Lecture Notes Continuous-Time Finance

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

Version from May 23, 2020, 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 **??** 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 **??**. The text closes with a discussion of interest rate models and gives applications to interest-rate and and currency derivatives (in Chapter **??**). 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 (2003) 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 $\phi = (\phi_n)_{n=1,...,N}$ is called admissible if ϕ_n^0 and ϕ_n^1 are $\mathcal{F}_{n-1}$-measurable for $n = 1, \ldots N$, 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 $n = 1, \ldots, N$

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

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 $n = 1, \ldots, N$. 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 $T \in \{1, \ldots, N\}$ 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*

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

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 $(\tilde{V}_n(\phi))_{n=0,\ldots,T}$ 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.

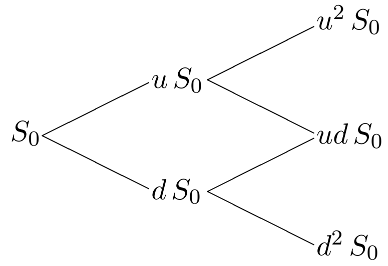
Theorem 1.14. *An arbitrage-free market $\mathcal{M}$ is complete if and only if there exists a* unique *equivalent martingale measure Q.*

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 $\omega_i \in \{u, d\}$, $i = 1, \dots, N$. 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 $\tau_0 < \tau_1 < \ldots$. 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

- $P(N_t = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$, $k = 0, 1, \ldots$, $t \geq 0$.
- $N_{t+u} - N_t$ 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 $\{\tau < t\} \in \mathcal{F}_t$, $\forall t \geq 0$.*

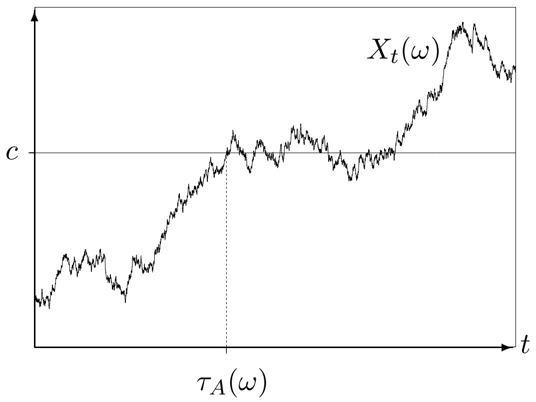
Proof. It holds $\{\tau \leq t\} = \bigcap_{\varepsilon>0}\{\tau < t+\varepsilon\}$. Let $\{\tau < t\} \in \mathcal{F}_t$, $\forall t \geq 0$. Hence $\{\tau < t+\varepsilon\} \in \mathcal{F}_{t+\varepsilon}$ and $\{\tau \leq t\} \in \bigcap\limits_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} = F_t$. For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

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

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 $t_0 \geq 0$ we have

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

Let S and T be stopping times, where $S \leq T$. 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 $S \leq T$. Then $\mathcal{F}_S \subseteq \mathcal{F}_T$.*

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

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

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 $\{S \wedge T \leq t\} = \{S \leq t\} \cup \{T \leq t\}$, $\{S \vee T \leq t\} = \{S \leq t\} \cap \{T \leq t\}$. As S and T are stopping times, that means $\{S \leq t\} \in \mathcal{F}_t$ and $\{T \leq t\} \in \mathcal{F}_t$, 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 $A \cap \{S \leq t\} \in \mathcal{F}_t$ and $A \cap \{T \leq t\} \in \mathcal{F}_t$. It holds

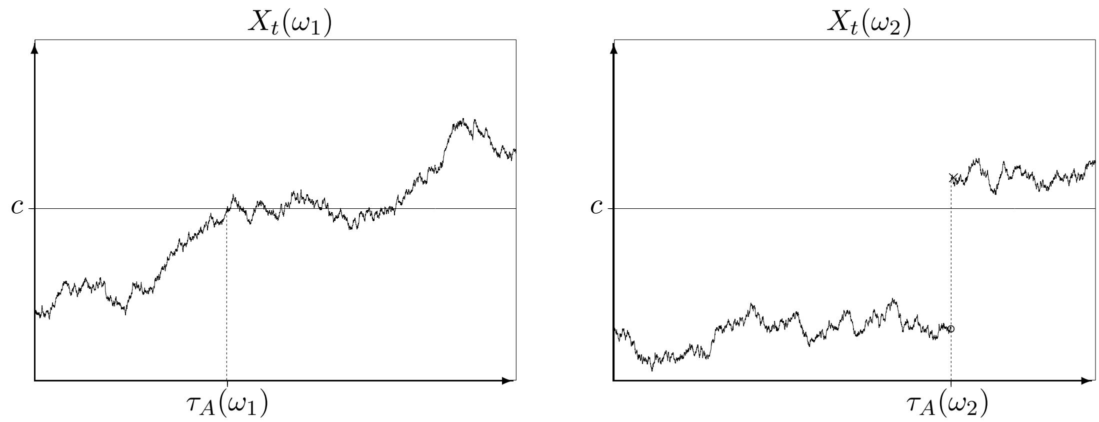
$$(A \cap \{S \leq t\}) \cup (A \cap \{T \leq t\}) = A \cap (\{S \leq t\} \cup \{T \leq t\}) = A \cap \{S \wedge T \leq t\} \in \mathcal{F}_t,$$

