A short course in Mathematical Finance

Josef Teichmann

Institute of mathematical Methods in Economics, Research Group of Financial and Actuarial Mathematics, Vienna University of Technology, Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria

E-mail address: josef.teichmann@fam.tuwien.ac.at

URL: www.fam.tuwien.ac.at/~jteichma

ABSTRACT. An introduction to no arbitrage theory and optimization of expected utility in discrete time models with a view towards geometric Brownian motion and the celebrated Black-Scholes formula.

I thank many of my students for the encouraging and careful feedback on my lectures. In particular I would like to thank Christian Bayer, Barbara Forster, Maria Siopacha and Richard Warnung for their eagle-eyed observations on several stages of preparation of these lecture notes.

February 28, 2007 (this draft).

Contents

CHAPTER 1

Discrete (time) Models

1. Topic of financial mathematics

Financial Mathematics is dealing with the analysis of price structures in financial markets. On the one hand one tries to understand the mechanism of making prices, on the other hand one tries to understand the stochastic behavior of price evolutions and the relations between different traded objects.

At the beginning we shall mainly focus on the second issue, hence we shall formulate models for price evolutions and conditions on them to be reasonable models from the point of view of economics: main focus lies on the introduction and meaning of trading portfolios within stochastic models for prices.

Even though we concentrate on discrete models first, the theory is challenging from the point of view of mathematics and also from the point of view of modeling. Discrete Models are stochastic processes modeled on discrete probability spaces. This seems to be a rough approximation, but all the main structural questions already appear in this setting and one can learn a lot about financial mathematics in this setting.

We shall also present the theoretical basis for utility optimization, which can be understood as a theory of making prices.

2. No Arbitrage Theory for discrete models

This is the main section for no arbitrage theory, all the mathematical notions can be found in chapter 3.

A *discrete model for a financial market* is an adapted $(d+1)$ -dimensional stochastic process S with $S_n := (S_n^0, \dots, S_n^d)$ for $n = 0, \dots, N$ on a finite probability space $(\Omega, \mathcal{F}, P)$ with filtration $\mathcal{F}_0 \subset \cdots \subset \mathcal{F}_N$ with $\mathcal{F}_N = \mathcal{F}$. We shall always assume that all σ-algebras contain all P-nullsets. The price process $(S_n^0)_{n=0,\dots,N}$ is assumed to be strictly positive and called the *risk-less asset* (even if it is stochastic) and we define $S_0^0 := 1$ (most of the formulas are valid even if $S_0^0 \neq 1$). We think of a bank account, where one can freely move money. The coefficients $\beta_n := \frac{1}{S_n^0}$ for $n = 0, \dots, N$ are called *discount factors*. The assets $S^1, \dots S^d$ are called *risky assets*.

A *trading strategy* is a predictable stochastic process ϕ with $\phi_n = (\phi_n^0, \dots, \phi_n^d)$ for $n = 0, \dots, N$. We think of a portfolio formed by an amount of ϕ_n^0 in the bank account and ϕ_n^i units of risky assets, at time n. The value or wealth at time n of such a portfolio is

$$V_n(\phi) = \phi_n S_n := \sum_{i=0}^{d} \phi_n^i S_n^i$$

for $n = 0, \dots, N$. We shall always assume that ϕ_0 is constant.

The *discounted value process* is given through

$$\widetilde{V}_n(\phi) = \beta_n(\phi_n S_n) = \phi_n \widetilde{S}_n$$

for $n = 0, \ldots, N$, where $\widetilde{S}_n = \beta_n S_n$ denotes the *discounted price process.*

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

$$\phi_n S_n = \phi_{n+1} S_n$$

for $n = 0, \ldots, N-1$. We interpret this condition that the readjustment at time n to new prices S_n is done without bringing in or consuming any wealth.

This condition is obviously equivalent to

$$\phi_{n+1}(S_{n+1} - S_n) = \phi_{n+1} S_{n+1} - \phi_n S_n,$$

for $n = 0, \ldots, N-1$, and therefore

$$V_{n+1}(\phi) - V_n(\phi) = \phi_{n+1}(S_{n+1} - S_n)$$

for $n = 0, \ldots, N-1$, which means that the changes of the value process are due to changes in the stock prices.

2.1. Proposition. *Let $S = (S^0, \ldots, S^d)$ be a discrete model of a financial market and ϕ a trading strategy, then the following assertions are equivalent:*

(1) *The strategy ϕ is self-financing.*
(2) *For $n = 0, \ldots, N$ we have*

$$V_n(\phi) = V_0(\phi) + (\phi \cdot S)_n.$$

(3) *For $n = 0, \ldots, N$ we have*

$$\widetilde{V}_n(\phi) = V_0(\phi) + (\phi \cdot \widetilde{S})_n.$$

Proof. The equivalence of 1. and 2. results from the previous remark. The equivalence of 1. and 3. results from the fact that ϕ is self-financing if and only if

$$\phi_n \widetilde{S}_n = \phi_{n+1} \widetilde{S}_n$$

for $n = 1, \ldots, N$, which leads to

$$\phi_{n+1}(\widetilde{S}_{n+1} - \widetilde{S}_n) = \phi_{n+1} \widetilde{S}_{n+1} - \phi_n \widetilde{S}_n$$

and therefore the result as in 2. Notice that therefore

$$\widetilde{V_n}(\phi) = V_0(\phi) + \sum_{j=1}^{n} \sum_{i=1}^{d} \phi_j^i (\widetilde{S}_{j+1}^i - \widetilde{S}_j^i).$$

The 0-th component does not enter in the calculation since $\widetilde{S}_{j+1}^0 = 1$ and therefore the increments vanishes. □

A self-financing trading strategy ϕ can also be given through the initial value $V_0(\phi)$ and $(\phi^1, \ldots, \phi^d)$, which is proved in the following proposition:

2.2. Proposition. *For any predictable process $(\phi^1, \ldots, \phi^d)$ and for any value V_0 there exists a unique predictable process ϕ^0 such that $(\phi^0, \ldots, \phi^d)$ is a self-financing trading strategy with $V_0(\phi) = V_0$ and $\widetilde{V}_n(\phi) = V_0 + (\phi \cdot \widetilde{S})_n$ for $n = 0, \ldots, N$.*

Proof. If we have a self-financing trading strategy the formula

$$\begin{aligned}\widetilde{V_n}(\phi) &= \phi_n^0 + \phi_n^1 \widetilde{S}_n^1 + \cdots + \phi_n^d \widetilde{S}_n^d \\ &= V_0 + (\phi \cdot \widetilde{S})_n\end{aligned}$$

holds, where-from we can calculate ϕ^0. The process ϕ^0 is predictable since

$$\phi_n^0 = V_0 + (\phi \cdot \widetilde{S})_{n-1} - \phi_n^1 \widetilde{S}_{n-1}^1 - \cdots - \phi_n^d \widetilde{S}_{n-1}^d.$$

□

2.3. Definition. *Let $S = (S^0, \ldots, S^d)$ be a discrete model for a financial market, then the model is called* arbitrage-free *if for any trading strategy ϕ the assertion*

$$V_0(\phi) = 0 \textit{ and } V_N(\phi) \geq 0, \textit{ then } V_N(\phi) = 0$$

holds true. We call a trading strategy ϕ an arbitrage opportunity (arbitrage strategy) *if $V_0(\phi) = 0$ and $V_N(\phi) \gneqq 0$.*

2.4. Definition. *A* contingent claim (derivative) *is an element of $L^2(\Omega, \mathcal{F}, P)$. We denote by $\widetilde{X}$ the discounted price at time N, i.e. $\widetilde{X} = \frac{1}{S_N^0} X$. We call the subspace of $\mathcal{K} \subset L^2(\Omega, \mathcal{F}, P)$*

$$\begin{aligned}\mathcal{K} &:= \{\widetilde{V_N}(\phi) | \ \phi \textit{ self-financing trading strategy, } \widetilde{V_0}(\phi) = 0\} \\ &= \{(\phi \cdot \widetilde{S})_N | \ \phi \textit{ predictable}\}\end{aligned}$$

the space of contingent claims attainable at price 0 *(see Proposition 2.2). We call the convex cone*

$$C = \{Y \in L^2(\Omega, \mathcal{F}, P) | \textit{ there is } X \in K \textit{ such that } X \geq Y\} = \mathcal{K} - L^2_{\geq 0}(\Omega, \mathcal{F}, P)$$

the cone of claims super-replicable at price 0 *or the outcomes with zero investment and consumption. A contingent claim X is called* replicable at price x *and at time N if there is a self-financing trading strategy ϕ such that*

$$\widetilde{X} = x + (\phi \cdot \widetilde{S})_N \in x + \mathcal{K}.$$

A contingent claim X is called super-replicable at price x *and at time N if there is a self-financing trading strategy ϕ such that*

$$\widetilde{X} \leq x + (\phi \cdot \widetilde{S})_N \in x + \mathcal{K},$$

in other words if $\widetilde{X} \in C$.

2.5. Remark. The set $\mathcal{K}$ is a subspace of $L^2(\Omega, \mathcal{F}, P)$ and the positive cone $L^2_{\geq 0}(\Omega, \mathcal{F}, P)$ is polyhedral, therefore by C is closed by Proposition 1.11.

