F}{\partial t}(t,x) + \mathcal{A}F(t,x) = r(t,x)F(t,x), \quad F(T,x) = \phi(x). \tag{5.16}$$

Then we can express F as an expectation involving the process X as we now show. Consider the process $Z_t = \exp(-\int_{t_0}^t r(s,X_s)ds)F(t,X_t)$ and let $D_t = \exp(-\int_{t_0}^t r(s,X_s)ds)$. An application of the Ito formula yields that

$$dZ_t = d(D_t F(t,X_t)) = F(t,X_t)dD_t + D_t dF(t,X_t) + d[D, F(t,X)]_t.$$

Now D is an FV-process which solves the ODE $dD_t = -r(t,X_t)D_t dt$, so that the covariation $[D, F(t,X)]_t \equiv 0$. Moreover, since $\mathcal{A}$ is the generator of X we have that

$$dF(t,X_t) = (\frac{\partial F}{\partial t} + \mathcal{A}F)(t,X_t)dt + dM_t$$

for some local martingale M. Hence we get, using the chain rule for stochastic integration

$$dZ_t = -\, r(t,X_t)F(t,X_t)D_t dt + \Big(\frac{\partial F}{\partial t} + \mathcal{A}F\Big)(t,X_t)D_t dt + D_t dM_t.$$

Using the PDE (5.16) for F we see that the dt terms vanish. Hence Z is a (local) martingale and a true martingale given sufficient integrality, which we assume form now on. Moreover, by definition $D_{t_0} = 1$ so that $Z_{t_0} = F(t_0, X_{t_0})$. The martingale property of Z now gives, using $Z_T = \exp\big(-\int_{t_0}^T r(s,X_s)ds\big)\phi(X_T)$ and the terminal condition $F(T, x_T) = \phi(X_T)$

$$F(t_0, X_{t_0}) = Z_{t_0} = E_{t_0,x}(Z_T) = E_{t_0,x}\Big(\exp\big(-\int_{t_0}^T r(s,X_s)ds\big)\phi(X_T)\Big). \tag{5.17}$$

Formula (5.17) is called *Feynman-Kac formula.* It can be used in two ways:

- We can use probabilistic techniques or Monte-Carlo simulation to compute the expectation on the rhs of (5.17) in order to solve numerically the terminal value problem (5.16).
- We can try to solve the terminal value problem (5.16) perhaps numerically, in order to compute the expectation on the rhs of (5.17).

Both approaches are frequently used in financial mathematics. For a generalization of (5.17) to multi-dimensional processes and for a precise statement of the necessary integrality conditions we refer to Section 5.7 of Karatzas & Shreve (1988).

Example 5.22 (The Black-Scholes PDE)**.** In Theorem 4.2 we showed that the fair price $V_t = \mu(t, S_t)$ of a terminal-value claim with payoff $h(S_T)$ in the Black-Scholes Model solves the terminal value problem

$$u_t + rSu_S + \frac{1}{2}\sigma^2 S^2 u_{SS} = ru, \quad u(T, S) = h(S).$$

In order to give a probabilistic interpretation we consider the SDE

$$dS_t = rS_t dt + \sigma S_t dW_t$$

with generator $\mathcal{A} = rS\frac{\partial}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2}$, and we obtain from (5.17)

$$u(t_0, S) = E_{t_0,S}\big(e^{-r(T-t_0)}h(S_T)\big), \quad t_0 \leq T. \tag{5.18}$$

This can be viewed as risk-neutral pricing formula in continuous time. In the next Chapter we give a derivation of (5.18) using only probabilistic techniques.

5.3 Change of Measure and Girsanov Theorem for Brownian motion

5.3.1 Motivation

In discrete-time models, it was shown that the price in $t < T$ of any attainable claim with $\mathcal{F}_T$-measurable payoff H is given by the risk-neutral pricing formula

$$H_t = S_{t,0}\, E^Q\left(\frac{H}{S_{T,0}} \mid \mathcal{F}_t\right), \quad t \leq T. \tag{5.19}$$

Here $S_{t,0} > 0$ represents the numeraire at time t (often $S_{t,0} = \exp(rt)$), and Q satisfies the following properties:

(i) $Q \sim P$, i.e. $Q(A) > 0 \Leftrightarrow P(A) > 0, \quad A \in \mathcal{F}_T$.

(ii) The discounted price processes $\tilde{S}_{t,i} = S_{t,i}/S_{t,0}$, $1 \leq i \leq n$, are Q-martingales, i.e. $E^Q\left(\tilde{S}_{t,i}|\mathcal{F}_s\right) = \tilde{S}_{s,i}\,, \; 0 \leq s \leq t \leq T$.

In this section we provide the mathematical tools for extending (5.19) to continuous-time models driven by Brownian motion.

By the Radon-Nikodym theorem, the equivalence of P and Q implies the existence of an $\mathcal{F}_T$-measurable rv Z with $P(Z > 0) = 1$ such that for all $A \in \mathcal{F}_T$

$$Q(A) = E^P\left(Z1_A\right) = E^P\left(Z; A\right). \tag{5.20}$$

In particular, we get $E^P(Z) = Q(\Omega) = 1$. Conversely, any $\mathcal{F}_T$-measurable rv Z with $E(Z) = 1$ and $P(Z > 0) = 1$ can be used to define a measure Q on $\mathcal{F}_T$ via (5.20); P and Q are obviously equivalent. Z is called the (Radon-Nikodym) density of Q wrt P, denoted by $Z = \frac{dQ}{dP}$. Note that (5.20) implies that for every $\mathcal{F}_T$-measurable $X \geq 0$ we have $E^Q(X) = E^P(ZX)$ (by measure theoretic induction).

Example 5.23. For finite $\Omega = \{\omega_1, ..., \omega_k\}$ two measures P and Q are equivalent if for $1 \leq i \leq k$: $P(\{\omega_i\}) > 0 \Leftrightarrow Q(\{\omega_i\}) > 0$; it is immediate from (5.20) that $\frac{dQ}{dP}(\omega) = Q(\{\omega\})/P(\{\omega\})$ in that case.

In the next Example we show how a change-of-measure can be used to alter the mean of a normally distributed random variable.

Example 5.24. Let $X \sim N(0, \sigma^2)$ on some $(\Omega, \mathcal{F}, P)$. Define a random variable Z by $Z = \exp\left(\frac{\mu}{\sigma^2}X - \frac{1}{2}\frac{\mu^2}{\sigma^2}\right)$. Then we get, using standard rules for the mean of lognormal random variables

$$\begin{aligned} E(Z) &= \exp\left(-\frac{1}{2}\frac{\mu^2}{\sigma^2}\right)\exp\left(E\left(\frac{\mu}{\sigma^2}X\right) + \frac{1}{2}\mathrm{var}\left(\frac{\mu}{\sigma^2}X\right)\right) \\ &= \exp\left(-\frac{1}{2}\frac{\mu^2}{\sigma^2}\right)\exp\left(\frac{1}{2}\frac{\mu^2}{\sigma^2}\right) = 1, \end{aligned}$$

so that we can define a measure Q by putting $\frac{dQ}{dP} = Z$. Note that we have for any bounded measurable $g : \mathbb{R} \to \mathbb{R}$:

$$\begin{aligned} E^Q(g(X)) &= E^P(Zg(x)) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} g(x)\exp\left(\frac{\mu}{\sigma^2}x - \frac{1}{2}\frac{\mu^2}{\sigma^2}\right)e^{-\frac{x^2}{2\sigma^2}}\,dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} g(x)\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx, \end{aligned}$$

which shows that $X \sim N(\mu, \sigma^2)$ under Q.

The main result of this section, Girsanov's theorem for Brownian motion, can be thought of as the infinite-dimensional analogue of this example: by a proper change of measure it is possible to alter the drift of an Itô process. In particular, by choosing $Q \sim P$ properly, it is usually possible to turn an asset price into a martingale.

5.3.2 Density martingales

Now we return to the change of measure for stochastic processes. Suppose that we have a filtered probability space $(\Omega, \mathcal{F}, P)$, $(\mathcal{F}_t)_{0\leq t\leq T}$, and a strictly positive, $\mathcal{F}_T$-measurable random variable Z with $E(Z) = 1$, and define the measure Q by $E^Q(X) = E^P(XZ)$, X bounded, $\mathcal{F}_T$-measurable. Define the associated density martingale by

$$Z_t = E^P(Z|\mathcal{F}_t), \quad 0 \leq t \leq T. \tag{5.21}$$

Lemma 5.25. *$(Z_t)_{0\leq t\leq T}$ is a martingale, and for every $\mathcal{F}_t$-measurable random variable Y we have $E^Q(Y) = E^P(YZ_t)$, $t \leq T$.*

Proof. The martingale property of $(Z_t)_{t\geq 0}$ is obvious. For the second claim note that by iterated conditioning

$$E^Q(Y) = E^P(YZ) = E^P\left(E^P(YZ|\mathcal{F}_t)\right) = E^P\left(YE^P(Z|\mathcal{F}_t)\right) = E^P(YZ_t).$$

□

The next lemma extends this result to conditional expectations.

Lemma 5.26 (abstract Bayes formula)**.** *Given $0 \le s \le t \le T$ and let Y be some $\mathcal{F}_t$-measurable integrable random variable. Then*

$$E^Q\left(Y|\mathcal{F}_s\right) = \frac{1}{Z_s} E^P\left(Y Z_t|\mathcal{F}_s\right) \tag{5.22}$$

Proof. We have to check that the right hand side of (5.22) satisfies the characterizing equation of conditional expectations, i.e.

$$E^Q\left(\frac{1}{Z_s} E^P\left(Y Z_t|\mathcal{F}_s\right) 1_A\right) = E^Q\left(Y 1_A\right), \quad A \in \mathcal{F}_s.$$

Using Lemma 5.25, the left hand side equals, as $A \in \mathcal{F}_s$,

$$E^P\left(Z_s \frac{1}{Z_s} E^P\left(Y Z_t|\mathcal{F}_s\right) 1_A\right) = E^P\left(E^P\left(Y Z_t 1_A|\mathcal{F}_s\right)\right) = E^P\left(Y Z_t 1_A\right),$$

which is equal to $E^Q(Y 1_A)$ by Lemma 5.25. □

Lemma 5.27. *Let $Q \sim P$ with $\frac{dQ}{dP} = Z$. An adapted process $(M_t)_{0\le t\le T}$ is a Q-martingale if and only if the process $(M_t Z_t)_{0\le t\le T}$ is a P-martingale.*

Proof. By the Bayes formula (5.22) we have for $t \le T$ with $Z_T := Z$

$$E^Q\left(M_T|\mathcal{F}_t\right) = \frac{1}{Z_t} E^P\left(M_T Z_T|\mathcal{F}_t\right). \tag{5.23}$$

If $(M_t Z_t)_{0\le t\le T}$ is a P-martingale the right hand side equals $\frac{1}{Z_t} M_t Z_t = M_t$ and M is a Q-martingale. Conversely, if M is a Q-martingale the left hand side equals M_t and multiplying (5.23) with Z_t gives $M_t Z_t = E^P\left(M_T Z_T|\mathcal{F}_t\right)$. □

5.3.3 The Girsanov Theorem

We begin with the one-dimensional version. Let $(W_t)_{0\le t\le T}$ be a one-dimensional Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, P)$, $(\mathcal{F}_t)_{0\le t\le T}$. Let $(\theta_t)_{0\le t\le T}$ be an adapted process, and define

$$Z_t = \exp\left(\int_0^t \theta_s\, dW_s - \frac{1}{2}\int_0^t \theta_s^2\, ds\right), \quad 0 \le t \le T. \tag{5.24}$$

We show that Z_t can be written as a stochastic integral with respect to W. Define $X_t = \int_0^t \theta_s\, dW_s - \frac{1}{2}\int_0^t \theta_s^2\, ds$. Then

$$[X]_t = [\int_0^\cdot \theta_s\, dW_s]_t = \int_0^t \theta_s^2\, ds,$$

and since $Z_t = \exp\left(X_t\right)$, Itô's formula gives with $f(x) = e^x$:

$$\begin{aligned} dZ_t &= f'(X_t)\, dX_t + \frac{1}{2} f''(X_t)\, d[X]_t \\ &= e^{X_t}\, dX_t + \frac{1}{2} e^{X_t} \theta_t^2\, dt \\ &= Z_t\, dX_t + \frac{1}{2}\theta_t^2 Z_t\, dt \\ &= \theta_t Z_t\, dW_t - \frac{1}{2}\theta_t^2 Z_t\, dt + \frac{1}{2}\theta_t^2 Z_t\, dt, \end{aligned}$$

so that $Z_t = Z_0 + \int_0^t \theta_s Z_s \, dW_s$. It follows that Z is a nonnegative local martingale and therefore a supermartingale by a version of the Fatou lemma. Hence Z is a true martingale if and only if the mapping $t \mapsto E(Z_t)$ is non-decreasing, that is for $E(Z_T) = Z_0 = 1$. In that case Z can be used as a Radon-Nikodym density.

Theorem 5.28 (Girsanov). *Suppose that $(Z_t)_{0\le t\le T}$ defined in (5.24) is a martingale. Define a probability measure Q on $\mathcal{F}_T$ by putting $\frac{dQ}{dP} = Z_T$. Then the process $\tilde{W}$ defined by*

$$\tilde{W}_t = W_t - \int_0^t \theta_s \, ds, \quad 0 \le t \le T \tag{5.25}$$

is a Brownian motion under Q.

Remark 5.29. The process $W_t = \tilde{W}_t + \int_0^t \theta_s \, ds$ is under Q a Brownian motion with drift $(\theta_t)_{0\le t\le T}$; in that sense Theorem 5.28 generalises Example 5.24.

Proof. The proof is based on the Levy-characterization of Brownian motion:

Theorem 5.30 (Levy). *Given a filtered probability space $(\Omega, \mathcal{F}, P)$, $(\mathcal{F}_t)_{t\ge 0}$. An adapted process $M = (M_t)_{t\ge 0}$ with continuous trajectories is a one-dimensional Brownian motion if and only if the following conditions hold:*

1. *M is a local martingale.*
2. *$[M]_t = t$, where $[M]$ is the pathwise quadratic variation defined for instance in (5.3).*

Condition 2 is easily verified: Since $\int_0^t \theta_s \, ds$ is FV, we have $[W]_t = [\tilde{W}]_t$, and since W is a P-Brownian motion, we have $[W]_t = t$ P-a.s. and hence also Q-a.s. (as $P \sim Q$). The theorem is proven if we can show that $\tilde{W}$ is a Q-local martingale. By Lemma 5.27 this is equivalent to showing that $(\tilde{W}_t Z_t)_{0\le t\le T}$ ia a P-local martingale. Recall that Z has the representation $Z_t = Z_0 + \int_0^t \theta_s Z_s \, dW_s$. Hence we get, using Itô's product formula

$$d\tilde{W}_t Z_t = \tilde{W}_t \, dZ_t + Z_t \, d\tilde{W}_t + d[\tilde{W}, Z]_t.$$

Now we have

$$\begin{aligned} [\tilde{W}, Z]_t &\overset{(i)}{=} [W, Z]_t = [W, \int_0^\cdot \theta_s Z_s \, dW_s]_t \\ &\overset{(ii)}{=} \int_0^t \theta_s Z_s \, d[W, W]_s \\ &= \int_0^t \theta_s Z_s \, ds, \end{aligned} \tag{5.26}$$

using (i) that W and $\tilde{W}$ differ only by a continuous FV process and (ii) the Kunita-Watanabe characterization (Theorem 5.17). Hence we get, using (5.26) and the chain-rule for stochastic integrals (5.13),

$$\begin{aligned} d\tilde{W}_t Z_t &= \tilde{W}_t \theta_t Z_t \, dW_t + Z_t \, dW_t - Z_t \theta_t \, dt + Z_t \theta_t \, dt \\ &= (\tilde{W}_t \theta_t Z_t + Z_t) \, dW_t. \end{aligned}$$

This shows that the product $\tilde{W}_t Z_t$ has a representation as an integral with respect to the P-Brownian motion $(W_t)_{0\le t\le T}$ and is therefore a P-local martingale, proving the result. ☐

Theorem 5.28 is easily extended to the case of a d-dimensional Brownian motion.

Theorem 5.31 (Girsanov, d-dimensional version.). *Consider a d-dimensional Bownian motion on $(\Omega, \mathcal{F}, P)$, $(\mathcal{F}_t)_{0\le t\le T}$ and some adapted process $\theta = (\theta_{t,1}, \ldots, \theta_{t,d})'$. Define*

$$Z_t = \exp\Big(\sum_{i=1}^{d}\int_0^t \theta_{s,i} dW_{s,i} - \frac{1}{2}\int_0^t \|\theta_s\|^2\, ds\Big).$$

Suppose that Z is a martingale (and not only a local martingale) so that $E(Z_T) = 1$. Define a probability measure Q on $\mathcal{F}_T$ by putting $\frac{dQ}{dP} = Z_T$. Then the process $\tilde{W}$ defined by

$$\tilde{W}_{t,i} = W_{t,i} - \int_0^t \theta_{s,i}\, ds, \quad 0 \le t \le T \tag{5.27}$$

is a Brownian motion under Q.

Remark 5.32. The following conditions are sufficient to ensure that $(Z_t)_{0\le t\le T}$ is a martingale, so that Theorem 5.28 respectively Theorem 5.31 applies.

(a) $E\left(\int_0^T \|\theta_s\|^2 Z_s^2\, ds\right) < \infty$ (integrable quadratic variation).

(b) $E(Z_T) = 1$. (This is necessary and sufficient.)

(c) Novikov-criterion: $E\left(\exp\left(\frac{1}{2}\int_0^T \|\theta_s\|^2\, ds\right)\right) < \infty$.

We discuss two mathematical applications of Theorem 5.28.

1. Solution of SDEs via change of measure. Given some function $b(s, X) : \mathbb{R}^+ \to \mathbb{R}$, bounded (but not Lipschitz). We want to construct a solution X of the SDE

$$dX_t = b(t, X_t)\, dt + dW_t, \quad 0 \le t \le T, \tag{5.28}$$

with W some Brownian motion. Start with some probability space $(\Omega, \mathcal{F}, P)$, $(\mathcal{F}_t)_{0\le t\le T}$, supporting a Brownian motion $(X_t)_{0\le t\le T}$. Define a measure Q via

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_T} = Z_T := \exp\left(\int_0^T \theta(s, X_s)\, dX_s - \frac{1}{2}\int_0^T \theta^2(s, X_s)\, ds\right).$$

By Theorem 5.28, $W_t := X_t - \int_0^t \theta(s, X_s)\, ds$ is a Q-Brownian motion; hence

$$X_t = W_t + \int_0^t \theta(s, X_s)\, ds$$

solves equation (5.28) for a Q-Brownian motion W.

2. Maximum-likelihood estimation for a Brownian motion with drift. We observe a process $X_t = W_t + \mu t$, $\mu \in \mathbb{R}$ a constant and W a Brownian motion and want to estimate the unknown parameter μ. A possible approach is maximum-likelihood estimation (MLE):

Consider some 'random variable' X with values in the space $(\tilde{\Omega}, \tilde{\mathcal{F}})$ whose distribution P^μ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ depends on an unknown parameter $\mu \in A^\mu$. Suppose moreover, that there is a common reference measure $\tilde{P}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ such that $P^\mu \sim \tilde{P}$ and

$$\frac{dP^\mu}{d\tilde{P}} = L(X;\mu), \quad \mu \in A^\mu.$$

Given some observation $\hat{X}$ of X, the (abstract) ML-estimator $\hat{\mu}$ for μ is given by

$$\hat{\mu} = \arg\max_{\mu \in A^\mu} L(\hat{X}, \mu), \tag{5.29}$$

i.e. $\hat{\mu}$ is the value of μ that makes the observation $\hat{X}$ most likely. ML-estimators typically have desirable properties such as asymptotic normality and consistency.