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 $X_T(\omega_2) > c$.

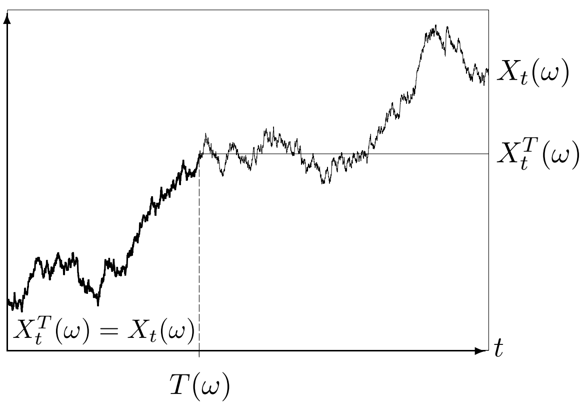


Lemma 2.14. *Let $(X_t)_{t\geq 0}$ be an $\{\mathcal{F}_t\}$-adapted and right-continuous stochastic process and let T be a stopping time. Then the rv X_T is $\mathcal{F}_T$-measurable.*

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 $E(X_T \,|\, \mathcal{F}_S) = X_S$.

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 $\mathcal{F}_t := \sigma(W_s, s \leq t)$. Then a) $(W_t)_{t\geq 0}$ b) $(W_t^2 - t)_{t\geq 0}$ 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 $|\tau| := \sup_{1\leq i\leq n} |t_i - t_{i-1}|$.

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

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

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

Remarks on notation: 1) Following standard conventions we denote by $\operatorname{Var}(f)$ the first variation of a function f, whereas $\operatorname{var}(Y)$ 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

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

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

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

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 $t_i \le t$ and denote by $(\Delta X)_{i,n}$ the increment $X_{t_i} - X_{t_{i-1}}$. We get from a Taylor-expansion of F

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

where $\tilde{t}$ is some point in the interval (t_{i-1}, t_i), and where $R_{i,n} := \frac{1}{2}(F''(X_{\tilde{t}}) - F''(X_{t_{i-1}}))(\Delta X)_{i,n}^2$. 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,

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

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

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

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

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

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

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

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

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

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

$$0 \le E\left((M_t - M_0)^2\right) = E\left(M_t^2 - 2M_tM_0 + M_0^2\right) = M_0^2 - 2M_0^2 + M_0^2 = 0\,,$$

which shows that $E\,(M_t - M_0)^2 = 0$ 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

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

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

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

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 (2003) 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 $V_t^n = \phi_t^n S_t^1 + \eta_t^n S_t^0$. 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 $t_i \in \tau_n$

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

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 $(\tau(t), z(t,S))$)

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

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.

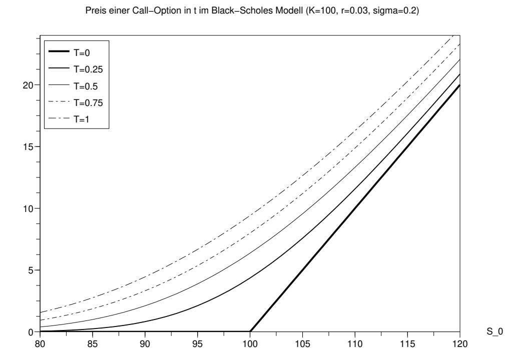
A probabilistic derivation of the Black-Scholes formula is given in Section **??** below.

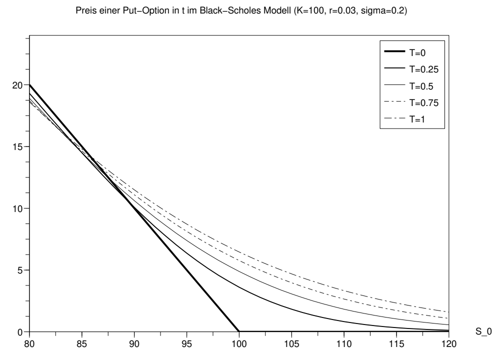
4.3.2 Properties of option prices and the Greeks

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 $(C_{BS}(t,S) - N(d_1)S)/e^{rt} = -e^{-rT}KN(d_2)$ 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

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

  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.