We see immediately that a discrete model for a financial market is arbitrage-free if

$$\mathcal{K} \cap L^2_{\geq 0}(\Omega, \mathcal{F}, P) = \{0\},$$

which is equivalent to

$$C \cap L^2_{\geq 0}(\Omega, \mathcal{F}, P) = \{0\}.$$

Given a discrete model for a financial market, then we call a measure Q equivalent to P an *equivalent martingale measure* with respect to the numeraire S^0 if the discounted price process $\widetilde{S^i}$ are Q-martingales for $i = 0, \ldots, N$. We denote the set of equivalent martingale measures with respect to the numeraire S^0 by $\mathcal{M}^e(S, S^0)$. If the numeraire satisfies $S^0 = 1$ we shall write $\mathcal{M}^e(S)$. In particular $\mathcal{M}^e(S, S^0) =$

$\mathcal{M}^e(\widetilde{S})$. We denote the absolutely continuous martingale measures with respect to the numeraire S^0 by $\mathcal{M}^a(S, S^0)$. If the numeraire process satisfies $S^0_n = 1$ we shall write $\mathcal{M}^a(S)$. In particular $\mathcal{M}^a(S, S^0) = \mathcal{M}^a(\widetilde{S})$.

2.6. Theorem. *Let S be a discrete model for a financial market, then the following two assertions are equivalent:*

1. *The model is arbitrage-free.*
2. *The set of equivalent martingale measures is non-empty, $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$.*

PROOF. We shall do the proof in two steps. First we assume that there is an equivalent martingale measure $Q \sim P$ with respect to the numeraire S^0. We want to show that there is no arbitrage opportunity. Let ϕ be a self-financing trading strategy and assume that

$$V_0(\phi) = 0, \ V_N(\phi) \geq 0,$$

then the discounted value process of the portfolio

$$\widetilde{V}_n(\phi) = (\phi \cdot \widetilde{S})_n$$

is a martingale with respect to Q by Theorem 4.6 and therefore

$$E_Q(\widetilde{V}_N(\phi)) = 0.$$

Hence we obtain by equivalence $V_N(\phi) = 0$ since $V_N(\phi) \geq 0$ Q-almost surely, so there is no arbitrage opportunity.

Next we assume that the market is arbitrage-free. Then

$$\mathcal{K} \cap L^2_{\geq 0}(\Omega, \mathcal{F}, P) = \{0\}$$

and therefore we find a linear functional l that separates $\mathcal{K}$ and the compact, convex set

$$\{Y \in L^2_{\geq 0}(\Omega, \mathcal{F}, P) | \ E_P(Y) = 1\},$$

i.e. $l(X) = 0$ for all $X \in \mathcal{K}$ and $l(Y) > 0$ for all $Y \in L^2_{\geq 0}(\Omega, \mathcal{F}, P)$ with $E_P(Y) = 1$. We define

$$Q(A) = \frac{l(1_A)}{l(1_\Omega)}$$

for measurable sets $A \in \mathcal{F}$ and obtain an equivalent probability measure $Q \sim P$, since $l(1_A) > 0$ for sets with $P(A) > 0$. We have in particular from separation

$$E_Q((\phi \cdot \widetilde{S})_N) = 0$$

for any predictable processes ϕ. Therefore $\widetilde{S}$ is a Q-martingale by Theorem 4.6. □

Now we can formulate a basic pricing theory for contingent claims.

2.7. Definition. *A* pricing rule *for contingent claims $X \in L^2(\Omega, \mathcal{F}, P)$ at time N is a map*

$$X \mapsto \pi(X)$$

where $\pi(X) = (\pi(X)_n)_{n=0,\ldots,N}$ is an adapted stochastic process, which determines the price of the claim at time N at time $n \leq N$. In particular $\pi(X)_N = X$ for any $X \in L^2(\Omega, \mathcal{F}, P)$. A pricing rule is arbitrage-free *if for any finite set of claims $X_1, \ldots, X_k$ the discrete time model of a financial market*

$$(S^0, S^1, \ldots, S^d, \pi(X_1), \ldots, \pi(X_k))$$

is arbitrage-free.

2.8. Lemma (arbitrage-free prices)**.** *Let π be an arbitrage-free pricing rule for a set of contingent claims $\mathfrak{X}$, then the discrete model $(S^0,\dots,S^d)$ is arbitrage-free and there is $Q \in \mathcal{M}^e(\widetilde{S})$ such that*

$$\pi(X)_n = E_Q(\frac{S_n^0}{S_N^0}X|\mathcal{F}_n),$$

for all $X \in \mathfrak{X}$. If the discrete time model S is arbitrage-free, then

$$\pi(X)_n = E_Q(\frac{S_n^0}{S_N^0}X|\mathcal{F}_n)$$

is an arbitrage-free pricing rule for all contingent claims $X \in L^2(\Omega,\mathcal{F},P)$. Hence the only arbitrage-free prices are conditional expectation of the discounted claims with respect to Q and pricing rules are always linear.

Proof. If the market $(S^0,S^1,\dots,S^d,\pi(X))$ is arbitrage-free, we know that there exists an equivalent martingale measure Q such that the discounted prices are Q-martingales. Hence in particular

$$\frac{\pi(X)_n}{S_n^0}$$

is a Q-martingale, so

$$E(\frac{\pi(X)_N}{S_N^0}|\mathcal{F}_n) = E(\frac{X}{S_N^0}|\mathcal{F}_n) = \frac{\pi(X)_n}{S_n^0}$$

which yields the desired relation.

Given an arbitrage-free discrete model S and define the pricing rules by the above relation for one equivalent martingale measure $Q \in \mathcal{M}^e(\widetilde{S})$, then the whole market is arbitrage-free by the existence of at least one equivalent martingale measure, namely Q. □

2.9. Remark. Taking not an equivalent but an absolutely continuous martingale measure $Q \in \mathcal{M}^a(\widetilde{S})$ means that there is at least one measurable set A such that $Q(A) = 0$ and $P(A) > 0$. Hence the claim 1_A with $P(A) > 0$ would have price 0, which immediately leads to arbitrage by entering this contract $X = 1_A$. Therefore only equivalent martingale measures are possible for pricing.

The set of equivalent martingale measures $\mathcal{M}^e(\widetilde{S})$ is convex and the set $\mathcal{M}^a(\widetilde{S})$ is compact and convex. Therefore the analysis of the extreme points of $\mathcal{M}^a(\widetilde{S})$ is of particular importance.

2.10. Remark. Given an arbitrage-free financial market such that $\mathcal{M}^e(\widetilde{S})$ contains more than one measure. Then an equivalent martingale measure $Q \in \mathcal{M}^e(\widetilde{S})$ can never be an extreme point of $\mathcal{M}^a(\widetilde{S})$. Assume that there were an extreme point $Q \in \mathcal{M}^e(\widetilde{S})$ of $\mathcal{M}^a(\widetilde{S})$ and take $Q_0 \neq Q$ with $Q_0 \in \mathcal{M}^e(\widetilde{S})$. Then we know that the segment $tQ + (1-t)Q_0 \in \mathcal{M}^e(\widetilde{S})$. For t near 1 we also have equivalent martingale measures. We also know that C is finitely generated by $\langle h_1,\dots,h_M,-h_1,\dots,-h_M,-e_1,\dots,-e_k\rangle_{con}$, where $h_1,\dots,h_M$ is a basis of $\mathcal{K}$ and $e_1,\dots,e_k$ generates the non-negative cone $L^2_{\geq 0}(\Omega,\mathcal{F},P)$. An equivalent martingale measure Q_t is defined via the relation $E_{Q_t}(C) \leq 0$. The Q_t are equivalent martingale measures since $E_{Q_t}(h_i) = 0$ and $E_{Q_t}(e_i) < 0$. So we can continue a little bit in t-direction beyond 1 and obtain again equivalent martingale measures, too. Hence

Q cannot be an extreme point, since it is a middle point of two equivalent martingale measures. Therefore an extreme point is necessarily absolutely continuous and not equivalent to P.