Let us apply the approach to the Brownian motion with drift X. In that case we choose $\tilde{\Omega} = C([0,T], \mathbb{R})$, $\tilde{P}$ equal to the Wiener measure, and we let $X_t(\tilde{\omega}) = \tilde{\omega}(t)$ (the so-called coordinate process on $C([0,T], \mathbb{R})$). Moreover, we define P^μ by

$$\frac{dP^\mu}{d\tilde{P}} = \exp\left(\mu X_T - \frac{1}{2}\mu^2 T\right), \quad \mu \in \mathbb{R}, \tag{5.30}$$

By definition of the Wiener measure, X is a $\tilde{P}$ Brownian motion. Moreover, as in the previous application, $X_t - \mu t$ is a P^μ-Brownian motion, so that under P^μ $X_t = (X_t - \mu t) + \mu t$ is a Brownian motion with drift μ. Given an observed trajectory $(\hat{X}_t)_{0 \le t \le T}$ of X, the ML-estimator is hence given by

$$\begin{aligned} \hat{\mu} &= \arg\max_{\mu \in \mathbb{R}} \exp\left(\mu \hat{X}_T - \frac{1}{2}\mu^2 T\right) \\ &= \arg\max_{\mu \in \mathbb{R}} \mu \hat{X}_T - \frac{1}{2}\mu^2 T. \end{aligned} \tag{5.31}$$

Differentiating (5.31) with respect to μ gives $\hat{X}_T - \hat{\mu} T = 0$ and hence the ML-estimator $\hat{\mu} = \frac{\hat{X}_T}{T}$.

Chapter 6

Financial Mathematics in Continuous-Time

6.1 Basic Concepts

6.1.1 The Model

We start with a general model for a frictionless security market with continuous trading. Fix some horizon date T and consider some probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\{\mathcal{F}_t\}_{0 \le t \le T}$. There are $d+1$ traded assets with price process $S = (S^0, \dots, S^d)$. We assume throughout that $S^0, \dots, S^d$ follow semimartingales with continuous trajectories. We assume that $S^0(t) > 0$ P a.s. for all t and use S^0 as numeraire; often $S^0(t) = \exp(\int_0^t r_s ds)$ for some adapted process r with $r_s \ge 0$. In that case S^0 represents the so-called *savings account* and r is the *short rate* of interest.

SELFFINANCING TRADING STRATEGIES: A $\mathbb{R}^{d+1}$-valued adapted and left-continuous process $\phi(t) = \left(\phi_t^0, \dots, \phi_t^d\right)$ is called a trading strategy; ϕ_t^i represents the number of units of security i in the portfolio at time t. We assume that ϕ^i is sufficiently integrable so that the stochastic integral $\int_0^t \phi_s^i dS_s^i$ is well defined.

Definition 6.1. (i) The *value* of the portfolio at time t equals $V_t = V_t^\phi = \sum_{i=0}^d \phi_t^i S_t^i$; $V^\phi = (V_t^\phi)_{0 \le t \le T}$ is called *value process* of the strategy.

(ii) The *gains process* of the strategy is given by

$$G_t^\phi = \int_0^t \phi_u dS_u := \sum_{i=0}^d \int_0^t \phi_u^i dS_u^i$$

(iii) The strategy is called *selffinancing*, if for all $t \in [0,T]$ $V_t^\phi = V_0^\phi + G_t^\phi$.

The economic justification of (ii) and (iii) is the same as in the case of the Markov strategies discussed in Chapter 4.

NUMERAIRE INVARIANCE: We choose S^0 as our numeraire and introduce the discounted price process

$$\tilde{S}_t = (1, S_t^1/S_t^0, \dots, S_t^d/S_t^0), \tag{6.1}$$

and the discounted value process $\tilde{V}_t^\phi = V_t^\phi / S_t^0 = \sum_{i=0}^d \phi_t^i \tilde{S}_t^i$. Intuitively, one would expect that a change of numeraire has no substantial economic implications. The next result confirms this fact.

Proposition 6.2 (Numeraire Invariance)**.** *Selffinancing strategies remain selffinancing after a change of numeraire.*

Proof. Let $X_t = \left(S_t^0\right)^{-1}$. We have, using Itô 's product formula, that

$$d\tilde{S}_t^k = dS_t^k X_t = S_t^k dX_t + X_t dS_t^k + d[S^k, X]_t. \tag{6.2}$$

By definition we have $V_t^\phi = \sum_{k=0}^d \phi_t^k S_t^k$. Moreover, as ϕ is selffinancing, it holds that $dV_t^\phi = \sum_{k=0}^d \phi_t^k dS_t^k$. This gives the following dynamics of the discounted wealth $\tilde{V}$:

$$\begin{aligned} d\tilde{V}_t^\phi &= V_t^\phi dX_t + X_t dV_t^\phi + d[X, V^\phi]_t \\ &= \Big(\sum_{k=0}^d \phi_t^k S_t^k\Big)\, dX_t + X_t\, d\sum_{k=0}^d \phi_t^k dS_t^k + \sum_{k=0}^d \phi_t^k d[X, S^k]_t \\ &= \sum_{k=0}^d \big(\phi_t^k S_t^k dX_t + X_t \phi_t^k dS_t^k + \phi_t^k d[S^k, X]_t\big) \\ &= \sum_{k=0}^d \phi_t^k d\big(S_t^k dX_t + X_t dS_t^k + d[S^k, X]_t\big) \qquad (6.3) \\ &= \sum_{k=0}^d \phi_t^k d\tilde{S}_t^k, \end{aligned}$$

as follows by comparing (6.3) and (6.2). Hence we have shown that $d\tilde{V}_t^\phi = \sum_{k=0}^d \phi_t^k d\tilde{S}_t^k$ so that the discounted value process can be represented as sum of the discounted initial value and of the gains process wrt the discounted securities prices. □

Corollary 6.3. *(i) A trading strategy ϕ is selffinancing if and only if*

$$\tilde{V}^\phi = \tilde{V}_0^\phi + \tilde{G}_t^\phi \quad \textit{with} \quad \tilde{G}_t^\phi = \sum_{k=1}^d \int_0^t \phi_t^k d\tilde{S}_t^k. \tag{6.4}$$

(ii) A selffinancing strategy is completely determined by the initial investment V_0 and the position $\phi^1, \dots, \phi^d$ in the assets $S^1, \dots, S^d$.

Proof. Statement (i) follows immediately from the numeraire invariance (note that $\tilde{S}_t^0 \equiv 1$, so that $\int_0^t \phi_t^0 d\tilde{S}_t^0 \equiv 0$).

For the second statement note that V_0 and $\phi^1, \dots, \phi^d$ uniquely determine the discounted wealth process $\tilde{V}_t^\phi$ by (6.4) and hence the undiscounted valued process V_t^ϕ; the position ϕ_t^0 is then given by

$$\phi_t^0 = (S_t^0)^{-1}\{V_t^\phi - \sum_{k=1}^d \phi_t^k S_t^k\}.$$

□

6.1.2 Martingale Measures and Arbitrage Opportunities

Definition 6.4 (Arbitrage). A selffinancing portfolio such that $P(V_t^\phi \geq 0 \quad \text{for all} \quad 0 \leq t \leq T) = 1$, $V_0^\phi = 0$ and $P(V_T^\phi > 0) > 0$ is an arbitrage opportunity.

Arbitrage opportunities represent a chance to create risk free profits and should not exist in a well-functioning market.

Definition 6.5 (Martingale measure). A probability measure Q is called an *equivalent martingale measure* if

(i) $Q \sim P$ on $\mathcal{F}_T$, i.e. for every $A \in \mathcal{F}_T$, $Q(A) = 0 \Leftrightarrow P(A) = 0$.

(ii) The discounted price processes $\tilde{S}^k$, $1 \leq k \leq d$ are Q-local martingales.

Note that a martingale measure is always related to a given numeraire (due to the discounting in (ii)). If the numeraire is of the form $S_t^0 = \exp(\int_0^t r_s ds)$, r the spot rate of interest, Q is also called *spot martingale measure*. Often we are interested in the situation where the asset prices are true martingales (and not just local ones); this will be mentioned where appropriate.

Lemma 6.6. *Q is a spot martingale measure, if and only if every price process S^i has Q-dynamics of the form*

$$dS_t^i = r_t S_t^i dt + dM_t^i, \quad 0 \leq i \leq d, \tag{6.5}$$

where $M^0, \ldots, M^d$ are Q-local martingales.

Proof. We get, as $(S_t^0)^{-1} = \exp(-\int_0^t r_s ds)$, that $d(S_t^0)^{-1} = -r_t (S_t^0)^{-1} dt$. Itô's product formula gives, as $(S_t^0)^{-1}$ is of finite variation, that

$$d\tilde{S}_t^i = -r_t \tilde{S}_t^i dt + (S_t^0)^{-1} dS_t^i,$$

which is a local martingale if and only if S^i has dynamics (6.5). □

The next result shows that the existence of an equivalent martingale measure excludes arbitrage opportunities.

Theorem 6.7. *If the model $S^0, \ldots, S^d$ admits an equivalent martingale measure Q, there are no arbitrage opportunities.*

Proof. Step 1: For any selffinancing strategy ϕ such that $V_t^\phi \geq 0$, the discounted wealth process $\tilde{V}_t^\phi$ is a Q-supermartingale. We have $\tilde{V}_t^\phi = \tilde{V}_0^\phi + \sum_{k=1}^d \int_0^t \phi_s^k d\tilde{S}_s^k$. The discounted price processes are Q-martingales by assumption, hence $\tilde{V}_t^\phi$ is a local Q-martingale, as it is a sum of stochastic integral wrt local martingales. Since $\tilde{V}_t^\phi \geq 0$, it follows that $\tilde{V}^\phi$ is a Q-supermartingale, using that every nonnegative local martingale is a supermartingale by the Fatou-Lemma for conditional expectations.

Step 2: Suppose now that $P(V_T^\phi > 0) > 0$. It follows that $Q(\tilde{V}_T^\phi > 0) > 0$, as $Q \sim P$. Since $\tilde{V}^\phi$ is a Q-supermartingale we get $\tilde{V}_0^\phi \geq E^Q(\tilde{V}_T^\phi)$, and this is strictly positive as $Q(V_T^\phi > 0) > 0$. Hence also $V_0^\phi > 0$, so that an arbitrage opportunity cannot exist. □

Remark 6.8. The converse statement (the fact that absence of arbitrage implies the existence of equivalent martingale measures) is 'in principle' correct as well. This is the famous first fundamental theorem of arbitrage, established in full generality by Delbaen & Schachermayer (1994). A good discussion of the more technical aspects surrounding this result is also given in Chapter 10 of Björk (2004).

6.1.3 Hedging and risk-neutral pricing of contingent claims

A contingent claim X is an $\mathcal{F}_T$-measurable random variable, to be interpreted as payoff of some financial claims. In this section we introduce basic concepts related to hedging and pricing of contingent claim. We assume throughout that the model admits some martingale measure related to the numeraire S^0, and denote a fixed such measure by Q.

Definition 6.9. A selffinancing trading strategy is called Q-admissible, if the discounted gains process $\tilde{G}^\phi$ is a Q-martingale; $\Phi(Q)$ denotes the linear space of all Q-admissible strategies.

From now on we restrict our attention to contingent claims X such that $E^Q(|X|/S_T^0) < \infty$ (i.e. $\tilde{X} := X/S_T^0 \in \mathcal{L}^1(\Omega, \mathcal{F}_T, Q)$).

Definition 6.10. (i) A contingent claim X such that $\tilde{X} \in \mathcal{L}^1(\Omega, \mathcal{F}_T, Q)$ is called *attainable*, if there is at least one admissible strategy $\phi \in \Phi(Q)$ such that $V_T^\phi = X$, Q a.s.; any such strategy is called replicating strategy for X.

(ii) The financial market model is called *complete*, if any contingent claim X with $\tilde{X} := X/S_T^0 \in \mathcal{L}^1(\Omega, \mathcal{F}_T, Q)$ is attainable.

The financial motivation for the definition is of course the fact that by initially investing V_0^ϕ and then following the trading strategy ϕ the claim can be replicated without any further cost or risk. Note that in reality the performance of the trading strategy ϕ may be sub-optimal due to market frictions or model errors, and perfect replication of the claim may be impossible.

The next result links attainability to martingale representation.

Lemma 6.11. *Consider a contingent claim X such that $\tilde{X} = X/S_T^0 \in \mathcal{L}^1(\Omega, \mathcal{F}_T, Q)$. Then X is attainable if and only if the Q-martingale M with $M_t = E^Q(\tilde{X}|\mathcal{F}_t)$ admits a representation as stochastic integral of the form*

$$M_t = x + \sum_{i=1}^d \int_0^t \phi_u^i d\tilde{S}_u^i, \quad 0 \le t \le T, \tag{6.6}$$

for some constant x and some Q-admissible strategy $\phi^1, \dots, \phi^d$.

Proof. Suppose that (6.6) holds. We define the replicating strategy by putting $V_0 = xS^0(0)$, and by using $\phi_t^1, \dots, \phi_t^d$ as position in $S^1, \dots, S^d$. The discounted wealth process of the strategy satisfies (by Corollary 6.3 (ii))

$$\tilde{V}_T^\phi = x + \sum_{i=1}^d \int_0^T \phi_s^i d\tilde{S}_s^i = \frac{X}{S_T^0},$$

where the last equality follows from (6.6). Again by Corollary 6.3, we thus have $X = V_0 + G_T^\phi$, and the 'if part' follows. The converse statement is true by Corollary 6.3. ☐

Market completeness can be characterized by the uniqueness of an equivalent martingale measure.

Theorem 6.12 (Second fundamental theorem of asset pricing)**.** *Consider a market that admits (at least) one martingale measure Q. Then the market is complete if and only if the measure Q is unique.*

The 'if part' will be proven below; for the more difficult 'only if' part we refer to the literature.

Now we turn to the risk-neutral pricing of contingent claims. Consider an arbitrage-free market and a contingent claim X. Denote by $\Pi(t;X)$, $0 \le t \le T$ a candidate price process for X.

Proposition 6.13. 1. *The extended model with price processes $\big(S^0, S^1, \dots S^d, \Pi(\cdot, X)\big)$ admits an equivalent martingale measure and is thus arbitrage-free, if $X/S_T^0 \in \mathcal{L}^1(\Omega, \mathcal{F}_T, Q)$ and*

$$\Pi(t,X) = S_t^0 E^Q\Big(\frac{X}{S_T^0}|\mathcal{F}_t\Big) \tag{6.7}$$

for some martingale measure Q for the original market $(S^0, \dots, S^d)$.

2. *Two different martingale measure Q_1, Q_2 will in general lead to two different price processes $\Pi_1(t,X)$, $\Pi_2(t,X)$ which are both consistent with absence of arbitrage.*

3. *If X is attainable, we have for any martingale measure Q with $X/S_T^0 \in \mathcal{L}^1(\Omega, \mathcal{F}_T, Q)$ the relation*

$$S_t^0 E^Q\Big(\frac{X}{S_T^0}|\mathcal{F}_t\Big) = V_t^\phi,$$

where ϕ is a Q-admissible replicating strategy for X. In particular, the fair arbitrage-free price of an attainable claim is uniquely defined.

Proof. 1) (6.7) ensures that there is some martingale measure for the extended market. (This condition is essentially also necessary)

2) We have seen examples for this in the discrete-time setup.

3) If X is attainable we have the representation $X/S_T^0 = x + \tilde{G}_T^\phi$ for some Q-admissible strategy ϕ. Taking conditional expectation gives, as $\tilde{G}^\phi$ is a martingale

$$E(\frac{X}{S_T^0}|\mathcal{F}_t) = x + \tilde{G}_t^\phi = \tilde{V}_t^\phi, \tag{6.8}$$

where the last equality follows form Corollary 6.3. Multiplying both side of (6.8) with S_t^0 proves the claim. □

Remark 6.14. If the market is complete, we must have for any two martingale measures Q_1 and Q_2 and any bounded random variable $\tilde{X}$ that

$$E^{Q_1}(\tilde{X}) = V^\phi(0) = E^{Q_2}(\tilde{X}),$$

where ϕ is a replicating strategy for $X = \tilde{X}S_T^0$ (which exists by market completeness). Hence we must have $Q_1 = Q_2$, which is the 'if part' of Theorem 6.12.

6.1.4 Change of numeraire

We now discuss the technique of change of numeraire, which will be very useful for the valuation of claims under stochastic interest rates in the next section. We consider a model with $d+1$ assets $S^0, S^1, \ldots, S^d$ and assume that there is a martingale measure Q such that the discounted asset prices $\tilde{S}_t^i = S_t^i/S_t^0, 1 \le i \le d$ are Q-martingales. Consider now a new asset $X_t > 0$. The following theorem shows which martingale measure we have to use in order to work with X as new numeraire.

Theorem 6.15 (Change of numeraire)**.** *Consider a non-dividend-paying numeraire X_t such that X_t/S_t^0 is a true Q-martingale. Define a measure Q^X by putting*

$$\frac{dQ^X}{dQ}|\mathcal{F}_t = \eta_t := \frac{X_t}{S_t^0} \cdot \frac{S_0^0}{X_0}. \tag{6.9}$$

Then the processes S_t^i/X_t, $1 \le i \le d$ (the basic security process discounted using X) are Q^X-martingales. Moreover, we have for every contingent claim H such that $H/S_T^0 \in \mathcal{L}^1(\Omega, \mathcal{F}_T, Q)$ the change-of-numeraire formula

$$S_t^0 E^Q(H/S_T^0|\mathcal{F}_t) = X_t E^{Q^X}(H/X_T|\mathcal{F}_t). \tag{6.10}$$

We mention that the theorem remains true for a general change of numeraire where one starts with an arbitrary numeraire X and goes to a new numeraire $\tilde{X}$.

Proof. First we note that η_t is in fact a martingale with mean one: we have

$$E^Q(\eta_T|\mathcal{F}_t) = \frac{S_0^0}{X_0} E^Q\Big(\frac{X_T}{S_T^0}|\mathcal{F}_t\Big) = \frac{S_0^0}{X_0} \cdot \frac{X_t}{S_t^0} = \eta_t;$$

moreover, $E^Q(\eta_T) = \eta_0 = 1$. Hence the measure Q^X is well-defined. Consider now an arbitrary stochastic process $(H_t)_{0 \le t \le T}$ such that $\tilde{H}_t := H_t/S_t^0$ is a Q-martingale; H_t could for instance be the price process of one of the traded securities. We get from the Bayes formula (Lemma 5.26) that

$$\begin{aligned} E^{Q^X}\Big(\frac{H_T}{X_T}|\mathcal{F}_t\Big) &= \frac{1}{\eta_t} E^Q\Big(\frac{H_T}{X_T}\eta_T|\mathcal{F}_t\Big) \stackrel{i}{=} \frac{X_0 S_t^0}{S_0^0 X_t} E^Q\Big(\frac{H_T S_0^0 X_T}{X_T X_0 S_T^0}|\mathcal{F}_t\Big) \\ &= \frac{S_t^0}{X_t} E^Q\Big(\frac{H_T}{S_T^0}|\mathcal{F}_t\Big) \stackrel{ii}{=} \frac{H_t}{X_t}, \end{aligned}$$

where we have used i the definition of η in (6.9) and ii the fact that $\tilde{H}_t$ is a Q-martingale. The above computation shows that H_t/X_t is a Q^X-martingale and thus proves the first part of the theorem.