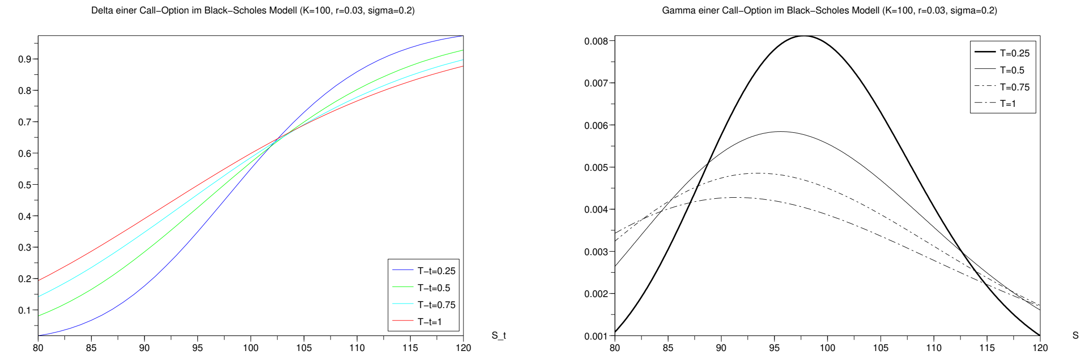


Figure 4.1: Delta (left) and Gamma (right) for a Call as a function of current price S for sevaral values of $T-t$

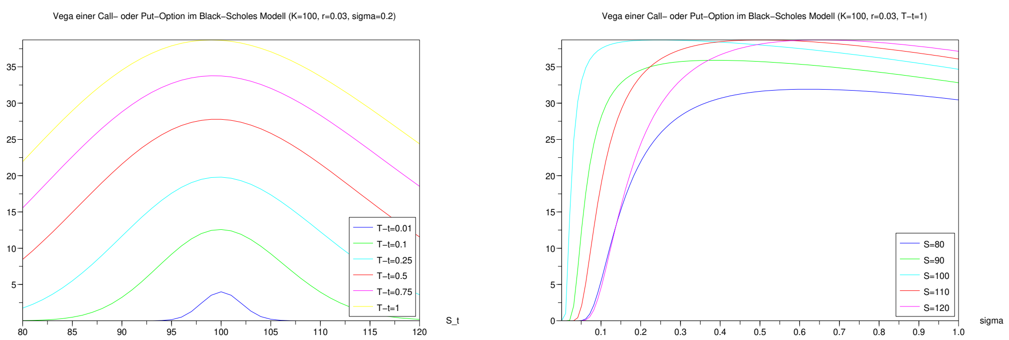


Figure 4.2: Vega of a call/put as a function of S for different $T-t$ (left) and as a function of σ for different S (right)

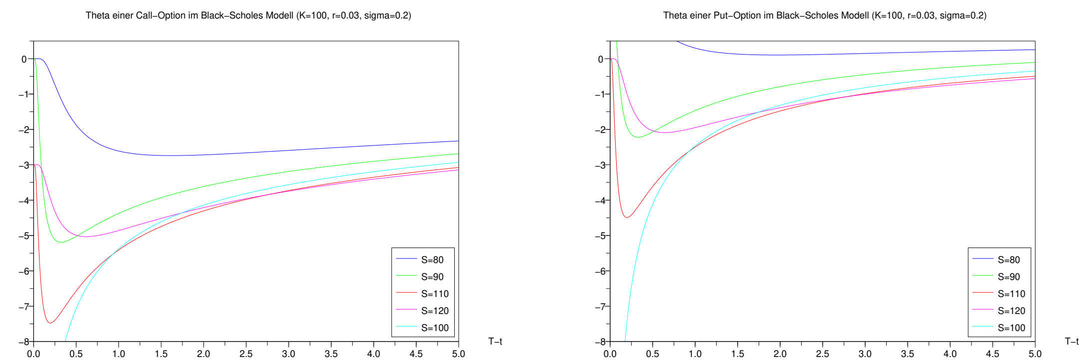


Figure 4.3: Theta of a call (left) and of a put (right). Note that the Theta of a put can become positive for small S.

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

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

□

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 $\mathcal{A} = \{A_1, \dots, A_n\}$ 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 $\mathcal{A} = \{A_1, \dots, A_n\}$ 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 $A_1, \dots, A_n$. 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. $\square$

Bibliography

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. (2003), *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.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

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.

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.

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.

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.