2.11. Theorem. *Let S be a discrete model for a financial market and assume $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$ and $\mathcal{M}^a(\widetilde{S}) = \langle Q_1, \dots, Q_m \rangle$ Then for all $X \in L^2(\Omega, \mathcal{F}, P)$ the following assertions are equivalent:*

1. *$X \in \mathcal{K}$ ($X \in C$).*
2. *For all $Q \in \mathcal{M}^e(\widetilde{S})$ we have $E_Q(X) = 0$ (for all $Q \in \mathcal{M}^e(\widetilde{S})$ we have $E_Q(X) \leq 0$).*
3. *For all $Q \in \mathcal{M}^a(\widetilde{S})$ we have $E_Q(X) = 0$ (for all $Q \in \mathcal{M}^a(\widetilde{S})$ we have $E_Q(X) \leq 0$).*
4. *For all $i = 1, \dots, m$ we have $E_{Q_i}(X) = 0$ (for all $i = 1, \dots, m$ we have $E_{Q_i}(X) \leq 0$).*

Proof. We shall calculate the polar cone of the cone C,

$$C^0 = \{Z \in L^2(\Omega, \mathcal{F}, P) \text{ such that } E_P(ZX) \leq 0\}$$

by definition. For $Q \in \mathcal{M}^a(\widetilde{S})$ we calculate the Radon-Nikodym-derivative $\frac{dQ}{dP}$ and see that

$$E_P(\frac{dQ}{dP}X) = E_Q(X) = E_Q((\phi \cdot \widetilde{S})_N + Y)$$

for $Y \leq 0$, hence – due to the fact that Q is a martingale measure (so the expectation of the stochastic integral vanishes) – we obtain

$$E_P(\frac{dQ}{dP}X) = E_Q(Y) \leq 0.$$

Consequently $\frac{dQ}{dP} \in C^0$. Given now $Z \in C^0$, then by the same reasoning we obtain

$$E_P(ZX) \leq 0$$

for all $X \in C$. Since the model is arbitrage-free, $Z \geq 0$, assume $Z \neq 0$, so

$$E_P(\frac{Z}{E_P(Z)}(\phi \cdot \widetilde{S})_N) \leq 0$$

for all self-financing trading strategies ϕ. Replacing ϕ by $-\phi$ we arrive at

$$E_P(\frac{Z}{E_P(Z)}(\phi \cdot \widetilde{S})_N) = 0,$$

which means that $\frac{Z}{E_P(Z)} \in \mathcal{M}^a(\widetilde{S})$. Since the polar cone of C is exactly given by the cone generated by $\frac{dQ}{dP}$ for $Q \in \mathcal{M}^a(\widetilde{S})$, all the assertion hold by the bipolar theorem.

$$C^0 = \left\langle \frac{dQ_1}{dP}, \dots, \frac{dQ_m}{dP} \right\rangle_{cone},$$
$$C^{00} = C = \{X \in L^2(\Omega, \mathcal{F}, P) \text{ such that } E_{Q_i}(X) \leq 0 \text{ for } i = 1, \dots, m\}.$$
$$\mathcal{K}^0 = \left\langle \frac{dQ_1}{dP}, \dots, \frac{dQ_m}{dP} \right\rangle_{vector},$$
$$\mathcal{K}^{00} = \mathcal{K} = \{X \in L^2(\Omega, \mathcal{F}, P) \text{ such that } E_{Q_i}(X) = 0 \text{ for } i = 1, \dots, m\}.$$

□

The last step of the general theory is the distinction between complete and incomplete markets and a renewed description of pricing procedures.

2.12. Definition. *Let S be a discrete model for a financial market and assume $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$. The financial market is called* complete *if $\mathcal{M}^e(\widetilde{S}) = \{Q\}$, i.e. the equivalent martingale measure is unique. The financial market is called* incomplete *if $\mathcal{M}^e(\widetilde{S})$ contains more than one element. In this case $\mathcal{M}^a(\widetilde{S}) = \langle Q_1, \dots, Q_m \rangle_{convex}$ for linearly independent measures Q_i, $i = 1, \dots, m$ and $m \geq 2$.*

2.13. Theorem (complete markets)**.** *Let S be discrete model of a financial market with $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$. Then the following assertions are equivalent:*

1. *S is complete financial market.*
2. *For every claim X there is a self-financing trading strategy ϕ such that the claim can be replicated, i.e.*
$$V_N(\phi) = X.$$
3. *For every claim X there is a predictable process ϕ and a unique number x such that the discounted claim can be replicated, i.e.*
$$\widetilde{X} = \frac{1}{S_N^0} X = x + (\phi \cdot \widetilde{S}).$$
4. *There is a unique pricing rule for every claim X.*

Proof. We can collect all conclusions from the previous results. 2. and 3. are clearly the same by discounting.

1.⇒2.,3.: If S is complete, then there is a unique equivalent martingale measure Q such that the discounted stock prices are Q-martingales. Take a claim X, then we know by Lemma 2.8 that
$$\pi(X)_n = \frac{S_n^0}{S_N^0} E_Q(X|\mathcal{F}_n)$$
is the only arbitrage-free price for X at time n, since there is only one martingale measure Q. The final value of the martingale $(\frac{\pi(X)_n}{S_n^0})_{n=0,\dots,N}$ can be decomposed into
$$\frac{\pi(X)_N}{S_N^0} = x + (\phi \cdot \widetilde{S})_N$$
Since $E_Q(\frac{\pi(X)_N}{S_N^0} - x) = 0$ means $\frac{\pi(X)_N}{S_N^0} - x \in \mathcal{K}$ by Theorem 4. So we have proved 3. and therefore also 2..

2.⇒4.: Given a claim X. If we are given a portfolio ϕ, which replicates the claim X, then we know that
$$\pi(X)_n = V_n(\phi)$$
for $n = 0, \dots, N$ defines a pricing rule. Therefore the pricing rule is uniquely given by the values of the portfolio, since the values of the portfolio are unique due to No-Arbitrage.

4.⇒1.: If we have a unique pricing rule $\pi(X)$ for any claim X, then we know by Lemma 2.8 that we have an equivalent martingale measure. □

2.14. Example. The Cox-Ross-Rubinstein model is a complete financial market model: The CRR-model is defined by the following relations
$$S_n^0 = (1+r)^n$$

for $n = 0, \dots, N$ and $r \geq 0$ is the bond-process.

$$S_{n+1} := \begin{cases} S_n(1+a) \\ S_n(1+b) \end{cases}$$

for $-1 < a < b$ and $n = 0, \dots, N$. We can write the probability space as $\{1+a, 1+b\}^N$ and think of $1+a$ as "down movement" and $1+b$ as up-movement. Every path is then a sequence of ups and downs. The σ-algebras $\mathcal{F}_n$ are given by $\sigma(S_0, \dots, S_n)$, which means that atoms of $\mathcal{F}_n$ are of the type

$$\{(x_1, \dots, x_n, y_{n+1}, \dots, y_N) \text{ for all } y_{n+1}, \dots, y_N \in \{1+a, 1+b\}\}$$

with $x_1, \dots, x_n \in \{1+a, 1+b\}$ fixed. Hence the atoms form a subtree, which starts after the moves $x_1, \dots, x_n$.

The returns $(T_i)_{i=1,\dots,N}$ are well-defined by

$$T_i := \frac{S_i}{S_{i-1}}$$

for $i = 1, \dots, N$. This process is adapted and each T_i can take two values

$$T_i = \begin{cases} 1+a \\ 1+b \end{cases}$$

with some specified probabilities depending on $i = 1, \dots, N$. We also note the following formula

$$S_n \prod_{i=n+1}^{m} T_i = S_m$$

for $m \geq n$. Hence it is sufficient for the definition of the probability on $(\Omega, \mathcal{F}, P)$ to know the distribution of $(T_1, \dots, T_N)$, i.e.

$$P(T_1 = x_1, \dots, T_N = x_N)$$

has to be known for each $x_i \in \{1+a, 1+b\}$.

2.15. Proposition. *Let $-1 < a < b$ and $r \geq 0$, then the CRR-model is arbitrage-free if and only if $r \in]a, b[$. If this condition is satisfied, then martingale measure Q for the discounted price process $(\frac{S_n}{(1+r)^n})_{n=0,\dots,N}$ is unique and characterized by the fact that $(T_i)_{i=1,\dots,N}$ are independent and identically distributed and*

$$T_i = \begin{cases} 1+a \text{ with probability } 1-q \\ 1+b \text{ with probability } q \end{cases}$$

for $q = \frac{r-a}{b-a}$.

Proof. The proof is done in several steps: First we assume that there is an equivalent martingale measure Q for the discounted price process $(\frac{S_n}{(1+r)^n})_{n=0,\dots,N}$. Then we can prove immediately that for $i = 0, \dots, N-1$

$$E_Q(T_{i+1} | \mathcal{F}_i) = 1 + r$$

simply by

$$E_Q(\frac{S_{i+1}}{(1+r)^{i+1}} | \mathcal{F}_i) = \frac{S_i}{(1+r)^i}$$

$$E_Q(\frac{S_{i+1}}{S_i} | \mathcal{F}_i) = 1 + r.$$

Taking this property we see by evaluation at $i = 0$ that

$$\begin{aligned} E_Q(T_1) &= 1 + r \\ &= Q(T_1 = 1 + a)(1 + a) + Q(T_1 = 1 + b)(1 + b), \\ r &= Q(T_1 = 1 + a)a + Q(T_1 = 1 + b)b, \end{aligned}$$

since $Q(T_1 = 1 + a) + Q(T_1 = 1 + b) = 1$ and both are positive quantities. Hence $r \in]a, b[$.

On the other hand the only solution of

$$(1 - q)(1 + a) + q(1 + b) = 1 + r$$

is given through $q = \frac{r-a}{b-a}$. Therefore under the martingale measure Q the condition on conditional expectations of the returns T_i reads as

$$\begin{aligned} E_Q(1_{\{T_{i+1}=1+a\}} | \mathcal{F}_i) &= 1 - q, \\ E_Q(1_{\{T_{i+1}=1+b\}} | \mathcal{F}_i) &= q \end{aligned}$$

and consequently the random variables are independent and identically distributed as described above under Q. Take the case of two returns, then

$$\begin{aligned} Q(T_1 = 1 + a, T_2 = 1 + a) &= E_Q(1_{\{T_2=1+a\}} 1_{\{T_1=1+a\}}) \\ &= E_Q(E_Q(1_{\{T_2=1+a\}} | \mathcal{F}_1) 1_{\{T_1=1+a\}}) = (1 - q) E_Q(1_{\{T_1=1+a\}}) \\ &= (1 - q)^2. \end{aligned}$$

Therefore the equivalent martingale measure is unique and given as above.