For the change-of-numeraire formula define the martingale $\tilde{H}_t = E^Q\big(H_T/S_T^0 \mid \mathcal{F}_t\big)$ and let $H_t = S_t^0 \tilde{H}_t$ (the risk-neutral price of H). Then we get from the first part of the proof that

$$E^{Q^X}\Big(\frac{H}{X_T}|\mathcal{F}_t\Big) = \frac{H_t}{X_t},$$

and multiplication with X_t gives the result. ◻

6.2 The Black-Scholes Model Revisited

Setup. As in the previous section we consider a model with two traded assets, stock and money market account, with dynamics given by

$$dS_t^0 = rS_t^0 dt, \quad S_0^0 = 1 \tag{6.11}$$
$$dS_t^1 = \mu S_t^1 dt + \sigma S_t^1 dW_t\,, \quad S_0^1 = S_0 > 0. \tag{6.12}$$

Note that (6.11) implies that $S_t^0 = \exp(rt)$. We define the discount factor by $D_t = (S_t^0)^{-1} = \exp(-rt)$. The discounted stock price is given by $\tilde{S}_t := D_t S_t^1$; the discounted price of S^0 is by definition equal to 1. In economic terms discounting corresponds to taking the money market account S^0 as new numeraire. We now use Itô's-formula to compute the dynamics of the discounted stock price. Since $dD_t = -r(D_t)dt$, the Itô'-product-formula yields

$$\begin{aligned} d\tilde{S}_t^1 &= S_t^1 dD_t + D_t dS_t^1 + d[S_t^1, D_t]_t = -r\tilde{S}_t^1 dt + \mu\tilde{S}_t^1 dt + \sigma\tilde{S}_t^1 dW_t \\ &= (\mu - r)\tilde{S}_t^1 dt + \sigma\tilde{S}_t^1 dW_t, \end{aligned}$$

where we have used that $[S^1, D]_t \equiv 0$ as (D_t) is of finite variation.

Equivalent Martingale Measure. Recall from Lemma 6.6 that the discounted price process $\tilde{S}^1$ is a Q-martingale if and only if S^1 itself has dynamics of the form $S_t^1 = S_0 + \int_0^t rS_s^1 ds + M_t^Q$, where M^Q is a Q-local martingale. Comparison with the Black-Scholes SDE shows that we need to define Q in such a way that by going from P to Q the drift is changed from μ to r .

Now we apply Girsanov's theorem to determine the[1] equivalent martingale measure in the Black-Scholes model. We have

$$\begin{aligned} S_t^1 &= S_0 + \int_0^t \mu S_s^1 ds + \int_0^t \sigma S_s^1 dW_s \\ &= S_0 + \int_0^t rS_s^1 ds + \int_0^t \left(\frac{\mu - r}{\sigma}\right) \sigma S_s^1 ds + \int_0^t \sigma S_s^1 dW_s \\ &= S_0 + \int_0^t rS_s^1 ds + \int_0^t \sigma S_s^1 d\tilde{W}_s \end{aligned}$$

where $\tilde{W}_t = W_t + \int_0^t \frac{\mu - r}{\sigma} ds$. Define $\lambda := \frac{\mu - r}{\sigma}$; λ is often referred to as *market price of risk* of the stock. In order for S to be of the form $S_t^1 = S_0 + \int_0^t rS_s^1 ds + \tilde{M}_t$ we need to turn $\tilde{W}_t = W_t + \int_0^t \lambda ds$ into a martingale by a change of measure. Girsanov's theorem tells us that under Q with

$$\frac{dQ}{dP} = \exp(-\lambda W_T - 1/2\lambda^2 T)$$

the process $\tilde{W}$ is a Q-Brownian motion, hence the Itô-integral $\tilde{M}_t = \int_0^t \sigma S_s^1 d\tilde{W}_s$ is a Q-local martingale (and in fact a true martingale as is easy to check). Summing up we have

Proposition 6.16. *Consider a Black-Scholes model with stock-price dynamics of the form* $dS_t^1 = \sigma S_t^1 dW_t + \mu S_t^1 dt$ *for constants* μ, σ. *Then the equivalent martingale measure* Q *is*

[1] We will see in Section 6.3.3 that the Black-Scholes model is complete, so that the equivalent martingale measure is unique by the second fundamental theorem.

given by the density $\frac{dQ}{dP} = G_T := \exp(-\lambda W_T - 1/2\lambda^2 T)$*, where* $\lambda = \frac{\mu - r}{\sigma}$ *is the market price of risk of the stock. Under* Q *the stock-price process solves the SDE*

$$dS_t^1 = rS_t^1 dt + \sigma S_t^1 d\tilde{W}_t \tag{6.13}$$

Black-Scholes via risk-neutral valuation. We now derive the Black-Scholes formula from the risk-neutral valuation principle; this is the easiest approach from a computational perspective. We know that S_t^1 solves equation (6.13) and is thus of the form $S_t^1 = S_0 \exp(\sigma \tilde{W}_t + (r - 1/2\sigma^2)t)$. From the risk-neutral pricing-rule – which applies as we already know from Chapter 4 that the call can be replicated – we get for C_0, the call price at time zero

$$C_0 \;=\; E^Q(e^{-rT}(S_T^1 - K)^+) = \underbrace{E^Q(e^{-rT} S_T^1 1_{(S_T^1 > K)})}_{(I)} - e^{-rT} K \underbrace{Q(S_T^1 > K)}_{(II)} \,.$$

We start with the second term. One has

$$Q(S_T^1 > K) = Q(\ln S_T^1 > \ln K) = Q\left(\frac{\sigma \tilde{W}_T}{\sqrt{\sigma^2 T}} > \frac{\ln K - \ln S_0 - (r - 1/2\sigma^2)T}{\sqrt{\sigma^2 T}} \right) .$$

Now note that $Z = \frac{\sigma \tilde{W}_T}{\sigma \sqrt{T}} \sim N(0,1)$. By the symmetry of centered normal distributions we have for $M \in \mathbb{R}$ that $Q(Z > M) = Q(Z < -M)$. Hence

$$Q(S_T^1 > K) = Q\left(Z < \frac{\ln S_0/K + (r - 1/2\sigma^2)T}{\sigma \sqrt{T}} \right) = N(d_2)\,.$$

To deal with term (I) we need to apply Girsanov's theorem once more. Note that $e^{-rT} S_T^1 = S_0 Y_T$ where $Y_T := \exp(\sigma \tilde{W}_T - 1/2\sigma^2 T)$ has the form of a Girsanov-density. Define a new measure Q^S by $dQ^S/dQ = Y_T$. Then we have

$$E^Q\left(e^{-rT} S_T^1 1_{\{S_T^1 > K\}}\right) = S_0 E^Q\left(Y_T 1_{\{S_T^1 > K\}}\right) = S_0 E^{Q^S}(1_{\{S_T^1 > K\}}) = S_0 Q^S(\ln S_T^1 > \ln K)\,.$$

Girsanov's theorem implies that $W_t^S := \tilde{W}_t - \sigma t$ is a Q^S-Brownian motion. Hence $\sigma \tilde{W}_T = \sigma W_T^S + \sigma^2 T$ and

$$Q^S(\ln S_t^1 > \ln K) = Q^S(\ln S_0 + (r - 1/2\sigma^2)T + \sigma W_T^S + \sigma^2 T > \ln K)\,.$$

The right hand side is now evaluated in exactly the same way as before, yielding

$$E^Q\left(e^{-rT} S_T^1 1_{(S_T^1 > K)}\right) = S_0 N(d_1)\,. \quad \square$$

6.3 Fundamental Theorems of Asset Pricing in a Generalized Black-Scholes Model

6.3.1 A multidimensional Black-Scholes model

The traded assets consist of one money-market account S^0 and d stocks $S^1, \dots, S^d$. Given a filtrated probability space $(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_t\}$, $0 \le t \le T$, we assume that the stock-price is an Itô-process of the form

$$dS_t^i = \alpha_t^i S_t^i dt + S_t^i \sum_{j=1}^n \sigma_{ij}(t) dW_t^j, \quad 1 \le i \le d, \tag{6.14}$$

for an n-dimensional standard Brownian motion $W_t = (W_t^1, \dots, W_t^n)$ and adapted process α_t^i, $1 \le i \le d$ and $\sigma_{ij}(t)$, $1 \le i \le d$, $1 \le j \le n$. The initial values are given by $S_0 = (S_0^1, \dots, S_0^d)$. The money-market account satisfies

$$S_t^0 = \exp(\int_0^t r_s ds)$$

for some adapted process $(r_s)_{0 \le s \le t}$ modelling the (random) short-rate of the interest. We now give a more intuitive description of the model (6.14). Define

$$\sigma_i(t) = \big(\sum_{j=1}^n \sigma_{ij}(t)^2\big)^{\frac{1}{2}} \tag{6.15}$$

and assume that $\sigma_i(t) > 0$ a.s.. Define processes B_t^i, $1 \le i \le d$, via

$$B_t^i = \sum_{j=1}^n \int_0^t \frac{\sigma_{ij}(s)}{\sigma_i(s)} dW_s^j. \tag{6.16}$$

It is easy seen that $[B^i]_t = t$ (exercise!), so that $B^1, \dots B^d$ are one-dimensional Brownian motions by the Lévy-characterization of Brownian motion. Moreover, one has for $i \neq j$ that

$$[B^i, B^j]_t = \int_0^t \rho_{ij}(s) ds \quad \text{with} \quad \rho_{ij}(s) = \frac{\sum_{l=1}^n \sigma_{il}(t)\sigma_{jl}(t)}{\sigma_i(t)\sigma_j(t)}.$$

Note that $\rho_{ij}(t) \in [-1, 1]$; it is termed instantaneous correlation of B^i and B^j. In terms of the B^i the asset price dynamics can be written in the form

$$dS_t^i = \alpha_t^i S_t^i dt + \sigma_i(t) S_t^i dB_t^i, \quad 1 \le i \le d, \tag{6.17}$$

so that $\sigma_i(t)$ is the volatility of S^i at time t. Moreover, we get from Itô's formula

$$d \ln S_t^i = (\alpha_t^i - \frac{1}{2}\sigma_i^2(t))dt + \sigma_i(t) dB_t^i,$$

so that

$$[\ln S^i, \ln S^j]_t = \int_0^t \sigma_i(s)\sigma_j(s) d[B^i, B^j]_s = \int_0^t \sigma_i(s)\sigma_j(s)\rho_{ij}(s) ds.$$

The quantity ρ_{ij} therefore models the instantaneous correlation of the logarithmic return processes: for h small we have

$$E\big((\ln S_{t+h}^i - \ln S_t^i)(\ln S_{t+h}^j - \ln S_t^j)|\mathcal{F}_t\big) \approx (\sigma_i \cdot \sigma_j \cdot \rho_{ij})h.$$

As in the case of the one-dimensional Black-Scholes model the dynamics of the discounted stock prices $\tilde{S}_t^i = S_t^i / S_t^0$ are given by

$$d\tilde{S}_t^i = (\alpha_t^i - r_t)\tilde{S}_t^i dt + \sigma_i(t)\tilde{S}_t^i dB_t^i = (\alpha_t^i - r_t)\tilde{S}_t^i dt + \tilde{S}_t^i d(\sum_{j=1}^n \sigma_{ij}(t) dW_t^j). \tag{6.18}$$

6.3.2 Existence of an equivalent martingale measure

The following Proposition gives sufficient conditions for the existence of an equivalent martingale measure, and hence for the absence of arbitrage in the model (6.14).

Proposition 6.17. *Suppose that for all* $0 \le t \le T$ *there is a solution* $\theta_t = (\theta_t^1, \dots, \theta_t^n)$ *of the so-called* market-price-of-risk equations

$$\alpha_t^i - r_t = \sum_{j=1}^{n} \sigma_{ij}(t)\theta_t^j, \quad 1 \le i \le d, \tag{6.19}$$

and that $E\big(\exp(\frac{1}{2}\int_0^T \|\theta_s\|^2 ds)\big) < \infty$. *Then the model* (6.14) *admits an equivalent martingale measure.*

Proof. Using the process θ_t from (6.19) we can write the SDE (6.18) for the discounted stock price $\tilde{S}^1, \dots \tilde{S}^d$ in the form in the form

$$d\tilde{S}_t^i = (\alpha_t^i - r_t)\tilde{S}_t^i dt + \tilde{S}_t^i \sum_{j=1}^{n} \sigma_{ij}(t) dW_t^j = \tilde{S}_t^i \sum_{j=1}^{n} \sigma_{ij}(t) d\Big(W_t^j + \theta_t^j dt\Big) \tag{6.20}$$

Define $\tilde{W}_t^j = W_t^j + \int_0^t \theta_s^j ds$, $1 \le j \le n$. We now use the Girsanov theorem to turn $\tilde{W}$ into a new Brownian motion. Define

$$Z_T := \exp(-\sum_{j=1}^{n} \int_0^T \theta_s^j dW_s^j - \frac{1}{2}\int_0^T \|\theta_s\|^2 ds). \tag{6.21}$$

The integrability condition on θ ensures $E(Z_T) = 1$. Define a new measure Q by $\frac{dQ}{dP}|_{\mathcal{F}_T} = Z_T$. The multidimensional Girsanov-theorem shows that under Q the process $\tilde{W}$ with is Q-Brownian motion, and the result follow from Lemma 6.6. ☐

Denote by $\sigma(t)$ the matrix $(\sigma_{ij}(t))_{1 \le i \le d, 1 \le j \le n}$ and by $\mathbf{1}$ the vector $(1, \dots, 1)' \in \mathbb{R}^d$. Suppose that the market price of risk equation does not admit a solution. This implies that the vector $(\alpha_t - r_t\mathbf{1})$ does not belong to $\mathrm{Im}\sigma(t)$ (the range of the matrix $\sigma(t)$) and hence we get for $\ker\sigma'(t) := \{x \in \mathbb{R}^d : \sigma'(t)x = 0\}$

$$\ker\sigma'(t) = \big(\mathrm{Im}\sigma(t)\big)^{\perp} \ne \{0\}.$$

Hence we can find $\xi_t \in \mathbb{R}^d$ such that

$$\sum_{i=1}^{d} \xi_t^i(\alpha_t^i - r_t) \ne 0, \quad \text{and} \quad \sigma'(t)\xi_t = 0,$$

otherwise $(\alpha_t - r_t\mathbf{1}) \in \ker\sigma'(t) = \mathrm{Im}\sigma(t)$, which contradicts the fact that the market-price-of-risk equation has no solution. Define a trading-strategy ϕ by $\phi_t^i = \xi_t^i/\tilde{S}_t^i$. The corresponding discounted wealth process $\tilde{V}^\phi$ satisfies

$$d\tilde{V}_t^\phi = \sum_{i=1}^{d} \xi_t^i(\alpha_t^i - r_t)dt + \sum_{j=1}^{n}(\sum_{i=1}^{d} \sigma_{ij}(t)\xi_t^i)dW_t^j = \sum_{i=1}^{d} \xi_t^i(\alpha_t^i - r_t)dt,$$

where we have used that $\sum_{i=1}^d \sigma_{ij}(t)\xi_t^i = \big(\sigma'(t)\xi_t\big)_j = 0$ as $\xi_t \in \ker\sigma'(t)$. Hence we have constructed a locally riskless selffinancing trading strategy whose discounted wealth-process is non-constant, and it is easy to construct arbitrage-opportunities from this. Hence we have shown the following result:

Theorem 6.18. *Modulo integrability conditions, the model* (6.14) *is arbitrage-free if and only if the market-price of risk equation* (6.19) *admits a solution.*

6.3.3 Market completeness in the generalized Black-Scholes model

In order to discuss market completeness we need the following martingale representation theorem for Brownian motion.

Theorem 6.19 (Ito-representation theorem for Brownian motion)**.** *Consider an n-dimensional Brownian motion $W = (W_{t,1}, \ldots, W_{t,n})'$ on some $(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_t\}$, $0 \le t \le T$ and denote by $\{\mathcal{F}_t^W\}$ the filtration generated by W. Then every $\{\mathcal{F}_t^W\}$-adapted square integrable martingale M can be written as a stochastic integral with respect to W, that is there is a n-dimensional adapted process $h_t = (h_t^1, \ldots, h_t^n)$ with $E\left(\int_0^T (h_s^i)^2 ds\right) < \infty$ for all i such that*

$$M_t = M_0 + \sum_{i=1}^{n} \int_0^t h_s^i dW_{s,i}, \quad 0 \le t \le T.$$

As before, we consider a model with $d+1$ traded assets and dynamics of the form $dS_t^0 = r_t S_t^0 dt$, and

$$dS_t^i = \alpha_t^i S_t^i dt + \sum_{j=1}^{n} \sigma_{ij}(t) S_t^i dW_t^j, \quad 1 \le i \le d, \tag{6.22}$$

for an n-dimensional Brownian motion $W_t = (W_t^1, \ldots, W_t^n)$. Moreover, we assume that our filtration is generated by the Brownian motion, that is $\mathcal{F}_t = \sigma(W_s^i : s \le t, 1 \le i \le n)$; this assumption restricts the contingent claims we may consider and is thus crucial for establishing completeness of the market.

Proposition 6.20. *Suppose that the model* (6.22) *admits at least one equivalent martingale measure Q, and that the underlying filtration is generated by the Brownian motion W. Then the market is complete if and only if* P.a.s. $\mathrm{Im}\sigma'(t) = \mathbb{R}^n$, $0 \le t \le T$, *where $\sigma(t)$ represents the matrix $\left(\sigma_{ij}(t)\right)_{1\le i\le d, 1\le j\le n}$.*

Proof. Denote by Q some equivalent martingale measure, which exists by assumption. As shown previously, under Q the discounted state prices have dynamics

$$d\tilde{S}_t^i = \tilde{S}_t^i \sum_{j=1}^{n} \sigma_{ij}(t) d\tilde{W}_t^j, \tag{6.23}$$

where $\tilde{W}^j$ is a Q-Brownian motion. By Lemma 6.11, the market is complete if and only if every Q-martingale M has a representation

$$M_t = M_0 + \sum_{i=1}^{d} \int_0^t \phi_s^i d\tilde{S}_s^i. \tag{6.24}$$

Using Itô's representation theorem (Theorem 6.19) and the assumption $\{\mathcal{F}_t\} = \{\mathcal{F}_t^W\}$ it is straightforward that M has a unique representation of the form

$$M_t = M_0 + \sum_{j=1}^{n} \int_0^t h_s^j d\tilde{W}_s^j. \tag{6.25}$$

On the other hand, writing (6.24) in short-hand notation and using (6.23) we get that the martingale representation of M in (6.24) is equivalent to

$$dM_t = \sum_{i=1}^{d} \tilde{S}_t^i \phi_t^i \sum_{j=1}^{n} \sigma_{ij}(t) d\tilde{W}_t^j = \sum_{j=1}^{n} \Big(\sum_{i=1}^{d} \tilde{S}_t^i \sigma_{ij}(t) \phi_t^i \Big) d\tilde{W}_t^j. \tag{6.26}$$

Comparing (6.25) and (6.26) we see that we must have

$$h_t^j = \sum_{i=1}^{d} \sigma_{ij}(t) \tilde{S}_t^i \phi_t^i = \sum_{i=1}^{d} \sigma_{ij}(t) \tilde{\phi}_t^i, \tag{6.27}$$

where $\tilde{\phi}_t^i = \phi_t^i \tilde{S}_t^i$. This equation does indeed have a solution since we assumed that $\mathrm{Im}\sigma'(t) = \mathbb{R}^n$ and we have shown the existence of a replication strategy.