To prove existence of Q we show that the returns satisfy

$$E_Q(T_{i+1} | \mathcal{F}_i) = 1 + r$$

for $i = 0, \dots, N - 1$ if we choose Q as above. If the returns are independent, then

$$E_Q(T_{i+1} | \mathcal{F}_i) = E_Q(T_i)$$

which equals $1 + r$ in the described choice of the measure. □

2.16. Example. We can calculate the limit of a CRR-model. Fix $\sigma > 0$ the time-normalized volatility, i.e. the standard deviation of the return of the stock. Therefore we assume

$$\begin{aligned} \ln(1 + a) &= -\frac{\sigma}{\sqrt{N}} \\ \ln(1 + b) &= \frac{\sigma}{\sqrt{N}}, \end{aligned}$$

which yields i.i.d random variables

$$T_i = \begin{cases} 1 + a \text{ with probability } 1 - q \\ 1 + b \text{ with probability } q \end{cases}$$

with $q = \frac{b}{b-a} = \frac{\exp(\frac{\sigma}{\sqrt{N}})-1}{\exp(\frac{\sigma}{\sqrt{N}})-\exp(-\frac{\sigma}{\sqrt{N}})}$ denotes the building factor of the martingale measure. The stock price in the martingale measure is given by

$$\begin{aligned} S_n &= S_0 \prod_{i=1}^{n} T_i \\ &= S_0 \exp(\sum_{i=1}^{n} \ln T_i). \end{aligned}$$

The random variables $\ln T_i$ take values $-\frac{\sigma}{\sqrt{N}}, \frac{\sigma}{\sqrt{N}}$ with probabilities q and $1-q$, so

$$\begin{aligned} E_Q(\ln T_i) &= \frac{\sigma}{\sqrt{N}} - \frac{\sigma}{\sqrt{N}} \frac{2\exp(\frac{\sigma}{\sqrt{N}})-2}{\exp(\frac{\sigma}{\sqrt{N}})-\exp(-\frac{\sigma}{\sqrt{N}})} \\ &= \frac{\sigma}{\sqrt{N}} \frac{2-\exp(\frac{\sigma}{\sqrt{N}})-\exp(-\frac{\sigma}{\sqrt{N}})}{\exp(\frac{\sigma}{\sqrt{N}})-\exp(-\frac{\sigma}{\sqrt{N}})} \\ E_Q(\ln(T_i)^2) &= \frac{\sigma^2}{N}. \end{aligned}$$

Therefore the sums $\sum_{i=1}^n \ln T_i$ satisfy the requirements of the central limit theorem, namely

$$\sum_{i=1}^N \ln T_i = \frac{1}{\sqrt{N}} \sum_{i=1}^N \sqrt{N} \ln T_i \to N(-\frac{\sigma^2}{2}, \sigma^2)$$

in law for $N \to \infty$, since $E_Q(N \ln T_i) \to -\frac{\sigma^2}{2}$ as $N \to \infty$ and $\sqrt{N}\ln T_i$ take values $-\sigma, \sigma$.

Consequently for every bounded, measurable function ψ on $\mathbb{R}_{\geq 0}$ we obtain

$$E_Q(\psi(\sum_{i=1}^n \ln T_i)) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(-\frac{\sigma^2}{2} + \sigma x) e^{-\frac{x^2}{2}} dx.$$

2.17. Example. We can write down in a concrete example the relevant quantities. Take $S_0 = 1, a = -\frac{1}{2}$ and $b = 1$, $r = 0$. In this case we want to calculate the attainable claims $\mathcal{K}$. We do this by calculating all stochastic integrals: We calculate first the increments and write them as vectors in $\mathbb{R}^4$,

$$(S_1 - S_0) = \begin{pmatrix} 1 \\ 1 \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}, (S_2 - S_1) = \begin{pmatrix} 2 \\ -1 \\ \frac{1}{2} \\ -\frac{1}{4} \end{pmatrix}.$$

Predictable processes are given by ϕ_1 and ϕ_2,

$$\phi_1 = \begin{pmatrix} a \\ a \\ a \\ a \end{pmatrix}, \phi_1 = \begin{pmatrix} b \\ b \\ c \\ c \end{pmatrix}$$

for real numbers $4a, 4b, 4c$. Therefore

$$\mathcal{K} = \{ \begin{pmatrix} 4a + 8b \\ 4a - 4b \\ -2a + 2c \\ -2a - c \end{pmatrix} \text{ for } a, b, c \in \mathbb{R} \}.$$

This in turn is a 3-dimensional subspace which can be expressed by one equation namely

$$x_1 + 2x_2 + 2x_3 + 4x_4 = 0,$$

where we can directly read of the equivalent martingale measure Q

$$Q = \begin{pmatrix} \frac{1}{9} \\ \frac{2}{9} \\ \frac{2}{9} \\ \frac{4}{9} \end{pmatrix}$$

which demonstrates the assertions.

To understand the situation for incomplete markets we have to work with the notion of cones and duality relations as described in Theorem 4.

2.18. Theorem (incomplete markets)**.** *Let S be discrete model of a financial market with $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$. Then the following assertions are equivalent:*

1. *S is incomplete financial market.*
2. *For every claim X there is a self-financing trading strategy ϕ such that the claim can be super-replicated, i.e.*
$$V_N(\phi) \geq X$$
and there is at least one claim, which cannot be replicated.
3. *For every claim X there is a predictable process ϕ and a unique number x such that the discounted claim can be super-replicated, i.e.*
$$\widetilde{X} = \frac{1}{S_N^0} X \leq x + (\phi \cdot \widetilde{S})$$
and there is at least one claim, which cannot be replicated.

In particular we have that the no arbitrage prices at time 0 *form an open interval* $]\pi_\downarrow(X), \pi_\uparrow(X)[$ *if* $\pi_\downarrow(X) < \pi_\uparrow(X)$ *with*

$$\pi_\downarrow(X) = \inf\{E_Q(\frac{X}{S_N^0}) \text{ for } Q \in \mathcal{M}^e(\widetilde{S})\},$$
$$\pi_\uparrow(X) = \sup\{E_Q(\frac{X}{S_N^0}) \text{ for } Q \in \mathcal{M}^e(\widetilde{S})\}.$$

The case $\pi(X)_\downarrow = \pi(X)_\uparrow$ (there is only one no-arbitrage price for the claim X) occurs if and only if X can be replicated.

Proof. We assume that the market is arbitrage-free. We then know that $\mathcal{M}^a(\widetilde{S}) = \langle Q_1, \ldots, Q_m \rangle_{convex}$ for $m \geq 2$ since the market is incomplete. The polar cone of C is generated by $\frac{dQ_1}{dP}, \ldots, \frac{dQ_m}{dP}$,

$$C^0 = \left\langle \frac{dQ_1}{dP}, \ldots, \frac{dQ_m}{dP} \right\rangle_{con}.$$

This means that we can find for each claim X numbers x such that

$$E_{Q_i}(\widetilde{X} - x) \leq 0$$

for $i = 1, \ldots, m$, so by the bipolar theorem

$$\widetilde{X} = \frac{1}{S_N^0} X = x + (\phi \cdot \widetilde{S}) - Y \leq x + (\phi \cdot \widetilde{S})_N,$$

where Y is some non-negative random variable. If every claim could be replicated, then by the main theorem on complete markets the martingale measure has to be unique. The equivalence of the second and third assertion is obvious. Given the super-replication principle of the third (or second assertion), we can immediately conclude by the main theorem on complete markets.

For the additional assertions we have to go into theory again. First we show that under the assumption of replication

$$\widetilde{X} = x + (\phi \cdot \widetilde{S})_N$$

there is only one price. This follows immediately from

$$E_Q(\frac{X}{S_N^0}) = x$$

for all $Q \in \mathcal{M}^e(\widetilde{S})$ by Doob's theorem. Assume that $\pi_\downarrow(X) < \pi_\uparrow(X)$ and that there is an arbitrage-free pricing rule $\pi(X)$ with $\pi(X)_0 \geq \pi_\downarrow(X)$. Then there is equivalent martingale measure $Q \in \mathcal{M}^e(\widetilde{S})$ such that $\pi(X)_0 = E_Q(\frac{X}{S_N^0})$, hence $\pi(X)_0 = \pi_\downarrow(X)$ by the no-arbitrage assumption. Therefore $E_Q(\frac{X}{S_N^0} - \pi_\downarrow(X)) \leq 0$ and so there is a predictable strategy ϕ such that $(\phi \cdot \widetilde{S})_N \geq \frac{X}{S_N^0} - \pi_\downarrow(X)$, but

$$0 = E_Q((\phi \cdot \widetilde{S})_N) \geq E_Q(\frac{X}{S_N^0} - \pi_\downarrow(X)) = 0$$

by assumption. Consequently $(\phi \cdot S)_N = \frac{X}{S_N^0} - \pi_\downarrow(X)$ and so $\pi_\downarrow(X) = \pi_\uparrow(X)$, which is contradiction. Therefore $\pi(X)_0 < \pi_\downarrow(X)$. For the inequality $\pi(X)_0 > \pi_\downarrow(X)$ we argue by the negative claim $-X$. If now the pricing interval degenerates, i.e.

$$\{x\} = \{E_Q(\frac{X}{S_N^0}) \text{ for } Q \in \mathcal{M}^e(\widetilde{S})\},$$

then we know that $E_Q(\frac{X}{S_N^0} - x) = 0$ for all $Q \in \mathcal{M}^a(\widetilde{S})$ and therefore there is a predictable strategy ϕ such that

$$\frac{X}{S_N^0} - x = (\phi \cdot \widetilde{S})_N.$$

□

2.19. Example (incomplete market)**.** We give ourselves a stochastic process representing a risky asset with 2 periods,

$$S_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, S_1 = \begin{pmatrix} 3 \\ 3 \\ 3 \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}, S_2 = \begin{pmatrix} 9 \\ 4 \\ 1 \\ 1 \\ \frac{1}{9} \end{pmatrix}$$

and interest rate $r = 0$. This leads to the increments

$$S_1 - S_0 = \begin{pmatrix} 2 \\ 2 \\ 2 \\ -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix}, S_2 - S_1 = \begin{pmatrix} 6 \\ 1 \\ -2 \\ \frac{2}{3} \\ -\frac{2}{9} \end{pmatrix}$$

and therefore

$$\mathcal{K} = \{ \begin{pmatrix} 2a + 6b \\ 2a + b \\ 2a - 2b \\ -\frac{2}{3}a + \frac{2}{3}c \\ -\frac{2}{3}a - \frac{2}{9}c \end{pmatrix} \text{ for } a, b, c \in \mathbb{R}\}.$$

The set of outcomes with 0 initial investment can be characterized by the two equations

$$x_1 + 3x_3 + 3x_4 + 9x_5 = 0,$$
$$8x_2 + 4xl3 + 9x_4 + 27x_5 = 0.$$

The set of outcomes with 0 initial investment and some consumption is characterized by

$$x_1 + 3x_3 + 3x_4 + 9x_5 \leq 0,$$
$$8x_2 + 4x_3 + 9x_4 + 27x_5 \leq 0.$$

Hence the set of absolutely continuous martingale measures is given by

$$\mathcal{M}^a(\widetilde{S}) = \mathcal{M}^a(S) = \{Q_t := t \begin{pmatrix} \frac{1}{16} \\ 0 \\ \frac{3}{16} \\ \frac{3}{16} \\ \frac{9}{16} \end{pmatrix} + (1-t) \begin{pmatrix} 0 \\ \frac{8}{48} \\ \frac{4}{48} \\ \frac{9}{48} \\ \frac{27}{48} \end{pmatrix} \text{ for } t \in [0,1]\}.$$

The set of equivalent martingale measures is given by

$$\mathcal{M}^e(\widetilde{S}) = \{Q_t \text{ for } t \in]0,1[\}.$$

In this example we nicely calculate the pricing interval for a European call with strike price $K = 6$ and $N = 2$. This yields the payoff

$$X = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and consequently

$$\pi_{\uparrow}(X) = \frac{1}{16},$$
$$\pi_{\downarrow}(X) = 0.$$

The set of arbitrage-free prices is therefore given by $]0, \frac{1}{16}[$. The price $\frac{1}{16}$ is the smallest price for super-replication and one can easily calculate the super-replicating strategy.

Given a financial market $(S_n^0, S_n^1, \ldots, S_n^d)_{n=0,\ldots,N}$ on a probability space $(\Omega, \mathcal{F}, P)$ with filtration $(\mathcal{F}_n)_{n=0,\ldots,N}$. Without restriction we assume $d = 1$ and $S_n^0 = 1$ for $n = 0, \ldots, N$, since

$$\mathcal{M}^a(\frac{S^1}{S^0}, \ldots, \frac{S^d}{S^0}) = \cap_{i=1}^d \mathcal{M}^a(\frac{S^i}{S^0}).$$

So if we are able to calculate the martingale measures for a one asset model, we can do it in general easily for an $\mathbb{R}^d$-valued process.

2.20. Example. Take now the defining definition for a martingale $(S_n)_{n=0,\ldots,N}$, namely

$$E_Q(S_n | S_{n-1} = x) = x$$

for all values x of S_{n-1} on Ω. We define for all values x of S_{n-1} the probabilities

$$E_Q(1_{\{A^n\}} | S_{n-1} = x) = q^x(A^n),$$

which is 0 if $S_{n-1}(A^n) \neq \{x\}$ for the atoms A^n of $\mathcal{F}_n$. Then we get the equations

$$\sum_{A^n \in \mathcal{A}(\mathcal{F}_n)}^{|\Omega|} q^x(A^n) S_n(A^n) = x,$$
$$\sum_{i=1}^{|\Omega|} q^x(A^n) = 1,$$
$$q^x(A^n) \geq 0,$$

which can be solved if the model is arbitrage-free. The martingale measures Q is then given through

$$Q(A^n) = E_Q(1_{A^n}) = E_Q(1_{A^n} | S_{n-1} = x) Q(S_{n-1} = x)$$

if $S_{n-1}(A^n) = \{x\}$. In the next step one reduces time by 1 and one does the same sort of calculus for the atoms of $\mathcal{F}_{n-1}$. By induction we arrive at $n = 0$, where from we can restart to calculate back all the absolutely continuous martingale measures.

Assume now that $\mathcal{M}^a(S^1) = \langle Q_1, \ldots, Q_m \rangle$ for $m \geq 1$ (both cases are included, complete or incomplete), then we want to calculate (super)replicating strategies. Given a claim X there is one Q_i for some $i \in \{1, \ldots, m\}$ such that

$$\pi_{\uparrow}(X) = E_{Q_i}(X),$$

which is trivial in the complete case and requires some reasoning in the incomplete one. Then calculate the conditional expectations of X with respect to Q_i

$$X_n := E_{Q_i}(X | \mathcal{F}_n)$$

for $n = 0, \ldots, N$. The difference $X_n - X_{n-1}$ for $n = 1, \ldots, N$ is then

$$X_n - X_{n-1} = \phi_n (S_n - S_{n-1})$$

for some predictable process ϕ_n, which can be easily calculated from this equation for $n = 1, \ldots, N$.

2.21. Example. We apply the receipt from the above summary to calculate the equivalent martingale measures for a trinomial tree with two steps (which would be somehow tiresome in the traditional way):

$$S_2 := \begin{cases} 3S_1 \\ 2S_1 \\ \frac{1}{2}S_1 \end{cases}, S_1 := \begin{cases} 3 \\ 2 \\ \frac{1}{2} \end{cases}$$

and $S_0 = 1$. S_1 can take three values $\frac{1}{2}, 2$ and 3. The random variable S_2 can take six values $\frac{1}{4}, 1, \frac{3}{2}, 4, 6$ and 9. We calculate at each node: Take $S_1 = 3$, $n = 2$, then we have to solve an equation of the type

$$q_1^3 3 + q_2^3 2 + q_3^3 \frac{1}{2} = 1$$
$$q_1^3 + q_2^3 + q_3^3 = 1$$
$$q_i^3 \geq 0$$

by factoring out $S_1 = 3$. For all other values of S_1 we obtain the same equation. The solution is given through

$$\begin{aligned} q_1 &= t \\ q_2 &= \frac{1}{3} - \frac{5}{3}t \\ q_3 &= \frac{2}{3} + \frac{2}{3}t \end{aligned}$$

for $0 \leq t \leq \frac{1}{5}$. Hence we obtain the following vector for martingale measures

$$Q(t_1, t_2, t_3, t_4) := \begin{pmatrix} t_1 t_2 \\ t_1(\frac{1}{3} - \frac{5}{3}t_2) \\ t_1(\frac{2}{3} + \frac{2}{3}t_2) \\ (\frac{1}{3} - \frac{5}{3}t_1)t_3 \\ (\frac{1}{3} - \frac{5}{3}t_1)(\frac{1}{3} - \frac{5}{3}t_3) \\ (\frac{1}{3} - \frac{5}{3}t_1)(\frac{2}{3} + \frac{2}{3}t_3) \\ (\frac{2}{3} + \frac{2}{3}t_1)t_4 \\ (\frac{2}{3} + \frac{2}{3}t_1)(\frac{1}{3} - \frac{5}{3}t_4) \\ (\frac{2}{3} + \frac{2}{3}t_1)(\frac{2}{3} + \frac{2}{3}t_4) \end{pmatrix}$$

for $0 \leq t_i \leq \frac{1}{5}$ and $i = 1, \dots, 4$. We get a nice 4 parameter family of probability measures, which can be written as convex hull.

We want to make clear in the sequel that the choice of the bank-account process was arbitrary. We can replace it by any strictly positive process in the financial market. This technique is particularly useful for the calculation of prices and is called "change of numeraire". Thinking about the market in dollars or euros should equivalent.