The proof of the converse direction is based on the observation that if $\mathrm{Im}\sigma'(t)$ is a strict subset of $\mathbb{R}^d$ we can find a process h^* such that (6.27) does not have a solution, so that the martingale $M_t = M_0 + \sum_{j=1}^{n} \int_0^t (h^*)_s^j d\tilde{W}_s^j$ cannot be represented as stochastic integral with respect to the stock price processes; we omit the details.

□

Remark 6.21 (On sources of risk and number of risky assets)**.** Assume that asset prices follow the model (6.22), denote by $\sigma(t)$ the matrix $\big(\sigma_{ij}(t)\big)_{1\le i\le d, 1\le j\le n}$, and let $\mathbf{1} = (1, \dots, 1)'$. Then we have the following two observations.

- If the model is arbitrage-free for *any* choice of the drift vector α_t, the market price of risk equation needs to have a solution for any excess return vector $\alpha_t - r_t\mathbf{1}$. For this $\sigma(t)$ needs to have rank d (as $\sigma(t)$ is a $d \times n$ matrix), and hence a necessary condition for absence of arbitrage for generic α_t is the inequality $d \le n$ (at least as many sources of risk (Brownian motions) as risky assets).
- By Proposition 6.20, a necessary (and essentially sufficient) condition for completeness of the generalized Black Scholes model is the condition that the rank of $\sigma'(t)$ equals n and hence the inequality $n \le d$ (at least as many risky assets as sources of risk).

Chapter 7

Optimization Problems in Continuous-Time Finance

7.1 Portfolio optimization and stochastic optimal control

In this section we give an introduction to portfolio optimization via stochastic control theory. The main result is the derivation of a nonlinear PDE for the value function of the control problem. Our exposition follows closely the textbook Björk (2004).

7.1.1 The portfolio optimization problem

Consider a market with two traded assets S^0 and S^1,; assume that $S_t^0 = \exp(rt), r > 0$, and that the dynamics of S^1 is given by the Black Scholes model, so that S^1 solves the SDE $dS_t^1 = \mu S_t^1 dt + \sigma S_t^1 dW_t$ for constants μ and $\sigma > 0$ and a Brownian motion W. We are interested in the trading strategy of an investor in this market with initial wealth v_0 who wants to invest in some sense optimally. The first step is to describe what we mean by an optimal investment strategy. Here we assume that the investor wants to maximize expected utility of the terminal wealth. Consider a horizon date T and a *utility function* $u : \mathbb{R}^+ \to \mathbb{R}$ with $u' > 0$ and $u'' < 0$. For concreteness and ease of computation we consider a utility function of the form

$$u(x) = \frac{x^\gamma}{\gamma}, \ \gamma < 1, \gamma \neq 0, \tag{7.1}$$

and $u(x) = \ln x$, if $\gamma = 0$. These utility functions have derivative $u'(x) = x^\gamma, \gamma < 1$, and the so-called and *Arrow-Pratt risk aversion* $-u''/u'$ is given by $(1-\gamma)x^{-1}$. Note that the risk aversion is inversely proportional to the wealth x of the investor. For this reason a utility function of the form (7.1) is said to exhibit *constant relative risk aversion.*

Formally, the problem of maximizing expected utility of terminal wealth amounts to finding some selffinancing strategy of the form $\hat\phi_t = (\hat\phi_t^0, \hat\phi_t^1)$ in stock and bond with wealth processes $V_t^{\hat\phi}$ and initial wealth $V_0^{\hat\phi} = v_0$ that maximizes $E\big(u(V_T^{\hat\phi})\big)$ over all admissible strategies with initial wealth no larger than v_0. Sometimes one considers also a modified version of the problem where some of the wealth is consumed before maturity. More precisely, it is assumed that that the investor consumes his wealth at a rate $c_t dt$ for an adapted process $c_t \geq 0$ so that the wealth dynamics of some admissible strategy $(\phi_t^0, \phi_t^1, c_t)$ become

$$dV_t = \phi_t^0 dS_t^0 + \phi_t^1 dS_t^1 - c_t dt, \quad V_0 \leq v_0 \, .$$

The objective of the investor is then to maximize the quantity

$$E\Big(\gamma \int_0^T e^{-\tilde{r}t} u(c_t)dt + u(V_T^{\hat{\phi}})\Big)$$

over all admissible (to be specified later) strategies (ϕ^0, ϕ^1, c). Here $\gamma \geq 0$ is some constant that models the relative importance of intermediate and terminal consumption and $\tilde{r}$ is a subjective discount rate.

Formulation as stochastic control problem. Next we reformulate the optimization problem as a stochastic control problem. We consider only strategies such that $V_t(\phi) > 0$. Hence we may define the *relative portfolio weights* by

$$\pi_t^0 = \frac{\phi_t^0 S_t^0}{V_t}, \quad \pi_t^1 = \frac{\phi_t^1 S_t^1}{V_t}.$$

Note that by definition π_t^i gives the proportion of the overall wealth V_t invested in asset i and that $\pi_t^0 + \pi_t^1 = 1$. It is quite natural to describe a portfolio strategy in terms of relative portfolio weights since rules such as"invest 70 % in stock and 30 % in bond" are typical descriptions of investment strategies used in practice. The wealth equation can be rewritten in terms of the relative portfolio weights as we now show. It holds that

$$\begin{aligned} V_t &= V_0 + \int_0^t \phi_s^0 dS_s^0 + \int_0^t \phi_s^1 dS_s^1 - \int_0^t c_s ds \\ &= V_0 + \int_0^t \pi_s^0 r V_s ds + \int_0^t \pi_s^1 \mu V_s ds + \int_0^t \pi_s^1 \sigma V_s dW_s - \int_0^t c_s ds \\ &= V_0 + \int_0^t r V_s ds + \int_0^t \pi_s^1 V_s(\mu - r) ds + \int_0^t \pi_s^1 V_s \sigma dW_s - \int_0^t c_s ds. \end{aligned} \tag{7.2}$$

Hence V_0, and the processes $\pi_t = \pi_t^1$ and c_t determine the evolution of the wealth process V_t; in particular the stock price does not appear in this description of the wealth process. The investor problem can now be rewritten in the form

$$\max_{\pi,c} E\Big(\gamma \int_0^T e^{-\tilde{r}t} u(c_t)dt + u(V_T)\Big) \text{ such that} \tag{7.3}$$

$$dV_t = rV_t dt + \pi_t V_t(\mu - r)dt + \sigma \pi_t V_t dW_t - c_t dt, \quad V_0 = v_0 \tag{7.4}$$

$$c_t \geq 0 \text{ for all } 0 \leq t \leq T \tag{7.5}$$

Here we have implicitly used that it is always optimal to invest the initial capital v_0 in full.

7.1.2 The Dynamic Programming Equation

In this section we consider a general control problem that contains the portfolio optimization problem (7.3) as a special case. Consider a Brownian motion W and functions $\mu\colon \mathbb{R} \times \mathbb{R}^n \to$

$\mathbb{R}$ and $\sigma\colon \mathbb{R}\times\mathbb{R}^n \to [0,\infty)$. For a given $x_0 \in \mathbb{R}$ we consider the so-called *state process* V with dynamics given by the controlled SDE

$$dV_t = \mu(V_t, \pi_t)dt + \sigma(V_t, \pi_t)dW_t, \quad V_0 = v_0. \tag{7.6}$$

. In this model the state process is controlled or steered by a proper choice of the n-dimensional control strategy π.

As a first step for setting up a formal control problem we need to define the class of admissible control processes. A minimal requirement is the fact that π_t should depend only on the past values of the state process V. A class of strategies that satisfy this requirement are *feedback control strategies* where

$$\pi_t = \pi(t, V_t)$$

for a suitable function π with values in $\mathbb{R}^n$. In most problems one has to require that π takes its values in some subset $\mathcal{K}$ of $\mathbb{R}^n$ (so-called control constraints). For instance in the portfolio optimization problem we required that $c_t \geq 0$.

Definition 7.1. An admissible feedback control strategy is a function $\pi\colon [0,\infty)\times\mathbb{R}\to\mathcal{K}$ and such that for all initial values $(t,v)\in[0,\infty)\times\mathbb{R}$ the SDE

$$dV_t = \mu(V_t, \pi(t, V_t))dt + \sigma(V_t, \pi(t, V_t))dW_t, \quad V_t = v \tag{7.7}$$

has a unique solution. The set of admissible feedback control strategies is denoted by $\mathcal{A}$.

Note that for a given strategy $\pi\in\mathcal{A}$ the state process (7.7) is a Markov process. Hence feedback control strategies are also known as *Markov control strategies.* For a given admissible strategy π and initial value v the solution of the SDE (7.7) will sometimes be denoted by $V^{v,\pi}$ or by V^{π}. Moreover, for a given admissible feedback strategy π we often write the SDE (7.7) in the short form $dV_t^{\pi} = \mu^{\pi}(t, V_t^{\pi})dt + \sigma^{\pi}(t, V_t^{\pi})dW_t$, where, of course, $\mu^{\pi}\left(t, V_t^{\pi}\right) = \mu\left(t, V_t^{\pi}, \pi(t, V_t^{\pi}\right)$, and similarly for σ^{π}; sometimes the arguments of μ^{π} and σ^{π} will even be omitted. Finally, for a given strategy $\pi\in\mathcal{A}$ we often write $\pi_t := \pi(t, V_t^{\pi})$.

Next we describe the objective of the control problem. Consider a pair of functions

$$F\colon [0,\infty)\times\mathbb{R}\times\mathbb{R}^n\to\mathbb{R} \qquad u\colon \mathbb{R}\to\mathbb{R}.$$

Now we define the *reward function* of the control problem as the function $J_0\colon \mathcal{A}\to\mathbb{R}$ with

$$J_0(\pi) = E\left(\int_0^T F(t, V_t^{v_0,\pi}, \pi(t, V_t^{v_0,\pi}))dt + u\left(V_T^{v_0,\pi}\right)\right) \tag{7.8}$$

The optimization problem is now to maximize J_0 over all $\pi\in\mathcal{A}$. The optimal value is thus given by $\hat{J}_0 = \sup\{J_0(\pi)\colon \pi\in\mathcal{A}\}$, and a strategy $\hat{\pi}\in\mathcal{A}$ is optimal if $\hat{J}_0 = J_0(\hat{\pi})$.

Of course, as with any optimization problem it is not clear if an optimal control strategy exists. In the sequel we will assume that an optimal strategy exists and we show how it can (in principle) be computed using PDE methods. Moreover, we give a verification result that gives sufficient conditions which ensure that a candidate solution produced by our approach is in fact an optimal strategy.

Step 1: A larger class of control problems. Fix $t \in [0,T], x \in \mathbb{R}$. Suppose that we solve the control problem only over the time period $[t,T]$, starting at time t with an initial value x. This leads to the control problem $P(t,x)$:

$$\max E\left(\int_t^T F(s, V_s^\pi, \pi(s, V_s^\pi))ds + u(V_T^\pi) \mid V_t^\pi = x\right)$$

over all strategies $\pi \in \mathcal{A}$. The original problem is of course the problem $P(0, v_0)$.

Definition 7.2. (i) The reward function associated to the problem $P(t,x)$ is the function $J\colon [0,T] \times \mathbb{R}^+ \times \mathcal{A} \to \mathbb{R}$ with

$$J(t,x,\pi) = E\left(\int_t^T F(s, V_s^\pi, \pi(s, V_s^\pi))ds + u(V_T^\pi) \mid V_t^\pi = x\right)$$

(ii) The value function is defined by $\hat{J}(t,x) = \sup\{J(t,x,\pi) : \pi \in \mathcal{A}\}$.

Step 2: The dynamic programming principle. The following result is crucial for the analysis of control problems with dynamic programming.

Proposition 7.3 (Dynamic Programming principle)**.** *Denote by $\mathcal{T}_{t,T}$ the set of all stopping times θ with $t \le \theta \le T$.*

1. For all admissible strategies $\pi \in \mathcal{A}$ and all $\theta \in \mathcal{T}_{t,T}$ it holds that

$$\widehat{J}(t,x) \ge E\left(\int_t^\theta F(s, V_s^\pi, \pi_s)\, ds + \widehat{J}(\theta, V_\theta^\pi) \mid V_t = x\right) \tag{7.9}$$

2. For all $\epsilon > 0$ there is some strategy $\tilde{\pi} \in \mathcal{A}$ such that for all $\theta \in \mathcal{T}_{t,T}$ it holds that

$$\widehat{J}(t,x) - \epsilon \le E\left(\int_t^\theta F(s, V_s^{\tilde{\pi}}, \tilde{\pi}_s)\, ds + \widehat{J}(\theta, V_\theta^{\tilde{\pi}}) \mid V_t = x\right)$$

As a corollary we obtain the classical dynamic programming principle.

Corollary 7.4. *For any $\theta \in \mathcal{T}_{t,T}$ it holds that*

$$\widehat{J}(t,x) = \sup_{\pi \in \mathcal{A}} E\left(\int_t^\theta F(s, V_s^{\tilde{\pi}}, \tilde{\pi}_s)\, ds + \widehat{J}(\theta, V_\theta^{\tilde{\pi}}) \mid V_t = x\right).$$

Proof of Proposition 7.3. Consider an arbitrary admissible strategy π. We get by iterated conditional expectations

$$J(t,x,\pi) = E\left(\int_t^\theta F(s, V_s^\pi, \pi_s)\, ds + J(\theta, V_\theta^\pi, \pi) \mid V_t = x\right) \tag{7.10}$$

Fix now some control $\pi \in \mathcal{A}$ and some $\theta \in \mathcal{T}_{t,T}$. Choose an ϵ-optimal strategy $\hat{\pi} \in \mathcal{A}$ for the problem 'starting at θ', that is a strategy $\hat{\pi}$ with $\widehat{J}(\theta,x) \le J(\theta,x,\pi) + \epsilon$ for all x. Define π^* by

$$\pi^*(s,y) = \begin{cases} \pi(s,y), s \in [t,\theta], y \in \mathbb{R}^+ \\ \hat{\pi}(s,y), s \in (\theta,T], y \in \mathbb{R}^+ \end{cases} \quad ;$$

Then $\widehat{J}(t,x) \geq J(t,x,\pi^*) \geq E\left(\int_t^\theta F(s,V_s^\pi,\pi_s)\,ds + \widehat{J}(\theta,V_\theta^\pi) - \epsilon \mid V_t = x\right)$, and for $\epsilon \to 0$ we get the first inequality.

Now we turn to the second inequality in the proposition. Using the inequality $J(\theta,V_\theta^\pi,\pi) \leq \widehat{J}(\theta,V_\theta^\pi)$ and (7.10) we get

$$J(t,x,\pi) \leq E\left(\int_t^\theta F(s,V_s^\pi,\pi_s)\,ds + \widehat{J}(\theta,V_\theta^\pi) \mid V_t = x\right) \quad \text{and thus}$$

$$J(t,x,\pi) \leq \inf_{\theta\in\mathcal{T}_{t,T}} E\left(\int_t^\theta F(s,V_s^\pi,\pi_s)\,ds + \widehat{J}(\theta,V_\theta^\pi) \mid V_t = x\right).$$

By taking the sup over π we get

$$\widehat{J}(t,x) \leq \sup_{\pi\in\mathcal{A}} \inf_{\theta\in\mathcal{T}_{t,T}} E\left(\int_t^\theta F(s,V_s^\pi,\pi_s)\,ds + \widehat{J}(\theta,V_\theta^\pi) \mid V_t = x\right).$$

By definition of the supremum there is some $\pi \in \mathcal{A}$ such that

$$\widehat{J}(t,x) - \epsilon \leq \inf_{\theta\in\mathcal{T}_{t,T}} E\left(\int_t^{t+h} F(s,V_s^{\tilde{\pi}},\tilde{\pi}_s)\,ds + \widehat{J}(\theta,V_\theta^{\tilde{\pi}}) \mid V_t = x\right),$$

which is the second inequality. □

Note that the dynamic programming principle implies that for every stopping time $\theta \in \mathcal{T}_{t,T}$ it holds that

$$\widehat{J}(t,x) = \sup_{\pi\in\mathcal{A}} E\left(\int_t^\theta F(s,V_s^\pi,\pi_s)\,ds + \widehat{J}(\theta,V_\theta^\pi) \mid V_t = x\right). \tag{7.11}$$

Step 3: The HJB equation. We now use heuristic arguments and the dynamic programming principle to derive a nonlinear PDE for the value-function $\hat{J}$; this PDE will provide a method for computing the optimal strategy $\hat{\pi}$ as well. For this we denote for $\lambda \in \mathcal{K}$ by $\mathcal{L}^\lambda$ the differential operator

$$\mathcal{L}^\lambda f(x) = \mu(x,\lambda)\frac{\partial f}{\partial x}(x) + \frac{1}{2}\sigma^2(x,\lambda)\frac{\partial^2 f}{\partial x^2}(x). \tag{7.12}$$

Note that $\mathcal{L}^\lambda$ is the generator of the state process V if a constant strategy $\pi_t \equiv \lambda$ is being used. For instance we get for the portfolio optimization problem with $\lambda = (\pi, c)$

$$\mathcal{L}^\lambda f(x) = \big((r + (\mu - r)\pi)x - c\big)\frac{\partial f}{\partial x}(x) + \frac{1}{2}\pi^2\sigma^2 x^2\frac{\partial^2 f}{\partial x^2}(x).$$

In order to proceed we assume that

A1) The optimal value function $\hat{J}$ is once differentiable in t and twice in x. Moreover, all stochastic integrals with respect to Brownian motion are true martingales.

In the next step we use Itô's Lemma to obtain a PDE for $\widehat{J}$ from the dynamic programming principle. It holds that

$$\hat{J}(t+h,V_{t+h}^\pi) = \hat{J}(t,x) + \int_t^{t+h} \hat{J}_t(s,V_s^\pi) + \mathcal{L}^\pi\hat{J}(s,V_s^\pi)\,ds + \int_t^{t+h} \hat{J}_x(s,V_s^\pi)\sigma_s^\pi dW_s.$$

By assumption the stochastic integral with respect to W is a true martingale and has expectation zero. Hence we obtain

$$E\left(\widehat{J}(t+h, V_{t+h}^{\pi}) \mid V_t = x\right) = \widehat{J}(t,x) + E\left(\int_t^{t+h} \hat{J}_t(s, V_s^{\pi}) + \mathcal{L}^{\pi}\hat{J}(s, V_s^{\pi}) ds\right)$$

Combining this with the inequality (7.9) we obtain

$$E\left(\int_t^{t+h} F(s, V_s^{\pi}, \pi_s) + \hat{J}_t(s, V_s^{\pi}) + \mathcal{L}^{\pi}\hat{J}(s, V_s^{\pi})\, ds\right) \leqslant 0. \tag{7.13}$$

Dividing (7.13) by h and letting $h \to 0$ in (7.13) we obtain from the fundamental theorem of calculus that for $\lambda \in \mathcal{K}$

$$F(t,x,\lambda) + \hat{J}_t(t,x) + \mathcal{L}^{\lambda}\hat{J}(t,x) \leq 0.$$

Moreover, we get from the second inequality in the dynamic programming principle that

$$\sup_{\pi \in \mathcal{A}} E\left(\int_t^{t+h} F(s, V_s^{\pi}, \pi_s) + \hat{J}_t(s, V_s^{\pi}) + \mathcal{L}^{\pi}\hat{J}(s, V_s^{\pi})\, ds\right) = 0\,.$$

For $h \to 0$ we thus obtain that

$$\hat{J}_t(t,x) + \sup_{\lambda \in \mathcal{K}}\{F(t,x,\lambda) + \mathcal{L}^{\lambda}\hat{J}(t,x)\} = 0.$$

Summarizing, we have derived the following

Proposition 7.5. *Under the regularity assumption **A1)** the optimal value function $\hat{J}(t,x)$ solves the following nonlinear PDE, known as dynamic programming or HJB (Hamilton-Jacobi-Bellman) equation:*

$$\hat{J}_t(t,x) + \sup_{\lambda \in \mathcal{K}}\{F(t,x,\lambda) + \mathcal{L}^{\lambda}\hat{J}(t,x)\} = 0,\ (t,x) \in [0,T] \times \mathbb{R} \tag{7.14}$$

with terminal condition $\hat{J}(T,x) = u(x)$. The supremum in (7.13) *is obtained by any optimal strategy $\hat{\pi}(t,x)$.*

Remark 7.6. Often one has to deal with stochastic minimization problems. In these case the above results remain valid if the 'sup' in the HJB equation is replaced by an 'inf'.

A verification result. Next we give a so-called verification theorem. This result tells us that if we have a function $H \in \mathcal{C}^{1,2}$ solving (7.14) with the right terminal condition, then $H = \hat{J}$, and $\hat{\pi}$ is given by the maximizer in (7.14).

Theorem 7.7. *Suppose that we have two functions $H(t,v)$ and $g(t,v) \in \mathcal{A}$ such that*

- *H is sufficiently integrable (see below) and solves the HJB equation*

$$H_t(t,x) + \sup_{\lambda \in \mathcal{K}}\{F(t,x,\lambda) + \mathcal{L}^{\lambda}H(t,x)\} = 0, \quad H(T,x) = u(x)$$

- *For each (t,x) fixed $g(t,x)$ is a maximizer of the expression $\mathcal{K} \ni \lambda \mapsto F(t,x,\lambda) + \mathcal{L}^{\lambda}H(t,x)$.*

Then it holds that $H = \hat{J}$. Moreover $g(\cdot)$ is an optimal feedback control strategy.

Proof. An application of Itô's formula shows that for an arbitrary strategy $\pi \in \mathcal{A}$ one has

$$u(V_T^\pi) = H(T, V_T^\phi) = H(t,x) + \int_t^T H_t(s, V_s^\pi) + \mathcal{L}^{\pi_s} H(s, V_s^\pi)\, ds$$
$$+ \int_t^T \sigma^\pi H_x(s, V_s^\pi) dW_s.$$

The HJB equation implies that $H_t(s, V_s^\pi) + \mathcal{L}^{\pi_s} H(s, V_s^\pi) + F(s, V_s^\pi, \pi_s) \le 0$. Hence we get that

$$\int_t^T H_t(s, V_s^\pi) + \mathcal{L}^{\pi_s} H(s, V_s^\pi)\, ds \le -\int_t^T F(s, V_s^\pi, \pi_s) ds$$

and thus

$$H(t,x) \ge \int_t^T F(s, V_s^\pi, \pi_s) ds + u(V_T^\pi) - \int_t^T \sigma^\pi H_x(s, V_s^\pi) dW_s$$

Assuming enough integrability so that the integral wrt dW is a martingale (this is the integrability condition on $H(\cdot)$) we get by taking expectations that

$$H(t,x) \ge E\Big(\int_t^T F(s, V_s^\pi, \pi_s) ds + u(V_T^\pi) \mid V_t = x \Big) = J(t, x, \pi),$$

and hence also $H(t,x) \ge \hat{J}(t,x)$. For $\pi = g(t,x)$ we obtain that equality holds in the above relations and hence we get

$$H(t,x) = J(t, x, g(\cdot)) \le \hat{J}(t,x).$$

Combining these inequalities gives $H \equiv \hat{J}$ and the optimality of $g(\cdot)$. □

Working with the dynamic programming equation. In order to solve the HJB equation one proceeds in two steps.

- Step 1. Solve the static optimization problem $sup_\pi \mathcal{L}^\pi H(t,x)$, making an educated guess about structural properties of H. This gives $\hat{\pi}$ in terms of H and its derivatives.
- Step 2. Substitute the solution $\hat{\pi}$ from Step 1 into the HJB equation and try to solve the resulting highly nonlinear equation, verifying the educated guess made in Step 1.

Note that explicit solutions to HJB equations exist only in very exceptional cases.

Example 7.8 (The stochastic regulator)**.** This is a classical example of a control problem that can be solved with the approach sketched above. The problem is as follows

$$\text{minimize} E\Big(\int_0^T \rho \pi_s^2 ds + X_T^2 \Big) \quad \text{subject to } dX_t = (aX_t + \pi_t)dt + dW_t \tag{7.15}$$

Here $a \in \mathbb{R}$ and $\rho > 0$ are given parameters. The control π_t can take arbitrary values, that is we take $\mathcal{K} = \mathbb{R}$. The interpretation of (7.15) is as follows: X_t represents the location of a particle that is moving randomly. The controller wants to steer the particle by a proper choice of π_t so that it is close to the origin at the terminal time T. However, he has to balance the deviation from the origin at T against the 'fuel cost' of his strategy as given by $\int_0^T \rho \pi_s^2 ds$.

The generator of X for given π is $\mathcal{L}^\pi f(x) = (ax+\pi) f_x + \frac{1}{2} f_{xx}$ so that the HJB equation is

$$\widehat{J}_t(t,x) + \inf_{\pi \in \mathbb{R}} \left\{ \rho \pi^2 + (ax+\pi)\widehat{J}_x(t,x) + \frac{1}{2}\widehat{J}_{xx}(t,x) \right\} = 0, \quad \widehat{J}(T,x) = x^2 \,.$$

i) We first consider the static problem $\min_{\pi\in\mathbb{R}}\{\rho\pi^2 + (ax+\pi)\widehat{J}_x(t,x)\}$. This is a quadratic minimization problem and since $\rho > 0$ the minimum is given by the FOC, that is we obtain

$$\widehat{\pi}(t,x) = -\frac{1}{2\rho}\widehat{J}_x(t,x).$$

ii) In order to determine $\widehat{\pi}(t,x)$ we need to find the value function $\widehat{J}_x(t,x)$. Since the terminal condition is quadratic we conjecture that $\widehat{J}$ is of the form $\widehat{J}(t,x) = c(t)x^2 + d(t)$ for deterministic functions $c(t)$ and $d(t)$. The terminal condition immediately gives $c(T) = 1$ and $d(T) = 0$. Our conjecture gives $\widehat{J}_x(t,x) = 2c(t)x$ so that $\widehat{\pi}$ becomes $\widehat{\pi}(t,x) = -\frac{c(t)}{\rho}x$. Moreover, under our conjecture for the form of $\widehat{J}$ we get

$$\widehat{J}_t(t,x) = c'(t)x^2 + d'(t) \text{ and } \widehat{J}_{xx}(t,x) = 2c(t).$$

Hence the HJB equation becomes

$$c'(t)x^2 + d'(t) + \rho(\frac{-c(t)x}{\rho})^2 + \big(ax + \frac{-c(t)x}{\rho}\big)2c(t)x + \frac{1}{2}2c(t) = 0$$

Simplifying and collecting the terms with an x^2 and the terms independent of x separately gives

$$x^2\Big(c'(t) - \frac{c^2(t)}{\rho} + 2ac(t)\Big) + (d'(t) + c(t)) = 0\,.$$

It follows that each of the brackets needs to be zero. This gives the following ODE for c: $c'(t) - c^2(t)/\rho + 2ac(t) = 0$, with terminal condition $c(T) = 1$. This is a so-called Ricatti equation that permits an explicit (but somewhat messy) solution; we omit the details. Given $c(t)$ we get from the second bracket that $d'(t) = -c(t)$ and hence $d(t) = \int_t^T c(s)ds$.

Example 7.9 (Limitations of classical solutions to HJB equation.)**.** In this example we show that the value function of a control problem is not always a classical solutions of the HJB equatio. Consider the state process S with dynamics $dS_t^\pi = \pi_t S_t^\pi dW_t$, the control set $\mathcal{K} = [0,\bar{\sigma}]$ and the problem

$$\max_{\pi\in\mathcal{A}} E\Big(h(S_T^\pi)\Big).$$

The financial interpretation is to find the worst-case price of the terminal-value claim $h(S_T)$ in a setup with uncertain volatility that takes values in the interval $[0,\bar{\sigma}]$. If h is concave, Jensen's inequality shows that the optimal strategy (the strategy that leads to the highest price of the claim) is $\pi_t \equiv 0$ (zero volatility) with value function $\widehat{J(t,S)} = h(S)$. This need not be differentiable; consider for instance the payoff $h(S) = \min\{S,K\}$).

To treat such problems a weaker solution concept is required. A possibility is to consider so-called *viscosity solutions*; see for instance Fleming & Soner (2006) or Pham (2009) for textbook treatments.

7.1.3 Back to the portfolio-optimization problem

Finally we apply Proposition 7.5 to the portfolio optimization problem. For simplicity we ignore the case of intermediate dividend payments, that is we put $C = 0$ in (7.3). In order to find a solution of the HJB equation (7.14) we proceed in two steps.

Step 1. In the first step we solve the static optimization problem $sup_\pi \mathcal{L}^\pi H(t,x)$. In this way we express the optimal strategy in terms of the candidate value function H and its derivatives. We have

$$\mathcal{L}^\pi H(t,x) = rxH_x + (\mu - r)x\pi H_x + \frac{1}{2}\sigma^2 x^2 \pi^2 H_{xx} \tag{7.16}$$

Assume now that H inherits the qualitative properties of the the utility function u via the terminal condition $H(T,x) = u(x)$, so that $H_x > 0$ and $H_{xx} < 0$. . In that case (7.16) is maximized by setting

$$\pi(t,x) = -\frac{(\mu - r)H_x}{\sigma^2 x H_{xx}}(t,x).$$

Step 2. (Solving the HJB equation) Plugging our candidate for $\hat{\pi}$ into the HJB-equation (7.14), we obtain the PDE

$$H_t(t,x) + rxH_x - \frac{(\mu-r)^2}{\sigma^2}\frac{H_x^2}{H_{xx}} + \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{H_x^2}{H_{xx}} = 0$$

This is a highly nonlinear equation. In order to solve this equation we conjecture that the solution is of the form $H(t,x) = f(t)x^\gamma/\gamma$, $\gamma \neq 0$ and $H(t,x) = f(t) + \ln(x)$ for $\gamma = 0$. With this conjecture we have that $H_x/(xH_{xx}) = -1/(1-\gamma)$ (recall that $\gamma < 1$ by assumption) and the optimal strategy takes the form

$$\hat{\pi}(t,x) = \frac{\mu - r}{(1-\gamma)\sigma^2}\,. \tag{7.17}$$

This form of the optimal strategy is quite intuitive: the investor should hold a long position if the growth rate μ is larger than the risk free rate r and he should hold a short position in the stocks whenever $\mu < r$. The size of his position is inversely proportional to his risk aversion coefficient $(1-\gamma)$ and to the instantaneous variance σ^2 of the stock. It remains to show that H is in fact a solution of the HJB equation for a proper choice of $f(\cdot)$. We consider only the case $\gamma \neq 0$. For $\hat{\pi}$ as in (7.17) the HJB equation becomes $H_t + \mathcal{L}^{\hat{\pi}} H = 0$. Substituting $H(t,x) = f(t)x^\gamma/\gamma$ we thus get

$$f'(t)(x^\gamma)/\gamma + f(t)\left(rxx^{\gamma-1} + \frac{(\mu-r)^2}{(1-\gamma)\sigma^2}xx^{\gamma-1} - \frac{1}{2}\frac{(\mu-r)^2}{(1-\gamma)\sigma^2}xx^{\gamma-1}\right) = 0.$$

Dividing this equation by x^γ shows that f satisfies a linear ODE. The terminal condition $H(T,x) = u(x)$ gives the terminal condition $f(T) = 1$, and f is easily computed. Hence we have in fact found a solution of the HJB equation of the form $H(t,x) = f(t)x^\gamma/\gamma$, and $\hat{\pi}$ given in (7.17) is the optimal strategy by the verification theorem.

For further information on stochastic control theory we refer to the recent textbook Pham (2009).

Chapter 8

Term-Structure Modelling

In this chapter we give an introduction to the modelling of interest-rate risk. Our presentation is largely based on Björk (2004); further information can be found in the textbook Filipović (2009).

8.1 Bonds and Interest-Rates

We begin by describing the market for interest-rate related products at a given point in time t.

Zero coupon bond. A zero coupon bond with maturity $T > t$, also called T-bond guarantees the holder €1 to be paid at the maturity date T. The price in $t < T$ is denoted by $p(t,T)$. Since we assume that bonds do not default it holds that $p(T,T) = 1$. The family $p(t,T)$, $T \geq t$ describes the term-structure of interest rates at time t. We assume that

- There is a frictionless market for T-bonds for all $T > t$
- For fixed t the mapping $T \mapsto p(t,T)$ is continuously differentiable; we write $p_T(t,T) = \partial_T p(t,T)$.

Note that the mapping $t \mapsto p(t,T)$ (the price trajectory of a bond with fixed maturity date) is usually quite irregular with sample path properties similar to those of Brownian motion.

Forward price of bonds. Consider time points $t < S < T$. The price, contracted in t, to be paid in S for €1 at time T is the forward price of the T-bond with maturity S, denoted by $F(t,S,T)$. It is well-known that $F(t,S,T) = p(t,T)/p(t,S)$. Alternatively, if we make a contract at t to invest €1 over the time period $[S,T]$ we obtain an amount of $1/F(t,S,T) = p(t,S)/p(t,T)$ at T.

From this relation we can define various interest rates.

Definition 8.1. (Interest rates, continuous compounding) Let $t < S < T$.

1. The (continuously compounded) forward rate over $(S,T]$, contracted at t, $R(t;S,T)$ is defined by the equation

$$\exp((T-S)R(t;S,T)) = \frac{p(t,S)}{p(t,T)},$$

i.e. as continuously compounded return on the forward investment. One has

$$R(t;S,T) = -\frac{\ln p(t,T) - \ln p(t,S)}{T-S}.$$

2. The (continuously compounded) spot rate or yield with maturity T, contracted at t is defined to be

$$R(t,T) := R(t;t,T) = -\frac{\ln p(t,T)}{T-t}.$$

3. The instantaneous forward rate with maturity T, contracted at t, is defined by

$$f(t,T) := \lim_{S\to T} R(t;S,T) = -\partial_T \ln p(t,T).$$

4. The (instantaneous) short-rate of interest is $r(t) = f(t,t)$, i.e. $r(t) = -\partial_T \ln p(t,T)|_{T=t}$.

As usual we define the *money market account* account process by $B_t = \exp(\int_0^t r_s ds)$, so that B has dynamics $dB_t = r_t B_t dt$ with initial value $B_0 = 1$.

Relation between $f(t,\cdot)$ and $p(t,\cdot)$. We get from the fundamental theorem of calculus

$$\ln p(t,T) = \ln p(t,S) + \int_S^T \partial_T \ln p(t,u)du = \ln p(t,S) - \int_S^T f(t,u)du.$$

and therefore $p(t,T) = p(t,S)\exp\left(-\int_S^T f(t,u)du\right)$. In particular

$$p(t,T) = \exp(-\int_t^T f(t,u)du), \tag{8.1}$$

as $p(t,t) = 1$. This shows that there is a one-to-one relation between the family $p(t,\cdot)$ of bond prices and the family $f(t,\cdot)$ of instantaneous forward rates.

Dynamic Modelling. Term-structure modelling is the task of constructing a probabilistic model for the evolution of the family of zero coupon bond price processes $p(\cdot,T)$. There are a number of requirements for such a model.

- $p(\cdot,T)$ needs to satisfy the final condition $p(T,T) = 1$.
- Absence of arbitrage between the security prices $\{p(\cdot,T), T \geq t\}$ (a challenge as we are dealing with many price processes simultaneously).
- Calibration. At the initial date t_0 the price $p(t_0,T)$ has to be consistent with the bond prices or interest rates observed in the market at time t_0).
- The short rate $r(t)$ is a nominal interest rate and should be nonnegative (but this requirement is sometimes relaxed).

In these lecture notes we consider two approaches for term-structure modelling, so-called *short-rate models* and *forward-rate-* or *HJM models*.

8.2 Short-Rate Models

8.2.1 Martingale Modelling and Term-Structure Equation

Consider a filtered probability space $(\Omega, \mathcal{F}, Q)$, $(\mathcal{F}_t)$ supporting an adapted process $(r_t)_{t\geq 0}$ which models the instantaneous short rate of interest. We view Q as risk-neutral measure and define the price of a T-bond at $t \leq T$ by the risk-neutral pricing rule, that is we put

$$p(t,T) = B_t \mathbb{E}\left(\frac{1}{B_T}\middle|\mathcal{F}_t\right) = \mathbb{E}\left(e^{-\int_t^T r_s ds}\middle|\mathcal{F}_t\right). \tag{8.2}$$

The approach of modelling the dynamics of traded securities directly under some martingale measure (and not under the historical measure P) is termed *martingale modelling*; it is motivated by the observation that at least in a complete market only the risk-neutral measure Q matters for the pricing of derivative securities. Martingale modelling has become very popular. The approach has the following advantages:

- The approach ensures that the model is arbitrage-free.
- Bond prices are automatically consistent with the terminal condition $p(T,T) = 1$.

Disadvantages/ problems:

- Bond prices defined via (8.2) are not automatically consistent with the prices observed in the market at time t; for this one needs to adjust parameters in the dynamics of the short rate process (calibration), which can be difficult from a computational viewpoint.
- The approach has conceptual difficulties in incomplete markets (but that is less relevant in the context of term-structure models).

Assumption 8.2. We assume that under the fixed martingale measure Q the short-rate has dynamics of the form $dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t$ for a Q-Brownian motion W.

Under Assumption 8.2 the short-rate is a Markov process, and we have for $p(t,T)$ as defined in (8.2)

$$p(t,T) = \mathbb{E}\left(\exp\left(-\int_t^T r_s ds\right)\middle|\mathcal{F}_t\right) = E_{t,r_t}\left(\exp\left(-\int_t^T r_s ds\right)\right), \tag{8.3}$$

and the right side is obviously a function of t and r_t which we denote by $F^T(t, r_t)$.

Proposition 8.3 (Term-structure equation.)**.** *The function F^T solves the parabolic PDE*

$$\partial_t F + \mu\, \partial_r F + \frac{1}{2}\sigma^2 \partial_{rr} F = rF,\ t < T \tag{8.4}$$

with terminal condition $F^T(T, r) = 1$.

Proof. The result follows immediately from the Feynman-Kac formula, see equation (5.17). □

Examples for short-rate dynamics. The following short-rate models are frequently used in the literature.

- Vasicek-model $dr_t = (b - ar_t)dt + \sigma dW_t$, $a, b > 0$.
- CIR-model $dr_t = (b - ar_t)dt + \sigma\sqrt{r_t}dW_t$.
- Hull-White $dr_t = (\varphi(t) - ar_t)dt + \sigma dW_t$.
- Extended CIR: $dr_t = (\varphi(t) - ar_t)dt + \sqrt{r_t}\sigma dW_t$.
- Black-Derman-Toy $dr_t = a(t)r_t dt + \sigma r_t dW_t$.

8.2.2 Affine Term Structure

Next we look for conditions such that the term structure equation is easy to solve. The main result in this direction is the existence of an affine term structure.