2.22. Remark. First we observe that the wealth process $(V_n(\phi))_{0 \leq n \leq N}$ of every portfolio ϕ, which finally produces the attainable claims $V_N(\phi)$, is equal to the price process $(\pi(V_N)_n)_{0 \leq n \leq N}$, which in turn is a Q-martingale for each $Q \in \mathcal{M}^e(\widetilde{S})$ if we discount with respect to S^0. If the portfolio is strictly positive we can take it as a new unit of calculation and calculate prices hence with respect to this portfolio. The formulas reflect this.

Therefore we interpret the above results with respect to numeraires for a discrete model of a financial market S.

- the discounted processes $(\frac{S_n^1}{S_n^0}, \dots, \frac{S_n^d}{S_n^0})_{0 \leq n \leq N}$ have to satisfy the condition that there is an equivalent martingale measure.
- calculating the price of a claim X at time n amounts to calculating the conditional expectation of the discounted claim $\frac{S_n^0 X}{S_N^0}$, discounted for time n. If one would discount this price to time 0 one obtains that the discounted price process is again a martingale.
- hence calculations are always done with respect to a numeraire and a discounting procedure to some fixed time 0. This is economically obvious since one cannot compare amounts of money at different times without taking care of the time changes in value of the money.

2.23. Definition. *Let S be a discrete model of a financial market. A* numeraire process $(C_n)_{0 \leq n \leq N}$ *is a strictly positive, adapted stochastic process, which has the* *property that there is a self-financing trading strategy* ϕ *such that* $C_n = V_n(\phi)$ *for* $0 \leq n \leq N$.

2.24. Theorem. *Let* S *be a discrete model of a financial market and* C *a numeraire process. If* S *is arbitrage-free, then also the market*

$$S_n^C = (1, \frac{S_n^0}{C_n}, \frac{S_n^1}{C_n}, \dots, \frac{S_n^d}{C_n})$$

for $n = 0, \dots, N$ *is arbitrage-free and we have*

$$\mathcal{M}^e(S^C, 1) = \{Q^C | \quad \frac{dQ^C}{dP} = \frac{1}{S_N^0} \frac{C_N}{C_0} \frac{dQ}{dP} \text{ for } Q \in \mathcal{M}^e(S, S^0)\}.$$

Every pricing rule π^C *for contingent claims* X *is of the form*

$$\pi^C(X)_n = E_Q(\frac{C_n}{C_N} X | \mathcal{F}_n)$$

for any $Q \in \mathcal{M}^e(S^C, 1)$ *and the discounted processes with respect to the numeraire* C *are martingales.*

2.25. Remark. By the last formula we see that calculating the price in the numeraire C yields a price

$$\pi^C(X)_n = E_{Q^C}(\frac{C_n}{C_N} X | \mathcal{F}_n)$$

for a contract X expiring at N at time n in the numeraire C. If we change the numeraire back to S^0, then we have to price X at time N with respect to the old numeraire. If we calculate the price in the numeraire S^0 we have to perform

$$\pi(X)_n = E_Q(\frac{S_n^0}{S_N^0} X | \mathcal{F}_n).$$

Are these prices the same if we change the numeraire? They have to be equal and indeed if we assume that Q^C is obtained from Q with the formula

$$\frac{dQ^C}{dP} = \frac{1}{S_N^0} \frac{C_N}{C_0} \frac{dQ}{dP},$$

then we obtain by Lemma 4.7 and the fact that

$$E_Q(\frac{1}{S_N^0} \frac{C_N}{C_0} | \mathcal{F}_n) = \frac{C_n}{C_0 S_n^0}$$

the final relation

$$\begin{aligned}
\pi(X)_n &= C_0 E_Q(\frac{S_n^0}{S_N^0} \frac{1}{C_0} X | \mathcal{F}_n) \\
&= C_n \frac{C^0 S_n^0}{C_n} E_Q(\frac{1}{S_N^0} \frac{1}{C_0} X | \mathcal{F}_n) \\
&= E_{Q^C}(\frac{C_n}{C_N} X | \mathcal{F}_n) \\
&= \pi^C(X)_n
\end{aligned}$$

as it should be.

PROOF. We have to show first that the measures Q^C are measures, which is – by the above remark – true since we know that discounted numeraire $(S_n^0)^{-1}C_n$ is a Q-martingale for $Q \in \mathcal{M}^e(S, S^0)$, therefore in particular

$$E_Q((S_N^0)^{-1}C_N) = C_0.$$

By strict positivity the measures are equivalent to P. By Lemma 4.7 we can calculate the conditional expectations with respect to Q^C of the discounted price processes

$$\begin{aligned} E_{Q^C}(\frac{S_N^i}{C_N}|\mathcal{F}_n) &= \frac{C_0 S_n^0}{C_n} E_Q(\frac{1}{S_N^0}\frac{C_N}{C_0}\frac{S_N^i}{C_N}|\mathcal{F}_n) \\ &= \frac{S_n^0}{C_n} E_Q(\frac{S_N^i}{S_N^0}|\mathcal{F}_n) \\ &= \frac{S_n^i}{C_n} \end{aligned}$$

for $0 \leq n \leq N$, since $(\frac{S_n^i}{S_n^0})_{0 \leq n \leq N}$ is a Q-martingale. Therefore the market is arbitrage-free by the main theorem and the measures Q^C are all martingale measures by reversing the calculations. The pricing formulas are clear by general theory. □

In the sequel we shall formulate most of the assertions with respect to a basis in $L^2(\Omega, \mathcal{F}, P)$. We shall assume (which is in our case not a real restriction), that $\mathcal{F} = 2^\Omega$ and $P(\omega_i) > 0$ for $i = 1, \ldots, |\Omega|$. We choose $(1_{\{\omega\}})_{\omega \in \Omega}$ and identify $L^2(\Omega, \mathcal{F}, P)$ with some $\mathbb{R}^{|\Omega|}$. Hence we can apply our duality theory for cones.

2.26. Proposition. *Let S be a discrete model for a financial market and assume $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$. Then there are linearly independent measures $Q_1, \ldots, Q_n$ such that*

$$\mathcal{M}^a(\widetilde{S}) = \langle Q_1, \ldots, Q_n \rangle_{convex},$$

the polar cone C^0 equals

$$C^0 = \left\langle \frac{dQ_1}{dP}, \ldots, \frac{dQ_n}{dP} \right\rangle_{cone}.$$

Furthermore the Q_i have at least $n-1$ zeros, where n equals the codimension of $\mathcal{K}$.

PROOF. The polar cone C^0 is polyhedral and therefore generated by finitely many elements $\frac{dQ_1}{dP}, \ldots, \frac{dQ_m}{dP}$. The set of absolutely continuous martingale measures $\mathcal{M}^a(\widetilde{S})$ is given by taking the correct normalization, since $C^0 \subset L^2_{\geq 0}(\Omega, \mathcal{F}, P)$. The codimension of $\mathcal{K}$ is denoted by n. We choose a maximal set of linearly independent measures $Q_1, \ldots, Q_r \in \mathcal{M}^a(\widetilde{S})$ (after reordering). We claim that $r = n$ and $\mathcal{K}^0 = \left\langle \frac{dQ_1}{dP}, \ldots, \frac{dQ_n}{dP} \right\rangle_{vector}$. Given Q_i for $i = 1, \ldots, r$, the expectation $E_{Q_i}(X) = E_P(\frac{dQ}{dP}X) = 0$ for $X \in \mathcal{K}$ by assumption, so $\left\langle \frac{dQ_1}{dP}, \ldots, \frac{dQ_n}{dP} \right\rangle_{vector} \subset \mathcal{K}^0$. Given $X \in \mathcal{K}^0$ such that $E_{Q_i}(X) = 0$ for $i = 1, \ldots, r$, we know by maximality of the set $Q_1, \ldots, Q_r$, that every extremal point Q_i for $i = 1, \ldots, m$ is a linear combination of $Q_1, \ldots, Q_r$, hence

$$E_{Q_i}(X) = 0 \text{ for } i = 1, \ldots, m$$

and therefore $X \in \mathcal{K}$, which yields $X = 0$. In particular $n \leq m$.

We want to prove $n = m$ and we proceed by induction with respect to $|\Omega|$. More precisely, we prove by induction that $m = n$ and that there is a permutation $\pi \in \mathfrak{S}_n$ such that for $i, j \in \{1, \dots, n\}$

$$\frac{dQ_{\pi(i)}}{dP}(\omega_i) > 0 \text{ for } j = i,$$
$$\frac{dQ_{\pi(i)}}{dP}(\omega_j) = 0 \text{ for } j \neq i$$

and

$$\#\{i | Q_i(\omega_k) = 0\} = \begin{cases} n-1 \\ 0 \end{cases}$$

holds.

The result holds for $|\Omega| = 2$. Fix $|\Omega| > 2$ and take $\mathcal{K} \subset L(\Omega, \mathcal{F}, P)$ with $\mathcal{K} \cap L(\Omega, \mathcal{F}, P) = \{0\}$. We know that there are $m \geq n$ extremal martingale measures. Without restriction we assume that the first component of one extremal martingale measure vanishes. The projection

$$p_1 : \mathbb{R}^{|\Omega|} \to \mathbb{R}^{|\Omega|-1}$$
$$(x_1, \dots, x_{|\Omega|}) \mapsto (x_2, \dots, x_{|\Omega|})$$

is injective on $\mathcal{K}$ by assumption. The polar cone of $p_1(\mathcal{K}) - \mathbb{R}^{|\Omega|-1}_{\geq 0}$ is generated by linearly independent extremal measures $\frac{dQ'_2}{dP'}, \dots, \frac{dQ'_n}{dP'}$ with the above conditions by the the induction hypothesis (notice that $p_1(\mathcal{K}) \cap \mathbb{R}^{|\Omega|-1}_{\geq 0} = \{0\}$). The obvious extensions $Q_2, \dots, Q_n$ via a vanishing first component are extremal martingale measures for the original problem. There is an extremal measure Q_1 with a non-vanishing first component by no arbitrage. The set $\frac{dQ_1}{dP}, \frac{dQ_2}{dP}, \dots, \frac{dQ_n}{dP}$ is then a basis of $\mathcal{K}^0$ by the first observation. Here we fix a numbering of the extremal martingale measures of $\mathcal{K}$ by $Q_1, \dots, Q_m$ with $m \geq n$.

Given ω_k with $Q_1(\omega_k) = 0$, then we either obtain $p_k(Q_i)$ for some $i = 1, \dots, m$ (except one) as extremal martingale measures for $p_k(\mathcal{K})$ or we need some additional $p_k(Q_{j_0})$ for some $j_0 = n+1, \dots, m$. The second case only occurs if there are two $i_1 \neq i_2 \in \{2, \dots, n\}$ such that the k-th component of Q_{i_1} and Q_{i_2} does not vanish, hence a contradiction to the assumption. Vice versa doing the construction with the $n-1$ zeros of $Q_2, \dots, Q_n$ we obtain at least $n-1$ zeros for Q_1. Consequently $Q_i, Q_2, \dots, Q_n$ satisfy the above properties, therefore $Q_{n+1}, \dots, Q_m$ are convex combinations of $Q_i, Q_2, \dots, Q_n$. This means – by construction – $m = n$. □

3. The baby examples of optimization

We consider the complete and incomplete case in a one period model with a general utility function and some particular examples. This section is preparatory and should provide a feeling for the type of problem, which we are going to treat.

3.1. Definition. *A real valued function* $u : I \to \mathbb{R}$ *is called* utility function *if* $I =]0, \infty[$ *or* $I =]-\infty, \infty[$ *and* u *is an increasing, strictly concave* C^2*-function. We shall denote* $\operatorname{dom}(u) := I$ *and we define* $u(x) = -\infty$ *for* $x \notin dom(u)$*. Furthermore we shall assume that* $\lim_{x \downarrow 0} u(x) = -\infty$ *if* $\operatorname{dom}(u) =]0, \infty[$.

3.2. Remark. In the sequel we shall impose further conditions on utility functions guaranteeing the existence of optimal solutions. For the presentation of the problem this is not necessary.

We consider a financial market $(S_n^0, \dots, S_n^d)_{n=0,1}$ on $(\Omega, \mathcal{F}, P)$ with one period and aim to solve the following optimization problem for a given utility function $u : \mathrm{dom}(u) \to \mathbb{R}$ and $x \in \mathrm{dom}(u)$.

$$E_P(u(\frac{1}{S_1^0} V_1(\phi)) \to \max,$$
$$V_0(\phi) = x,$$

where ϕ is running over all self-financing trading strategies. This leads to the following one dimensional optimization problem

$$a \mapsto E_P(u(x + a(\widetilde{S}_1 - \widetilde{S}_0))),$$

which can be solved by classical analysis. We see immediately that the existence of an optimal strategy $\widehat{a}(x)$ for a fixed $x \in dom(u)$ leads to

$$E_P(u'(x + \widehat{a}(x)(\widetilde{S}_1 - \widetilde{S}_0))(\widetilde{S}_1 - \widetilde{S}_0)) = 0.$$

This is in turn means that the vector can be normalized to a probability measure Q, i.e.

$$\frac{dQ}{dP} = \frac{1}{\lambda} u'(x + \widehat{a}(x)(\widetilde{S}_1 - \widetilde{S}_0)),$$

which is a martingale measure since $E_Q(\widetilde{S}_1 - \widetilde{S}_0) = 0$. Therefore the existence of an optimizer leads to arbitrage-free markets.

4. Basic concepts of Utility optimization

Given a financial market $(S_n^0, \dots, S_n^d)_{n=0,\dots,N}$ on $(\Omega, \mathcal{F}, P)$ and a utility function u, then we define the *utility optimization problem* as determination of $U(x)$ for $x \in \mathrm{dom}(u)$, i.e.

$$\sup_{\substack{\phi \text{ trading strategy} \\ \phi \text{ self financing} \\ V_0(\phi)=x}} E(u(\frac{1}{S_N^0} V_N(\phi)) =: U(x).$$

We say that the utility optimization problem at $x \in \mathrm{dom}(u)$ is *solvable* if $U(x)$ is finitely valued and if we find an optimal self financing trading strategy $\widehat{\phi}(x)$ for $x \in \mathrm{dom}(u)$ such that

$$U(x) = E(u(\frac{1}{S_N^0} V_N(\widehat{\phi}(x))),$$
$$V_0(\widehat{\phi}(x)) = x.$$

We shall introduce three methods for the solution of the utility optimization problem, where the number of variables involved differ.

We assume that $\mathcal{F} = 2^\Omega$ and $P(\omega) > 0$ for $\omega \in \Omega$. We then have three characteristic dimensions: the dimension of all random variables $|\Omega|$ (the number of paths), then the dimension of discounted outcomes at initial wealth 0, denoted by $\dim \mathcal{K}$, and the number of martingale measures m. We have the basic relation

$$m + \dim \mathcal{K} = |\Omega|.$$

- the pedestrian method is an unconstraint extremal value problem in $\dim \mathcal{K}$ variables.
- the Lagrangian method yields an unconstraint extremal value problem in $|\Omega| + m$ variables.

- the duality method (martingale approach) yields an unconstraint extremal value problem in m variables. Additionally one has to transform the dual value function to the original, which is a one dimensional extremal value problem.

In financial mathematics usually $\dim \mathcal{K} \gg m$, which means that the duality method is of particular importance.

4.1. Pedestrian's method. We can understand utility optimization as unrestricted optimization problem. Define $\mathcal{S}$ the vector space of all predictable strategies $(\phi_n)_{n=0,\ldots,N}$, then the utility optimization problem for $x \in \operatorname{dom}(u)$ is equivalent to solving the following problem

$$F_x : \begin{cases} \mathcal{S} \to \mathbb{R} \cup \{-\infty\} \\ (\phi_n)_{n=0,\ldots,N} \mapsto E(u(x + (\phi \cdot \widetilde{S})_N)) \end{cases}$$

$$\sup_{\phi \in \mathcal{S}} F_x(\phi) = U(x)$$

This is an ordinary extremal value problem for every $x \in \operatorname{dom}(u)$. Let $(\widehat{\phi_n})_{n=0,\ldots,N}$ be an optimal strategy, then necessarily

$$\operatorname{grad} F_x((\widehat{\phi_n})_{n=0,\ldots,N}) = 0$$

and therefore we can in principle calculate the optimal strategy. From this formulation we take one fundamental conclusion.

4.1. Theorem. *Let the utility optimization problem at $x \in \operatorname{dom}(u)$ be solvable and let $(\widehat{\phi_n})_{n=0,\ldots,N}$ be an optimal strategy, so*

$$\sup_{\phi \in \mathcal{S}} F_x(\phi) = U(x) = F_x(\widehat{\phi}),$$

then $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$.

Proof. We calculate the directional derivative with respect to 1_A for $A \in \mathcal{A}(\mathcal{F}_{i-1})$ for $i = 1, \ldots, N$,

$$\begin{aligned} &\frac{d}{ds}|_{s=0} E(u(x + (\widehat{\phi} \cdot \widetilde{S})_N + s 1_A \Delta S_i)) \\ &= E(u'(x + (\widehat{\phi} \cdot \widetilde{S})_N) 1_A \Delta S_i). \end{aligned}$$

Since $(\widehat{\phi_n})_{n=0,\ldots,N}$ is an optimizer we necessarily have that the directional derivatives in direction of the elements $1_A \Delta S_i$ vanish. We define

$$\lambda := E(u'(x + (\widehat{\phi} \cdot \widetilde{S})_N)) > 0$$

since $u'(y) > 0$ for $y \in \operatorname{dom}(U)$. Consequently

$$\frac{dQ}{dP} := \frac{1}{\lambda} u'(x + (\widehat{\phi} \cdot \widetilde{S})_N)$$

defines a probability measure equivalent to P. Hence we obtain from the gradient condition that

$$E_Q(1_A(S_i - S_{i-1})) = 0$$

for all $A \in \mathcal{A}(\mathcal{F}_{i-1})$ and $i = 1, \ldots, N$, which means

$$E(S_i | \mathcal{F}_{i-1}) = S_{i-1}$$

for $i = 1, \ldots, N$, therefore $Q \in \mathcal{M}^e(\widetilde{S})$. □

Besides baby examples the pedestrian's method is not really made for the solution of the utility optimization problem, since equations become very complicated and the internal structure does not really get clear. Nevertheless the above conclusion is of high importance, since it will be a basic assumption from now on.