Definition 8.4. A short-rate model has an affine term structure (ATS), if the bond-prices are of the form

$$p(t,T) = F^T(t, r_t) = \exp(A(t,T) - B(t,T)r_t) \tag{8.5}$$

for deterministic functions $A(t,T)$ and $B(t,T)$, $0 \le t \le T$.

Note that in a model with an ATS the continuously compounded yield $R(t,T) = -\frac{1}{T-t}\ln p(t,T)$ is an affine function of r_t; this explains the name. It is easy to give sufficient conditions for the existence of an ATS.

Proposition 8.5. *Assume that μ and σ^2 are of the form*

$$\mu(t,r) = \alpha(t)r + \beta(t),\ \sigma^2(t,r) = \gamma(t)r + \delta(t) \tag{8.6}$$

Then the model has an ATS (8.5) where A and B satisfy the ODE-system

$$\begin{aligned} B_t(t,T) + \alpha(t)B(t,T) - \frac{1}{2}\gamma(t)B^2(t,T) &= -1 \\ A_t(t,T) - \beta(t)B(t,T) + \frac{1}{2}\delta(t)B^2(t,T) &= 0, \end{aligned} \tag{8.7}$$

with terminal conditions $B(T,T) = A(T,T) = 0$.

Proof. Assume that F^T is of the form (8.5). This gives,

$$\begin{aligned} \partial_t F^T(t,r) &= (A_t(t,T) - B_t(t,T)r)F^T(t,r), \\ \partial_r F^T(t,r) &= -B(t,T)F^T(t,r), \\ \partial_{rr} F^T(t,r) &= B^2(t,T)F^T(t,r). \end{aligned}$$

Plugging these terms into the term-structure equation gives after division by $F^T(t,r)$

$$A_t(t,T) - B_t(t,Tr - \mu(t,r)B(t,T) + \frac{1}{2}\sigma^2(t,r)B^2(t,T) = r. \tag{8.8}$$

If μ and σ^2 are affine in r, we get by substituting (8.7) into (8.8)

$$\big(A_t(t,T) - \beta B(t,T) + \frac{1}{2}\delta B^2(t,T)\big) - r\big(B_t(t,T) + 1 + \alpha(t,T) - \gamma B^2(t,T)\big) = 0.$$

This equation holds for all values of r if and only if both brackets are zero, which gives (8.7). The terminal condition on A and B ensures that $F^T(T,T) = 1$. □

Example 8.6. (Term-structure equation in the Vasicek-model.) Here the ODE-system (8.7) becomes, as $\alpha = -a$, $\beta = b$, $\gamma = 0$, $\delta = \sigma^2$,

$$B_t(t,T) - aB(t,T) = -1,\ A_t(t,T) = bB(t,T) - \frac{1}{2}\sigma^2 B^2(t,T).$$

Since the ODE for B is linear, one has, as $B(T,T) = 0$,

$$B(t,T) = \frac{1}{a}(1 - \exp\left(-a(T-t)\right)),$$

and, by integration, $A(t,T) = -\int_t^T bB(s,T)ds + \frac{\sigma^2}{2}\int_t^T B^2(s,T)ds\,.$

8.2.3 Calibration

All short-rate models considered so-far have dynamics which depend on a number of unknown parameters. In order to apply the model these parameters need to be determined. Here two different approaches can be distinguished.

a) Statistical estimation from a time series $(r_s)_{s\le t}$ of the past values of the short-rate. Here we face the following conceptual problem: in order to compute the bond prices we use the Q-dynamics of the short-rate, whereas the statistical estimation gives a estimate of the parameters under P, and the drift of the short-rate may change in the transition from P to Q.

b) Calibration to market prices. Schematically, the approach is as follows: Denote current time by t_0.

- Fix a concrete short-rate model, say, the CIR model with parameter vector $\theta = (b, a, \sigma)$.
- Solve the term-structure equation for all maturities T, denote the solution by $F^T(t_0, r_{t_0}; \theta)$.
- Go to the market and obtain observed prices $\{p^*(t_0,T_1),\dots,p^*(t_0,T_m)\}$ for maturities $T_1,\dots,T_m$ and obtain moreover an estimate of the current short-rate r_{t_0}.
- Determine θ^* as solution of the minimization problem

  $$\min_\theta \sum_{i=1}^m w_i\left(p^*(t_0,T_i) - F^{T_i}(t_0, r^*_{t_0};\theta)\right)^2, \tag{8.9}$$

  where $w_1,\dots,w_n$ is a vector of weights.

This approach is widely adapted in practice, as it is less 'subjective' and often easier than statistical estimation. Moreover the approach (approximately) aligns model and market prices which is important for the use of market-consistent valuation methods. Nonetheless there are a number of problems as-well.

- On the practical side, the model may not be flexible enough so that model- and market prices diverge widely even for 'optimal' θ^*.
- With very little price observation there may be many solutions to (8.9), i.e. θ can not be determined from observable prices.
- The parameter vector θ^* may fluctuate a lot over time, whereas in pricing it is assumed that this parameter is constant.
- Time series properties of the short rate (or of observed bond prices) are ignored completely.

8.3 HJM-Models

Basic approach. In the HJM-approach to term structure modelling one models simultaneously the dynamics of all forward-rates $f(t,T)$, $T \geq t$ and computes bond-price dynamics from the relation

$$p(t,T) = \exp\left(-\int_t^T f(t,u)du\right) \tag{8.10}$$

(recall that $f(t,T) = -\partial_T \ln p(t,T)$). In this way the calibration to the observed bond prices is ensured by design, as the current forward-rate curve $f^*(0,T)$ is taken as initial value of the forward-rate dynamics; moreover (8.10) ensures that $p(T,T) = \exp(0) = 1$. On the other hand, it is a priori not clear if the model generated in this way is actually arbitrage-free, as one models infinitely many securities (all bond prices), but has typically only finitely many sources of randomness in the model; this issue is taken up below.

Assumption 8.7. (Forward-rate dynamics) Given a probability space $(\Omega, \mathcal{F}, P)$, $(\mathcal{F}_t)$, P the historical measure, let W be a d-dimensional Brownian motion on this space. Then for each fixed T the forward rate $f(t,T)$ has a stochastic differential of the form

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_t \tag{8.11}$$

where $\alpha(\cdot,T)$ and $\sigma(\cdot,T)$ are adapted process with values in $\mathbb{R}$ respective $\mathbb{R}^d$, and where $\int_0^T \sigma(t,T)dW_t$ is shorthand for $\sum_{i=1}^d \int_0^t \sigma_i(t,T)dW_{t,i}$. Moreover α and σ are continuously differentiable in the T-variable.

In order to give conditions on $\alpha(t,T)$ and $\sigma(t,T)$ in (8.11) ensuring that the model is free of arbitrage we need to compute the dynamics of the bond prices using (8.10).

Proposition 8.8 (Short-rate dynamics and bond price dynamics in HJM)**.** *If the family of forward rates satisfies Assumption 8.7, the short rate $r(t) = f(t,t)$ has the differential*

$$dr_t = a(t)dt + b(t)dW_t, \quad \text{with } a(t) = \partial_T f(t,t) + \alpha(t,t) \text{ and } b(t) = \sigma(t,t). \tag{8.12}$$

Moreover, $p(t,T) = \exp\left(-\int_t^T f(t,u)du\right)$ satisfies

$$dp(t,T) = p(t,T)\{r(t) + A(t,T) + \frac{1}{2}\left\|S(t,T)\right\|^2\}dt + p(t,T)S(t,T)dW_t, \text{ with} \tag{8.13}$$

$$A(t,T) = -\int_t^T \alpha(t,u)du, \text{ and } S(t,T) = -\int_t^T \sigma(t,u)du \tag{8.14}$$

Proof. As $r(t) = f(t,t)$ we get from (8.11)

$$r(t) = f(0,t) + \int_0^t \alpha(s,t)ds + \int_0^t \sigma(s,t)dW_s. \tag{8.15}$$

Now write $\alpha(s,t) = \alpha(s,s) + \int_s^t \partial_T \alpha(s,u)du$, and similarly for σ. Substitution into (8.15) gives

$$\begin{aligned} r(t) = &\int_0^t \alpha(s,s)ds + \int_0^t \sigma(s,s)dW_s \\ &+ f(0,t) + \int_0^t \int_s^t \partial_T \alpha(s,u)duds + \int_0^t \int_s^t \partial_T \sigma(s,u)dudW_s. \end{aligned} \tag{8.16}$$

Changing the order of integration we get

$$\int_0^t \int_s^t \partial_T \alpha(s,u)duds = \int_0^t \int_0^t 1_{\{u>s\}} \partial_T \alpha(s,u)duds = \int_0^t \int_0^u \partial_T \alpha(s,u)dsdu, \tag{8.17}$$

and similarly $\int_0^t \int_s^t \partial_T \sigma(s,u)dudW_s = \int_0^t \int_0^u \partial_T \sigma(s,u)dW_s du$. Moreover, $f(0,t) = r(0) + \int_0^t \partial_T f(0,u)du$. On the other hand one has

$$\partial_T f(u,u) = \partial_T f(0,u) + \int_0^u \partial_T \alpha(s,u)ds + \int_0^u \partial_T \sigma(s,u)dW_s.$$

Hence the sum in (8.16) equals $r(0) + \int_0^t \partial_T f(u,u)du$, which proves he first part of the proposition.

In order to identify the bond price dynamics (8.13) and (8.14) one argues as follows: Define $Y(t,T) := -\int_t^T f(t,s)ds$ so that $p(t,T) = \exp(Y(t,T))$. Recall that under Assumption 8.7,

$$f(t,s) = f(0,s) + \int_0^t \alpha(u,s)du + \int_0^t \sigma(u,s)dW_u.$$

This gives, after changing the order of integration,

$$Y(t,T) := -\int_t^T f(0,s)ds - \int_0^t \int_t^T \alpha(u,s)dsdu - \int_0^t \int_t^T \sigma(u,s)dsdW_u.$$

Now we can decompose these integrals into

$$-\int_0^T f(0,s)ds - \int_0^t \int_u^T \alpha(u,s)dsdu - \int_0^t \int_u^T \sigma(u,s)dsdW_u \tag{8.18}$$

$$+\int_0^t f(0,s)ds + \int_0^t \int_u^t \alpha(u,s)dsdu + \int_0^t \int_u^t \sigma(u,s)dsdW_u$$

$$= Y(0,T) + \int_0^t A(u,T)du + \int_0^t S(u,T)dW_u$$

$$+\int_0^t f(0,s)ds + \int_0^t \int_0^s \alpha(u,s)duds + \int_0^t \int_0^s \sigma(u,s)dW_u ds, \tag{8.19}$$

where we have used the definition of A and S in (8.14) and a similar argument as in (8.17). Now, (8.19) is obviously equal to $\int_0^t f(s,s)ds = \int_0^t r(s)ds$, and we get that $dY(t,T) = \{r(t) + A(t,T)\}dt + S(t,T)dW_t$. The proof of (8.13) is completed by applying the Ito-formula to $p(t,T) = \exp(Y(t,T))$. □

Absence of arbitrage Recall that we consider a market with infinitely many assets but only finitely many sources of uncertainty, namely $W_1, \dots, W_d$. Hence we need special conditions on the forward rate dynamics in order to ensure that the model is free of arbitrage.

Proposition 8.9. *Suppose that there is a process* $\lambda_t = (\lambda_{t,1}, \dots, \lambda_{t,d})'$ *such that for all* $T > 0$ *and all* $0 \le t \le T$

$$\alpha(t,T) = \sigma(t,T) \int_t^T \sigma(t,s)' ds + \sigma(t,T)\lambda(t) \tag{8.20}$$

and such that $Z_t = \exp\left(\int_0^t -\lambda_s dW_s - \frac{1}{2}\int_0^t \|\lambda_s\|^2 ds\right)$ *is a true martingale. Then the model is free of arbitrage.*

Note that Condition (8.20) is a restriction on the drift of the forward rates, as this equation has to hold simultaneously for all maturity dates T.

Proof. According to Propostion 8.8, bond prices are of the form

$$dp(t,T) = p(t,T)r(t)dt + p(t,T)\left(A(t,T) + \frac{1}{2}\|S(t,T)\|^2\right)dt + p(t,T)S(t,T)dW_t.$$

Define a new measure Q by $\frac{dQ}{dP}|\mathcal{F}_T = Z_T$; Q is well-defined as Z is a true martingale. Moreover, $\widetilde{W}_t = W_t + \int_0^t \lambda_s ds$ is a Q-Brownian Motion by the Girsanov theorem. Now note that under Q the finite-variation part of $p(\cdot,T)$ equals

$$\int_0^t p(s,T)\left\{r(s) + A(s,T) + \frac{1}{2}\|S(s,T)\|^2 - S(s,T)\lambda_s\right\} ds. \tag{8.21}$$

We get from (8.20) and the definition of $S(s,T)$ in (8.14) that

$$\begin{aligned} S(s,T)\lambda_s &= -\int_s^T \sigma(s,u)\lambda_s du \text{ (by (8.14))} \\ &= -\int_s^T \left\{\alpha(s,u) - \sigma(s,u)\int_s^u \sigma(s,\tau)' d\tau\right\} du \text{ (by (8.20))} \\ &= A(s,T) + \frac{1}{2}\|S(s,T)\|^2 . \end{aligned}$$

The last relation follows from (by (8.14)) and the observation that

$$\frac{\partial}{\partial u}\frac{1}{2}\|S(s,u)\|^2 = S(s,u)\frac{\partial}{\partial u}S(s,u)' = -\int_s^u \sigma(s,\tau)d\tau(-\sigma(s,u)').$$

Hence the finite variation part of $p(\cdot,T)$ in (8.21) equals $\int_0^t p(s,T)r(s)ds$, and discounted bond prices are Q-local martingales. □

Obviously the measure Q constructed in the proof of Proposition 8.9 is a risk-neutral measure. It is interesting to study the forward-rate and price dynamics under Q

Corollary 8.10. *Under Q the forward-rate dynamics are*

$$\begin{aligned} df(t,T) &= \left(\sigma(t,T)\int_t^T \sigma(t,s)' ds\right) dt + \sigma(t,T)d\widetilde{W}_t \\ dp(t,T) &= r(t)p(t,T)dt + p(t,T)S(t,T)d\widetilde{W}_t, \end{aligned}$$

$\widetilde{W}$ *a Q-Brownian motion.*

This follows immediately from the proof of Proposition 8.9. The fact that the risk-neutral drift of $f(t,T)$ is of the form

$$\alpha^Q(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)'ds$$

is also known as HJM drift condition.

Martingale modelling in HJM The previous corollary has shown that the Q-dynamics of bonds and forward-rates are fully determined from the forward-rate volatilies $\sigma(t,T)$. For practical purposes the HJM model is therefore used as follows:

1.) Specify the volatilities $\sigma(t,T)$ (the modelling part)

2.) The forward-rate drift $\alpha^Q(t,T)$ is specified via the HJM-drift condition.

3.) Go to the market and observe current forward rates $f^*(0,T)$.

4.) The forward-rates are given by

$$f(t,T) = f^*(0,T) + \int_0^T \alpha^Q(s,T)ds + \int_0^t \sigma(t,T)d\widetilde{W}_t.$$

5.) Bond prices are given via $p(t,T) = \exp\{-\int_t^T f(t,u)du\}$ or as solutions of

$$dp(t,T) = p(t,T)r(t)dt + p(t,T)S(t,T)d\widetilde{W}_t,$$

for $S(t,T) = -\int_t^T \sigma(t,u)du$.

Example 8.11 (The Ho-Lee model)**.** Assume that $\sigma(t,T) = \sigma$. Then the drift of the forward-rates equals $\alpha^Q(t,T) = \sigma\int_t^T \sigma ds = \sigma^2(T-t)$ and the forward rates are given by

$$\begin{aligned} f(t,T) &= f^*(0,T) + \int_0^t \sigma^2(T-s)ds + \sigma W_t \\ &= f^*(0,T) + \sigma^2 t\left(T - \frac{t}{2}\right) + \sigma W_t. \end{aligned}$$

In particular, the short-rate satisfies

$$r(t) = f(t,t) = f^*(0,t) + \sigma^2\frac{t^2}{2} + \sigma W_t,$$

or, in differential form, $dr_t = \left(\partial_T f^*(0,t) + \sigma^2 t\right)dt + \sigma dW_t$. Short-rate dynamics of this form are known as Ho-Lee model (the Vasicek or Hull-White model for $a = 0$). Note that the HJM-approach automatically ensures that the model is calibrated to the initial yield curve.

Example 8.12 (The Vasicek-model)**.** Here we take $\sigma(t,T) = \sigma e^{-a(T-t)}$ for some $a > 0$. Hence we get

$$\alpha^Q(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds = \frac{\sigma^2}{a}e^{-a(T-t)}\left(1 - e^{-a(T-t)}\right).$$

In order to identify the corresponding short-rate dynamics we use Proposition 8.8. The volatility of the short-rate dynamics is given by $\sigma(t,t) = \sigma$. Identifying the drift $\mu(t,r)$ is more involved. According to Proposition 8.8, we have $\mu(t,r_t) = \alpha^Q(t,t) + \partial_T f(t,t) = \partial_T f(t,t)$, as $\alpha^Q(t,t) = 0$. Now we have, using the forward rate dynamics

$$\partial_T f(t,t) = \partial_T f(0,T) + \int_0^t \partial_T \alpha(s,t) ds + \int_0^t \partial_T \sigma(s,t) dW_s.$$

As $\sigma(s,t) = \sigma e^{-a(t-s)}$, we get $\partial_T \sigma(s,t) = -a\sigma(s,t)$. Moreover, $\partial_T \alpha^Q(t,s) = -a\alpha^Q(t,s) + \sigma^2 e^{-2a(t-s)}$. This gives

$$\begin{aligned}\partial_T f(t,t) &= \partial_T f(0,t) - a\int_0^t \alpha(s,t) ds - a\int_0^t \sigma(s,t) dW_s + \int_0^t \sigma^2 e^{-2a(t-s)ds} \\ &= -af(t,t) + g(t)\end{aligned}$$

with $g(t) = \partial_T f(0,t) - af(0,t) + \int_0^t \sigma^2 e^{-2a(t-s)} ds$. Summarizing we thus have the following short-rate dynamics

$$dr(t) = (g(t) - ar(t))dt + \sigma dW_t,$$

so that the above form of σ leads to the Vasicek or the Hull-White model.

8.3.1 The Musiela parametrization.

still to do

8.4 Pricing and Hedging of Multi-Currency Derivatives with Interest Rate Risk

This section, which is based on Frey & Sommer (1996), discusses the valuation and hedging of non path-dependent European options on one or several underlyings in a model of an international economy which allows for both, *interest rate risk* and *exchange rate risk*. We study options on stocks, bonds, future contracts, interest rates and exchange rates; their payoff may be in any currency and a relatively complex function of one or several underlyings. We use an international economy model similar to the one introduced by Amin & Jarrow (1991) as framework of our analysis. Their model combines a fully developed stochastic theory of the term structure of interest rates in the sense of Heath, Jarrow & Morton (1992) with models for the valuation of exchange rate and stock options. The main tools of our analysis are stochastic methods and in particular the change of numeraire technique introduced in Section 6.1.4. Since pricing formulas are of limited practical use without knowledge of the corresponding hedge portfolio we present a systematic approach to computing hedging strategies. In order to illustrate the flexibility of our method we derive explicit formulas for prices and hedge portfolios for a wide range of examples containing among others currency options or guaranteed-exchange-rate options or options on interest rates.

8.4.1 The Model

Here we introduce an arbitrage-free model of an international economy that incorporates stochastic interest rates and exchange rates. This model will serve as our framework for the valuation of derivatives. We consider N countries indexed by $n \in \{0, \dots, N\}$. Country 0 will be the *domestic* country. The exchange rate between country 0 and country $n \in \{1, \dots, N\}$ will be denoted by e^n, that is e^n units of the domestic currency can be exchanged for one unit of the foreign currency. When working with only two countries we simply talk about the domestic and the foreign country and index them with d and f. The choice of the domestic country is arbitrary and depends on the particular pricing and hedging problem under consideration. We call an asset a *domestic* asset if its payoffs are denominated in the domestic currency. Notice that every asset whose payoffs are not originally denominated in this currency can be transformed into a domestic asset by translating its payoffs into the domestic currency using the corresponding exchange rate.

We assume that in all countries zero coupon bonds of all maturities $T \in [0, T_F]$ are traded. The zero coupon bond in country n with maturity date T shall be denoted by $p^n(t,T)$ for $t \in [0,T]$. By assumption $p^n(T,T) \equiv 1 \; \forall T, n$. The short rate in country n, r^n, is given by

$$r_t^n = -\frac{\partial}{\partial T}\Big|_{T=t} \ln p^n(t,T) \,. \tag{8.22}$$

By $\beta_{t,T}^n := \exp(\int_t^T r_s^n \, ds)$ we denote the savings-account of country n. Apart from zero coupon bonds we consider other primitive assets such as dividend free stocks. They are denoted by $S^{n,j}$, $0 \le n \le N$, $0 \le j \le i_n$ where $S^{n,j}$ is the price of asset j in country n.

We now introduce our model of asset price dynamics. When modelling asset price processes one usually starts from assumptions on their dynamics under the so-called historical probabilities which govern the actual evolution of asset prices. Since we are only interested in the pricing of derivatives by no-arbitrage arguments it is legitimate to model the asset price dynamics directly under a domestic risk-neutral measure Q. Under such a measure all non-dividend paying domestic assets are martingales after discounting with the domestic savings account. This implies that their drift is equal to r^0.

Assumption 8.13. Let there be given a filtered probability space $(\Omega, \mathcal{F}, Q)$, $(\mathcal{F}_t)_{t \in [0,T_F]}$ supporting a d-dimensional Brownian Motion $W = (W_t)_{0 \le t \le T_f}$. We work with the following assumptions on asset price dynamics: We put for the domestic assets

$$\begin{aligned} dp^0(t,T) &= r_t^0 p^0(t,T)dt + \eta^0(t,T)p^0(t,T)dW_t \\ dS_t^{0,j} &= r_t^0 S_t^{0,j} dt + \eta^{0,j}(t) S_t^{0,j} dW_t \,, \end{aligned} \tag{8.23}$$

and for the foreign assets

$$\begin{aligned} dp^n(t,T) &= (r_t^n - \eta^n(t,T) \cdot \eta^{e^n}(t))p^n(t,T)dt + \eta^n(t,T)p^n(t,T)dW_t \\ dS_t^{n,j} &= (r_t^n - \eta^{n,j}(t) \cdot \eta^{e^n}(t))S_t^{n,j} dt + \eta^{n,j}(t) S_t^{n,j} dW_t \,. \end{aligned} \tag{8.24}$$

Finally the dynamics of the exchange rates are given by

$$de^n(t) = (r_t^0 - r_t^n)e^n(t)dt + \eta^{e^n}(t)e^n(t)dW_t \,. \tag{8.25}$$

Here $\eta^n(t,T), \eta^{n,j}(t), \eta^{e^n}(t) : [0, T_F] \to \mathbb{R}^d$ are *deterministic* square integrable functions of time. For the bonds we require moreover that $\eta^n(t,T) = 0 \; \forall t \geq T$ and that $\eta^n(t,T)$ is smooth in the second argument.

Note that the assumption of deterministic dispersion coefficients is essential as we want to obtain explicit pricing formulas. An explicit construction of the bond price model is easily done using the HJM approach explained in the previous section. As shown by Amin & Jarrow (1991) the dynamics of asset prices and exchange rates given in Assumption 8.13 actually specify an arbitrage-free model of an international economy with Q representing a domestic risk-neutral measure. The drift terms of the exchange rate and the foreign assets are determined by absence of arbitrage considerations. As an example we derive the drift of e^n. Consider the domestic asset $Y := e^n \beta^n_{(0,\cdot)}$. By absence of arbitrage its drift must equal r^0. Using Itô's Lemma to compute the dynamics of Y it is immediate that the drift of Y equals r^0 if and only if the drift of the exchange rate equals the interest rate differential.

The volatility of asset $S^{n,j}$ is given by $\sigma^{n,j}(t) := |\eta^{n,j}(t)|$. The instantaneous correlations between the assets in our economy are given by

$$\rho(S^{n_1,j_1}, S^{n_2,j_2}) := \frac{\eta^{n_1,j_1} \cdot \eta^{n_2,j_2}}{\sigma^{n_1,j_1}\sigma^{n_2,j_2}} .$$ [1]

Only volatilities and instantaneous correlations matter for the pricing of derivatives, since they determine the law of the asset prices under the domestic risk neutral measure. In our analysis this is reflected by the fact that only inner products of the dispersion coefficients η and hence instantaneous covariances enter the pricing formulas. Nonetheless we start with independent Brownian motions and model correlations by means of the dispersion coefficients η because this facilitates the use of stochastic calculus. To compute these coefficients from the estimated instantaneous covariance matrix of the processes one may use the Cholesky decomposition of this matrix.

Finally we note that the price process of a discounted foreign asset is not a martingale under the domestic risk-neutral measure as can be seen from (8.24); hence this measure must not be used for the valuation of derivatives paying off in foreign currencies.

For our pricing theory we need to assume that the markets in our economy are complete.

Assumption 8.14. There are d traded domestic assets such that for all $t \in [0, T_F]$ the instantaneous covariance matrix of these assets is strictly positive definite.

As shown in Proposition 6.20, his assumption guarantees that every contingent claim adapted to the filtration generated by the asset prices can be replicated by a dynamic trading strategy in the d assets and the domestic savings-account. Hence the domestic risk neutral measure is unique and the price at time t of every domestic contingent claim H with single $\mathcal{F}_T$ measurable and integrable payoff H_T at time T is given by

$$H_t := E^Q \left[\left(\beta^0_{t,T}\right)^{-1} \cdot H_T \;\middle|\; \mathcal{F}_t \right] . \tag{8.26}$$

We now introduce the class of admissible underlyings for the derivative contracts considered in this section. A typical example of the kind of options we want to analyze is the guaranteed-exchange-rate call. This contract is defined by its terminal payoff $[\bar{e}S^f_T - K]^+$, where S^f_T is some primitive foreign asset and $\bar{e}$ is a guaranteed exchange rate which will be

[1] Of course similar formulas hold for bonds and exchange rates.

applied at time T to convert the price of the foreign asset into domestic currency. Now, $\bar{e}S_T^f$ is not the time T value of a traded domestic asset. However, it defines a domestic contingent claim X whose price $X_t = E^Q[(\beta_{t,T})^{-1}\bar{e}S_T^f|\mathcal{F}_t]$ is given by

$$X_t = X_0 \exp\left(\int_0^t \eta_s^X dW_s - \frac{1}{2}\int_0^t |\eta_s^X|^2 ds + \int_0^t r_s^d ds\right) \text{ with } X_0 = E^Q\left[(\beta_{0,T})^{-1} X_T\right]$$

and η_s^X a deterministic $\mathbb{R}^d$-valued function of time. We will see in section 8.4.3 below that this structure is found in many ostensibly complex option contracts. This motivates the following definition.

Definition 8.15. A domestic contingent-claim X with a single payoff X_T at a certain date T is called a *lognormal claim*[2] if its price process $(X_t)_{0\le t\le T}$ given by

$$X_t := E^Q\left[\left(\beta_{t,T}^d\right)^{-1} \cdot X_T \mid \mathcal{F}_t\right]$$

admits a representation of the form

$$X_t = X_0 \cdot \exp\left(\int_0^t \eta_s^X dW_s - \frac{1}{2}\int_0^t |\eta_s^X|^2 ds + \int_0^t r_s^d ds\right) \tag{8.27}$$

with some constant X_0 and with *deterministic* dispersion coefficients $\eta^X : [0,T] \to \mathbb{R}^d$.

Remark 8.16. The main restriction made in the definition of a lognormal claim is the assumption of η^X being deterministic. In fact, whenever X_T is strictly positive,

$$X_t := E^Q\left[\left(\beta_{t,T}^d\right)^{-1} X_T \mid \mathcal{F}_t\right]$$

is always of the form (8.27) with possibly stochastic "volatility" η^X, as can easily be shown by means of the martingale representation theorem (Theorem 6.19). Note that the solution of the SDE $dX_t = r_t^0 X_t dt + \eta_t^X X_t dW_t$ is given by

$$X_t := X_0 \exp\left(\int_0^t \eta_s^X dW_s - 1/2\int_0^t |\eta_s^X|^2 ds + \int_0^t r_s^0 ds\right) .$$

Hence under our assumption on asset price dynamics every primitive domestic asset, interpreted as contingent claim with payoff equal to the asset's price at time T, is a lognormal claim. However, the class of contingent claims that satisfy Definition 8.15 is much larger. For instance products and quotients of lognormal claims remain lognormal claims.

8.4.2 Exchange Options on Lognormal Claims

Next we give a rather general theorem which leads to a unified treatment of the pricing of European options on various underlyings such as foreign and domestic zero coupon bonds, foreign or domestic stocks or forward and future contracts on foreign and domestic assets.

[2] This name is motivated by the fact that X_T is lognormally distributed. This is immediate if one writes $X_T = X_T/p^0(T,T)$ and then expresses the right hand side using (8.27) and the corresponding expression for $p^0(\cdot,T)$.

Theorem 8.17. *Let X, Y be lognormal claims. Consider an option to exchange X for Y at the maturity date T, i.e. a European option with payoff $[X_T - Y_T]^+$.*

1. The price process $C = (C_t)_{0\le t\le T}$ of this option is given by

$$C_t = C(t, X_t, Y_t) \quad := \quad X_t \mathcal{N}(d_t^1) - Y_t \mathcal{N}(d_t^2)$$

where $\mathcal{N}$ denotes the one-dimensional standard normal distribution function, and where d_t^1 and d_t^2 are given by

$$d_t^1 = \frac{\ln(X_t/Y_t) + \frac{1}{2}\int_t^T |\eta_s^X - \eta_s^Y|^2 ds}{\sqrt{\int_t^T |\eta_s^X - \eta_s^Y|^2 ds}}, \quad d_t^2 = d_t^1 - \sqrt{\int_t^T |\eta_s^X - \eta_s^Y|^2 ds}\,.$$

2. The hedge portfolio for this option in terms of the lognormal claims X and Y consists of

$$\delta_X^C(t) := \mathcal{N}(d_t^1) \text{ units of } X \quad \text{and } \delta_Y^C(t) := -\mathcal{N}(d_t^2) \text{ units of } Y\,.$$

Proof. The main tool in the proof is the change of numeraire technique as introduced in Section 6.1.4. We now recall a few facts from this theory. Define for a lognormal claim X a new equivalent probability measure Q^X on $\mathcal{F}_T$ by

$$\frac{dQ^X}{dQ} \quad = \quad \frac{X_T \cdot \left(\beta_{0,T}^d\right)^{-1}}{X_0}.$$

Then for every domestic asset Z whose discounted price process is a martingale under Q — that is for every asset that pays no dividends in $[0, T)$ — the process Z/X is a martingale under Q^X, i.e. Q^X is the martingale measure corresponding to the numeraire X. Moreover we have the transition formula

$$E^Q\left[\left(\beta_{t,T}^d\right)^{-1} \cdot X_T \cdot Z_T \,|\mathcal{F}_t\right] = X_t \cdot E^{Q^X}[Z_T|\mathcal{F}_t] \tag{8.28}$$

In our setup it is easy to determine the law of the asset price processes under Q^X by means of the Girsanov theorem. Applying this theorem to dQ^X/dQ immediately yields that $W_t^X := W_t - \int_0^t \eta_s^X ds$ is a Brownian Motion under Q^X.

Now it is easy to proof the first part of the theorem. According to (8.26) the price of the option is given by

$$\begin{aligned} C_t &= E^Q\left[\left(\beta_{t,T}^d\right)^{-1} [X_T - Y_T]^+ \,\Big|\, \mathcal{F}_t\right] \\ &= E^Q\left[\left(\beta_{t,T}^d\right)^{-1} X_T \cdot 1_{\{Y_T/X_T<1\}} \,\Big|\, \mathcal{F}_t\right] - E^Q\left[\left(\beta_{t,T}^d\right)^{-1} Y_T \cdot 1_{\{X_T/Y_T>1\}} \,\Big|\, \mathcal{F}_t\right] \\ &= X_t \cdot E^{Q^X}\left[1_{\{Y_T/X_T<1\}} \,\Big|\, \mathcal{F}_t\right] - Y_t \cdot E^{Q^Y}\left[1_{\{X_T/Y_T>1\}} \,\Big|\, \mathcal{F}_t\right] \end{aligned}$$

The last line follows from (8.28) if we take once X and once Y as numeraire. Now we get under Q^X for Y_T/X_T

$$\frac{Y_T}{X_T} = \frac{Y_t}{X_t} \cdot \exp\left(\int_t^T (\eta_s^Y - \eta_s^X) dW_s^X - \frac{1}{2}\int_t^T |\eta_s^Y - \eta_s^X|^2 ds\right)$$

Hence

$$Q^X\left[\frac{Y_T}{X_T}<1\,\Big|\,\mathcal{F}_t\right]=Q^X\left[\ln Y_T-\ln X_T<0\,\big|\,\mathcal{F}_t\right]$$
$$=Q^X\left[\frac{\int_t^T(\eta_s^Y-\eta_s^X)dW_s^X}{\sqrt{\int_t^T|\eta_s^Y-\eta_s^X|^2ds}}<\frac{\ln X_t-\ln Y_t+\frac{1}{2}\int_t^T|\eta_s^Y-\eta_s^X|^2ds}{\sqrt{\int_t^T|\eta_s^Y-\eta_s^X|^2ds}}\right]$$

Since η^X and η^Y are deterministic, $\int_t^T(\eta_s^Y-\eta_s^X)dW_s^X/\sqrt{\int_t^T|\eta_s^Y-\eta_s^X|^2ds}$ is a standard normally distributed random variable so that

$$Q^X\left[\frac{Y_T}{X_T}<1\,\Big|\,\mathcal{F}_t\right]=\mathcal{N}(d_t^1)\,.$$

Analogously we get $Q^Y\left[X_T/Y_T>1\,\big|\,\mathcal{F}_t\right]=\mathcal{N}(d_t^2)$, and the first part of the theorem follows.

Now we turn to the hedging part. Let $(Z)^M$ denote the (uniquely determined) martingale part of a continuous semimartingale Z. To prove the second claim we note that the proposed selffinancing hedge portfolio duplicates the option if it holds that

$$d\,(C)^M=\mathcal{N}(d_t^1)d(X)_t^M-\mathcal{N}(d_t^2)d(Y)_t^M, \tag{8.29}$$

and if moreover the value of the hedge portfolio equals the option's price for all $0\leq t\leq T$. We now check these two conditions. (i) As C_t is a function only of X_t and Y_t we get from Itô's Lemma

$$d\,(C)_t^M=\frac{\partial C}{\partial x}(t,X_t,Y_t)d\,(X)_t^M+\frac{\partial C}{\partial y}(t,X_t,Y_t)d\,(Y)_t^M\,.$$

Now following El Karoui, Myneni & Viswanathan (1992) we may compute the derivatives of the option price:

$$\begin{aligned}\frac{\partial C}{\partial x}(t,X_t,Y_t) &= E_t^Q\left[\frac{\partial}{\partial X_t}\left(\left(\beta_{t,T}^d\right)^{-1}[X_T-Y_T]^+\right)\right]\\ &= E_t^Q\left[\left(\beta_{t,T}^d\right)^{-1}1_{\{X_T\geq Y_T\}}\frac{\partial X_T}{\partial X_t}\right]\\ &= \frac{1}{X_t}E_t^Q\left[\left(\beta_{t,T}^d\right)^{-1}1_{\{X_T\geq Y_T\}}X_T\right]\end{aligned}$$

As shown in the first part of the proof this expression equals $\mathcal{N}(d_t^1)$. Similarly we get $\partial C/\partial y\,(t,X_t,Y_t)=-\mathcal{N}(d_t^2)$, and hence (8.29).

(ii) By Euler's Theorem we get from the linear homogeneity of C in X_t and Y_t

$$C_t=\frac{\partial C}{\partial x}(t,X_t,Y_t)\cdot X_t+\frac{\partial C}{\partial y}(t,X_t,Y_t)\cdot Y_t=\mathcal{N}(d_t^1)X_t-\mathcal{N}(d_t^2)Y_t\,,$$

which shows that also the second condition is satisfied. □

Whenever the lognormal claims X and Y are assets for which liquid markets exist, Theorem 8.17 is sufficient for the construction of a hedge portfolio. Otherwise we must go on and duplicate X and Y by a dynamic trading strategy. The existence of such a strategy is guaranteed by Assumption 8.14; it can be computed as in the proof of Theorem 8.17. The following observation then shows how to construct hedging strategies for C from the hedge portfolios for X and Y. Suppose that the hedge portfolios for X and Y in terms of domestic assets H_i^X and H_i^Y for which we assume the existence of liquid markets are given by

$$P_t^X = \sum_{i=1}^{L^X} \delta_i^X(t) H_i^X \quad \text{and} \quad P_t^Y = \sum_{i=1}^{L^Y} \delta_i^Y(t) H_i^Y \, .$$

Then the hedge portfolio for the exchange option on X and Y in terms of H_i^X and H_i^Y is given by

$$P_t = \sum_{i=1}^{L^X} \mathcal{N}(d_t^1) \cdot \delta_i^X(t) H_i^X - \sum_{i=1}^{L^Y} \mathcal{N}(d_t^2) \cdot \delta_i^Y(t) H_i^Y \, .$$

The application of this principle is illustrated in certain examples presented below.

8.4.3 Options on Lognormal Claims: Examples

Now we want to use Theorem 8.17 in order to price a number of practically relevant contracts.

Currency Options: The payoff of a plain vanilla currency option equals $[e_T - K]^+$. Define the domestic assets $X := e \cdot p^f(\cdot, T)$ and $Y := Kp^d(\cdot, T)$; the parameters of their price processes can be read off from the asset price dynamics and are given by $X_0 = e_0 p^f(0,T)$, $\eta^X(t) = \eta^e(t) + \eta^f(t,T)$ and $Y_0 = Kp^d(0,T)$, $\eta^Y(t) = \eta^d(t,T)$, respectively. Since $p^d(T,T) = p^f(T,T) = 1$ the option's payoff equals $[X_T - Y_T]^+$, and its price can be computed by means of Theorem 8.17. We obtain

$$C_t = e_t p^f(t,T) \mathcal{N}(d_1) - Kp^d(t,T) \mathcal{N}(d_2) \text{ with}$$

$$d_1 = \frac{\ln(e_t p^f(t,T)/p^d(t,T)) - \ln K + \int_t^T |\eta^e(s) + \eta^f(s,T) - \eta^d(s,T)|^2 ds}{\sqrt{\int_t^T |\eta^e(s) + \eta^f(s,T) - \eta^d(s,T)|^2 ds}}$$

and $d_2 = d_1 - \sqrt{\int_t^T |\eta^e(s) + \eta^f(s,T) - \eta^d(s,T)|^2 ds}$. Since we assume p^f and p^d to be traded assets we can use directly Theorem 8.17 to compute a feasible hedge portfolio.

Currency Converted Options: There are two types of currency converted options. The payoff of a *Foreign Asset/ Domestic Strike Option* equals $[e_T S_T^f - K]^+$. To deal with this claim we set $X := eS^f$ and notice that this is a lognormal claim with $X_t = e_t S_t^f$ and $\eta^X = \eta^e + \eta^{S^f}$. Next set $Y := K \cdot p^d(\cdot, T)$. Theorem 8.17 can now be directly applied to give the price and the hedging strategy of this contract. Similarly for a *Domestic Asset/ Foreign Strike Option* with payoff $[S_T^d - e_T K]^+$, where K is in foreign currency we use the lognormal claims $X := S^d$ and $Y = Kep^f(\cdot, T)$.

Guaranteed-Exchange-Rate Options: The payoff of this derivative equals $[\bar{e} S_T^f - \bar{e} K]^+$, where S^f is a foreign asset and $\bar{e}$ some predetermined exchange rate. This contract can be interpreted as an option to exchange the lognormal claims X and Y with payoff $X_T = \bar{e} S_T^f$ and $Y_T := \bar{e} K$. Whereas Y_T equals the time T value of $K \cdot \bar{e}$ units of

$p^f(\cdot,T)$, there is no traded asset whose value at T is equal to X_T. To price and hedge the option we therefore have to compute the parameters of X. We have

$$\begin{aligned} X_0 &= \bar{e}S_0^f \frac{p^d(0,T)}{p^f(0,T)} \exp\Bigg\{ \int_0^T |\eta^f(s,T)|^2 + \eta^{S^f}(s)\cdot\eta^d(s,T) + \eta^e(s)\cdot\eta^f(s,T) - \qquad (8.30) \\ &\qquad \eta^{S^f}(s)\cdot\eta^f(s,T) - \eta^f(s,T)\cdot\eta^d(s,T) - \eta^{S^f}(s)\cdot\eta^e(s)ds \Bigg\} \\ \eta_s^X &= \eta^{S^f}(s) + \eta^d(s,T) - \eta^f(s,T) \end{aligned}$$

The price of the option can now be computed by plugging these parameters into the pricing formula of Theorem 8.17. Next we want to determine the hedge portfolio for the option. As X_T is not the terminal value of a traded asset we have to go through the procedure outlined after the proof of Theorem 8.17. To replicate X_T by a dynamic trading strategy we first note that by (8.30) X_t is given by a function $\tilde{X}$ of the domestic assets $e\cdot S^f$, $p^d(t,T)$ and $e\cdot p^f(t,T)$ with derivatives $\partial\tilde{X}/\partial eS^f = \tilde{X}/(e_tS_t^f)$, $\partial\tilde{X}/\partial ep^f = -\tilde{X}/(e_tp^f(t,T))$ and $\partial\tilde{X}/\partial p^d = \tilde{X}/p^d(t,T)$. As $\tilde{X}$ is linear homogenous in the prices of these assets, an argument similar to the proof of the second part of Theorem 8.17 shows that the hedge portfolio for X equals

$$\delta_{e\cdot S^f}^X(t) = \frac{X_t}{e_t\cdot S_t^f}, \qquad \delta_{e\cdot p^f}^X(t) = -\frac{X_t}{e_t\cdot p^f(t,T)}, \qquad \delta_{p^d}^X(t) = \frac{X_t}{p^d(t,T)}\,.$$

Options on Forwards: Assume that $(\tilde{X}_t)_{0\le t\le\bar{T}}$ is the price process of a lognormal claim. Consider two points in time T_1 and T_2 with $0 < T_1 \le T_2$. The payoff in T_1 of an option with maturity date T_1 on a forward contract on $\tilde{X}$ with maturity date T_2 is given by $\left[\tilde{X}_{T_1} - p^d(T_1,T_2)K\right]^+$. For $t < T_1$ price and hedge portfolio of this contract immediately follow from Theorem 8.17, if we use the lognormal claims $\tilde{X}_t$ and $Kp^d(t,T_2)$.

Options on Interest Rates: We are mainly interested in contracts where one of the underlying assets is a foreign or domestic LIBOR rate. For a fixed $\alpha > 0$ (in practice usually $\alpha = 0.25$ or $\alpha = 0.5$) the LIBOR rate $L^n(t,\alpha)$ prevailing in country n over the period $[t,t+\alpha]$ is defined by the equation

$$(1+\alpha\cdot L^n(t,\alpha))p^n(t,t+\alpha) = 1\,,$$

that is $L^n(t,\alpha) = \alpha^{-1}(1/p^n(t,t+\alpha)-1)$.

Caps: Perhaps the most important LIBOR derivatives are caps and floors. A cap is a portfolio of caplets. The payoff of a caplet with face value V, underlying interest rate process $L^d(t,\alpha)$, level K and maturity date $T+\alpha$ equals

$$V\cdot\alpha\cdot[L^d(T,\alpha)-K]^+ = V\cdot\left[\frac{1}{p^d(T,T+\alpha)} - (\alpha K+1)\right]^+$$

As the payoff of this caplet is known already at T we may compute its present value at T which equals $V[1-(\alpha K+1)\cdot p^d(T,T+\alpha)]^+$. From this we see that the price and the hedge portfolio for caplets can be inferred directly from Theorem 8.17 if we use the lognormal claims $X = p^d(\cdot,T)$ and $Y = (\alpha K+1)\cdot p^d(\cdot,T+\alpha)$. Of course this choice of X and Y reflects the well-known fact that caplets can be considered as options on zero coupon bonds.

Appendix A

Mathematical Background

A.1 Conditional Expectation

Given a probability space $(\Omega, \mathcal{F}, P)$ and a random variable X. A priori, the best prediction for X is $E(X)$. If we have additional information about the outcome of the experiment modelled by $(\Omega, \mathcal{F}, P)$, we can give a better prediction of X. The best-possible prediction – in an L^2-sense – is the conditional expectation. We first study this idea in an elementary setting where information is modelled by a (finite) partition of Ω, which leads us to an explicit formula for the conditional expectation. In a second step we will use the properties of this elementary conditional expectation to extend the notion to general probability spaces.

A.1.1 The elementary case

Definition A.1. A set $\mathcal{A} = \{A_1, \ldots, A_n\}$ of measurable subsets of Ω with $P(A_i) > 0$ for all i is called a partition of Ω if $A_i \cap A_j = \emptyset$ for $i \neq j$ and if moreover $\Omega = A_1 \cup \cdots \cup A_n$.

Now consider a partition $\mathcal{A}$ of Ω. Suppose that we have the additional information that the result ω of our random experiment belongs to a particular subset $A_{i_0} \in \mathcal{A}$. Our best prediction for the rv X is now

$$E(X|A_{i_0}) := \frac{1}{P(A_{i_0})} E(X 1_{A_{i_0}})$$

The 'prediction-mechanism' which gives the prediction of X if we are given the additional information which set from the partition $\mathcal{A}$ actually occurs is therefore given by the random variable

$$\sum_{i=1}^{n} 1_{A_i}(\omega) E(X|A_i) = \sum_{i=1}^{n} 1_{A_i}(\omega) \frac{E(X 1_{A_{i_0}})}{P(A_i)}.$$

Example: Consider a 2-period binomial model with $P(\text{'up'}) = P(\text{'down'}) = \frac{1}{2}$. Then $E(S_2) = S_0(\frac{1}{4}u^2 + \frac{1}{2}ud + \frac{1}{4}d^2) = S_0(\frac{1}{2}u + \frac{1}{2}d)^2$. Define the partition $\mathcal{A} = \{A_1, A_2\}$ with $A_1 = \{S_1 = uS_0\}$ and $A_2 = \{S_1 = dS_0\}$, i.e. the partition is formed by the value of the stock-price at $t = 1$. Then the forecast of S_2 given the information contained in $\mathcal{A}$ is given by

$$1_{A_1}(\omega)E(S_2|A_1) + 1_{A_2}(\omega)E(S_2|A_2) = 1_{A_1}(\omega)S_0\left(\frac{1}{2}u^2 + \frac{1}{2}ud\right) + 1_{A_2}(\omega)S_0\left(\frac{1}{2}ud + \frac{1}{2}d^2\right).$$

Obviously, this forecast differs depending on whether A_1 or A_2 actually occurs in $t = 1$.

Formally, the additional information is described by the σ-field $\mathcal{F}^A$ generated by the partition $\mathcal{A}$.

Definition A.2. Given a partition $\mathcal{A} = \{A_1, \ldots, A_n\}$ of Ω. The σ-field $\mathcal{F}^A$ generated by $\mathcal{A}$ is the set of all unifications $\bigcup_{j=1}^{k} A_j$, $A_j \in \mathcal{A}$, $k \in \mathbb{N}$.

Remark: Usually the σ-field $\mathcal{F}^{\mathcal{A}}$ is defined as the smallest σ-field containing all the sets $A_1, \ldots, A_n$. It is easily seen that the two definitions are equivalent.

Definition A.3. Given a partition $\mathcal{A} = \{A_1, ..., A_n\}$ of Ω with *sigma*-field $\mathcal{F}^{\mathcal{A}}$ and a random variable X. The conditional expectation of X given $\mathcal{F}^{\mathcal{A}}$ is the random variable

$$E(X|\mathcal{F}^{\mathcal{A}})(\omega) = \sum_{i=1}^{n} 1_{A_i}(\omega) E(X|A_i) = \sum_{i=1}^{n} 1_{A_i}(\omega) \frac{E(X 1_{A_i})}{P(A_i)}.$$

Proposition A.4. *Given a partition $\mathcal{A}$ of Ω and a rv X. The conditional expectation $E(X|\mathcal{F}^{\mathcal{A}})$ has the following properties*

(i) $E(X|\mathcal{F}^{\mathcal{A}})$ is $\mathcal{F}^{\mathcal{A}}$-measurable.

(ii) For every random variable Y which is $\mathcal{F}^{\mathcal{A}}$-measurable (i.e. Y is constant on the sets A_i, $i = 1, ..., n$) we have $E(XY) = E(E(X|\mathcal{F}^{\mathcal{A}})Y)$.

Proof. The property (i) is clear, as $E(X|\mathcal{F}^{\mathcal{A}})$is constant on each A_j. As Y is $\mathcal{F}^{\mathcal{A}}$-measurable it is of the form $Y = \sum_{i=1}^{n} c_j 1_{A_j}$ for constants c_j. Hence

$$\begin{aligned} E(XY) &= \sum_{i=1}^{n} c_j E(X 1_{A_j}) = \sum_{i=1}^{n} c_j P(A_j) \frac{E(X 1_{A_j})}{P(A_j)} \\ &= E\left(\sum_{i=1}^{n} c_j \frac{E(X 1_{A_j})}{P(A_j)} 1_{A_j} \right) = E(Y E(X|\mathcal{F}^{\mathcal{A}})) . \quad \square \end{aligned}$$

□

A.1.2 Conditional Expectation - General Case

The explicit definition of the conditional expectation works only if $P(A_i) \geq 0$ for all sets in our partition. However, in continuous models this is usually not the case. We therefore use the properties of the conditional expectation obtained in Proposition A.4 to define the conditional expectation in more general situations.

Definition A.5. Given an integrable rv X on $(\Omega, \mathcal{F}, P)$ and a sigma-field $\mathcal{G} \subset \mathcal{F}$. A random variable Z is called conditional expectation of X given $\mathcal{G}$, $Z = E(X|\mathcal{G})$, if

(i) Z is $\mathcal{G}$-measurable.

(ii) $E(YX) = E(YZ)$ for all rvs Y which are $\mathcal{G}$-measurable.

Theorem A.6. *There is exactly one random variable Z which satisfies (i), (ii).*

The proof can be found in any standard textbook on probability theory.

Examples:

(1) If $\mathcal{G} = \{\emptyset, \Omega\}$ we have $E(X|\mathcal{G}) = E(X)$.

(2) If X is $\mathcal{G}$-measurable we have $E(X|\mathcal{G}) = X$

(3) If X_1, X_2 are independent and $\mathcal{G} := \sigma(X_2)$ we get for any bounded measurable function f that $E(f(X_1)|\mathcal{G}) = E(f(X_1))$.

Proposition A.7. *The conditional expectation has the following properties:*

(1) Linearity: $E(c_1X_1 + c_2X_2|\mathcal{G}) = c_1E(X_1|\mathcal{G}) + c_2E(X_2|\mathcal{G})$.

(2) If Y *is* $\mathcal{G}$*-measurable we have* $E(YX|\mathcal{G}) = YE(X|\mathcal{G})$.

(3) Projectivity of the conditional expectation: Consider sigma-fields $\mathcal{G}_0 \subseteq \mathcal{G}_1 \subseteq \mathcal{F}$. *Then we have* $E(X|\mathcal{G}_0) = E(E(X|\mathcal{G}_1)|\mathcal{G}_0)$, *in particular* $E(X) = E(X|\mathcal{G})$ *for every sub-*σ*-field* $\mathcal{G}$.

Property (3) is often referred to as law of iterated expectations.

Proof. ad 2) We have to check the property (ii) of the definition of the conditional expectation. Let Z be $\mathcal{G}$-measurable. Then

$$E(YXZ) = E((YZ)X) = E(YZE(X|\mathcal{G})) = E(Y(ZE(X|\mathcal{G}))\,,$$

as the product (YZ) is $\mathcal{G}$-measurable.

ad 3) $E(X|\mathcal{G}_0)$ is obviously $\mathcal{G}_0$-measurable. Consider a $\mathcal{G}_0$-measurable random variable Y. As Y is also $\mathcal{G}_1$-measurable, we have $E(Y(E(X|\mathcal{G}_1)) = E(XY)$. On the other hand we get from the definition of $E(X|\mathcal{G}_0)$ that $E(XY) = E(YE(X|\mathcal{G}_0))$. This shows that $E((X|\mathcal{G}_1)Y) = E(YE(X|\mathcal{G}_0))$ so that Definition (A.5)(ii) holds. □

Bibliography

Amin, K. I. & Jarrow, R. A. (1991), 'Pricing foreign currency options under stochastic interest rates', *Journal of International Money and Finance* **10**, 310–330.

Bachelier, L. (1900), Théorie de la spéculation, *in* P. H. Cootner, ed., 'The Random Character of Stock Market Prices (1964)', MIT Press, Cambridge, MA, pp. 17–78.

Bingham, N. H. & Kiesel, R. (1998), *Risk-Neutral Valuation*, Springer, Berlin.

Björk, T. (2004), *Arbitrage Theory in Continuous Time*, second edn, Oxford University Press, Oxford.

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**(3), 637–654.

Canter, M. S., Cole, J. B. & Sandor, R. S. (1996), 'Insurance derivatives: A new asset class for the capital markets and a new hedging tool for the insurance industry', *Journal of Derivatives* **Winter 1996**, 89–104.

Cont, R. (2006), 'Model uncertainty and its impact on the pricing of derivative instruments', *Mathematical Finance* **16**(3), 519–542.

Cont, R. & Tankov, P. (2004), *Financial Modelling with Jump Processes*, Chapman & Hall.

Cox, J. C., Ross, S. A. & Rubinstein, M. (1979), 'Option pricing: A simplified approach', *Journal of Financial Economics* **7**, 229–263.

Cox, J. C. & Rubinstein, M. (1985), *Options Markets*, Prentice Hall, Englewood Cliffs, NJ.

Delbaen, F. & Schachermayer, W. (1994), 'A general version of the fundamental theorem of asset pricing', *Mathematische Annalen* **300**, 463–520.

El Karoui, N., Jeanblanc-Picqué, M. & Shreve, S. E. (1998), 'Robustness of the Black and Scholes formula', *Mathematical Finance* **8**, 93–126.

El Karoui, N., Myneni, R. & Viswanathan, R. (1992), 'Arbitrage pricing and hedging of interest rate claims with state variables: 2, applications', Preprint, University of Paris VI.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

Filipović, D. (2009), *Term-Structure Models A Graduate Course*, Springer, Berlin Heidelberg.

Fleming, W. H. & Soner, H. M. (2006), *Controlled Markov Processes and Viscosity Solutions*, 2nd edn, Springer, New York.

Föllmer, H. (1981), Calcul d'Itô sans probabilités., *in* 'Séminaire Probabilités XV, Lecture Notes in Mathematics Vol. 850', Springer, Berlin, pp. 144–150.

Föllmer, H. & Schweizer, M. (1993), 'A microeconomic approach to diffusion models for stock prices', *Mathematical Finance* **3**(1), 1–23.

Frey, R. (1997), 'Derivative asset analysis in models with level-dependent and stochastic volatility', *CWI Quarterly* **10**, 1–34.

Frey, R. & Sommer, D. (1996), 'A systematic approach to pricing and hedging of international derivatives with interest rate risk', *Applied Mathematical Finance* **3**, 295–317.

Gibson, R., ed. (2000), *Model Risk: Concepts, Calibration and Pricing*, Risk Books, London.

Harrison, J. M. & Pliska, S. R. (1981), 'Martingales and stochastic integrals in the theory of continuous trading', *Stochastic Processes and their Applications* **11**, 215–260.

He, H. & Leland, H. E. (1993), 'On equilibrium asset price processes', *Review of Financial Studies* **6**(3), 593–617.

Heath, D., Jarrow, R. A. & Morton, A. (1992), 'Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation', *Econometrica. Journal of the Econometric Society* **60**(1), 77–105.

Hull, J. C. (1997), *Options, Futures and Other Derivatives*, 3rd edn, Prentice Hall, Englewood Cliffs, NJ.

Jacod, J. & Protter, P. (2004), *Probability Essentials*, Springer.

Jarrow, R. A. & Turnbull, S. M. (1996), *Derivative Securities*, South Western Publishing, Cincinnati.

Karatzas, I. & Shreve, S. E. (1988), *Brownian Motion and Stochastic Calculus*, Springer, Berlin.

Karatzas, I. & Shreve, S. E. (1998), *Methods of Mathematical Finance*, Springer, Berlin.

Lamberton, D. & Lapeyre, B. (1996), *Introduction to Stochastic Calculus Applied to Finance*, Chapman & Hall, London.

Liu, J. & Longstaff, F. (2000), Losing money on arbitrages. Preprint, UCLA.

McNeil, A. J., Frey, R. & Embrechts, P. (2005), *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press, Princeton.

Musiela, M. & Rutkowski, M. (1997), *Martingale Methods in Financial Modelling*, Springer, Berlin.

Oksendal, B. (1998), *Stochastic Differential Equations*, Universitext, 5th edn, Springer, Berlin.

Pham, H. (2009), *Continuous-Time Stochastic Control and Optimization with Financial Applications*, number 61 *in* 'Stochastic Modelling and Applied Probability', Springer, Berlin.

Protter, P. (2005), *Stochastic Integration and Differential Equations: A New Approach*, second edn, Springer, Berlin.

Revuz, D. & Yor, M. (1994), *Continuous Martingales and Brownian Motion*, Springer, Berlin.

Samuelson, P. A. (1965), 'Rational theory of warrant pricing', *Industrial Management Review* **6**, 13–31.

Sandmann, K. (1999), *Einführung in die Stochastik der Finanzmärkte*, Springer, Berlin.

Shreve, S. E. (2004), *Stochastic Calculus for Finance II*, Springer, New York.

Williams, D. (1991), *Probability with Martingales*, Cambridge University Press, Cambridge.

Wilmott, P., Dewynne, J. & Howison, S. (1993), *Option Pricing, Mathematical Models and Computation*, Oxford Financial Press, Oxford.