4.2. Condition. *We shall always assume* $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$.

Furthermore we can easily formulate a basis existence and regularity result by the pedestrian's method (which allows to make nice general conclusions).

4.3. Proposition. *Assume* $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$ *and* $\lim_{x\to\infty} u'(x) = 0$ *if* $\operatorname{dom}(u) = \mathbb{R}$, *then the utility optimization problem for* $x \in \operatorname{dom}(u)$ *has a unique solution* $\widehat{X}(x) \in x + \mathcal{K}$, *which is also the unique local maximum, and* $x \mapsto \widehat{X}(x)$ *is* C^1 *on* $\operatorname{dom}(u)$. *If* $x \notin \operatorname{dom}(u)$, *then* $\sup_{\phi\in\mathcal{S}} F_x(\phi) = -\infty$.

Proof. The functional $X \mapsto E_P(u(X))$ is C^2, strictly concave and increasing. Assume that there are two optimizers $\widehat{X}_1(x) \neq \widehat{X}_2(x) \in x + \mathcal{K}$, then

$$E_P(u(t\widehat{X}_1(x) + (1-t)\widehat{X}_2(x))) > tE_P(u(\widehat{X}_1(x))) + (1-t)E_P(u(\widehat{X}_2(x))) = U(x)$$

for $t \in]0,1[$, which is a contradiction. The argument also yields that two local maxima have to coincide. Therefore the optimizer is also the unique local maximum.

Since $\widetilde{S}$ is a martingale, the space $\mathcal{K}$ of outcomes with zero investment has the property that for $X \in L^2(\Omega, \mathcal{F}, P)$

$$X \in \mathcal{K} \Longleftrightarrow E_Q(X) = 0$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$. Given an equivalent martingale measure $Q \in \mathcal{M}^e(\widetilde{S})$, then we prove that for any $x \in \operatorname{dom}(u)$

$$\lim_{\substack{Y\in\mathcal{K} \\ E_Q(|Y|)\to\infty}} E_P(u(x+Y)) = -\infty.$$

Assume that it were bounded from below by M, so we can find $Y_n \in \mathcal{K}$ such that $E_P(u(x+Y_n)) \geq M$ and $E_Q(|Y|) \geq n$. Since $Y_n \in \mathcal{K}$ we have

$$E_Q(Y_n) = 0$$

and Y_n has positive and negative components. Hence

$$E_Q((Y_n)_+) \geq \frac{n}{2}, -E_Q((Y_n)_-) \geq \frac{n}{2}.$$

We can choose the sequence Y_n such that the smallest components form a sequence decreases to $-\infty$ and the sequence of largest components form a sequence increasing to ∞. We have

$$|\frac{\max Y_n}{\min Y_n}| \leq M_1 < \infty$$

for all $n \geq 1$. If $\operatorname{dom}(u) =]0,\infty[$, the assertion is trivial since $-\infty$ is reached after finitely many steps. If $\operatorname{dom}(u) = \mathbb{R}$, then

$$E_P(u(x+Y)) \leq E_P(\max u(Y_n)) - E_P(u(Y_n)_-) \leq u(a_n) - b_n u(c_n)$$

with $a_n \uparrow \infty$ (largest component of Y_n), $c_n \downarrow -\infty$ (smallest component of Y_n), $b_n \in]\epsilon, 1]$ (probability $Q(Y_n = \min Y_n) > 0$) and $|\frac{a_n}{c_n}| \leq M_1$. Since obtain the result, since u' increases in negative direction strictly more than in positive direction.

Consequently the function $Y \mapsto E_P(u(x+Y))$ has a maximum on $\mathcal{K}$.

If $x \notin \mathrm{dom}(u)$, then for any $Y \in \mathcal{K}$, there are negative components and therefore $E_P(u(x+Y)) = -\infty$.

For the regularity assertion we take a basis of $\mathcal{K}$ denoted by $(f_i)_{i=1,\dots,\dim\mathcal{K}}$ and calculate the derivative with respect to this basis at the unique existing optimizer $\widehat{Y}(x) = \widehat{X}(x) - x$,

$$E_P(u'(x+\widehat{Y}(x))f_i) = 0$$

for $i = 1, \dots, \dim\mathcal{K}$. Calculating the second derivative we obtain the matrix

$$(E_P(u''(x+Y)f_i f_j))_{i,j=1,\dots,\dim\mathcal{K}}$$

which is invertible for any $Y \in \mathcal{K}$, since u'' is strictly negative. Therefore $x \mapsto \widehat{X}(x)$ is C^1 on $\mathrm{dom}(u)$. □

4.2. Duality methods. Since we have a dual relation between the set of martingale measures and the set $\mathcal{K}$ of claims attainable at price 0, we can formulate the optimization problem as constraint problem: for any $X \in L^2(\Omega, \mathcal{F}, P)$

$$X \in \mathcal{K} \Longleftrightarrow E_Q(X) = 0$$

for $Q \in \mathcal{M}^a(\widetilde{S})$ and for any probability measure Q

$$Q \in \mathcal{M}^a(\widetilde{S}) \Longleftrightarrow E_Q(X) = 0$$

for all $X \in \mathcal{K}$. Therefore we can formulate the problem as constraint optimization problem and apply the method of Lagrangian multipliers.

First we define a function $H : L^2(\Omega, \mathcal{F}, P) \longrightarrow \mathbb{R}$ via

$$H(X) := E_P(u(X))$$

for a utility function u. For $x \in \mathrm{dom}(u)$ we can formulate the constraints

$$U_x := \mathcal{K} + x = \{X \in L^2(\Omega, \mathcal{F}, P) \text{ such that } E_Q(X) = x \text{ for } Q \in \mathcal{M}^a(\widetilde{S})\}.$$

Consequently the utility optimization problem reads

$$\sup_{X \in U_x} E_P(u(X)) = U(x)$$

for $x \in \mathrm{dom}(u)$. Hence we can treat the problem by Lagrangian multipliers, i.e. if $\widehat{X} \in U_x$ is an optimizer, then

$$\text{(LM)} \qquad u'(\widehat{X}) - \sum_{i=1}^{m} \widehat{\eta}_i \frac{dQ_i}{dP} = 0$$

$$E_{Q_i}(\widehat{X}) = x$$

for $i = 1, \dots, m$, $\mathcal{M}^a(\widetilde{S}) = \langle Q_1, \dots, Q_m \rangle$ and some values $\widehat{\eta}_i$. This result is obtained by taking the gradient of the function

$$X \mapsto E_P(u(X) - \sum_{i=1}^{m} \eta_i (\frac{dQ_i}{dP} X - x))$$

with respect to some basis. We can choose the $\widehat{\eta}_i$ positive, since $u'(\widehat{X})$ represents a positive multiple of an equivalent martingale measure. Notice that by assumption $u'(x) > 0$ for all $x \in \mathrm{dom}(u)$, and $u'(\widehat{X})$ is finitely valued.

4.4. Lemma. *If $(\widehat{X}, \widehat{\eta_1}, \dots, \widehat{\eta_m})$ is a solution of the Lagrangian multiplier equation (LM), then the multipliers $\widehat{\eta}_i > 0$ are uniquely determined and $\sum_{i=1}^m \widehat{\eta}_i > 0$. Given $x \in \mathrm{dom}(u)$, the map $x \mapsto (\widehat{\eta}_i(x))_{i=1,\dots,m}$ is C^1.*

PROOF. By Proposition 2.26 we know that the $\widehat{\eta}_i$ are uniquely determined and the inverse function theorem together with the previous result yields the C^1-dependence. □

The Lagrangian $\widetilde{L}$ is given through

$$\widetilde{L}(X,\eta_1,\dots,\eta_m) = E_P(u(X)) - \sum_{i=1}^{m} \eta_i(E_{Q_i}(X) - x)$$

for $X \in L^2(\Omega,\mathcal{F},P)$ and $\eta_i \geq 0$. We introduce $y := \eta_1 + \dots \eta_m$ and $\mu_i := \frac{\eta_i}{y}$ (we can assume $y > 0$ since the value for η_i we are looking for has to satisfy $y > 0$). Therefore

$$L(X,y,Q) = E_P(u(X)) - y(E_Q(X) - x)$$

for $X \in L^2(\Omega,\mathcal{F},P)$, $Q \in \mathcal{M}^a(\widetilde{S})$ and $y > 0$. We define

$$\Phi(X) := \inf_{\substack{y>0 \\ Q\in\mathcal{M}^a(\widetilde{S})}} L(X,y,Q)$$

for $X \in L^2(\Omega,\mathcal{F},P)$ and

$$\psi(y,Q) = \sup_{X\in L^2(\Omega,\mathcal{F},P)} L(X,y,Q)$$

for $y > 0$ and $Q \in \mathcal{M}^a(\widetilde{S})$. We can hope for

$$\sup_{X\in L^2(\Omega,\mathcal{F},P)} \Phi(X) = \inf_{y>0} \inf_{Q\in\mathcal{M}^a(\widetilde{S})} \psi(y,Q) = U(x).$$

by a mini-max consideration.

4.5. Remark. Where does the minimax consideration stem from? Look at $X \mapsto \widetilde{L}(X,\eta_1,\dots,\eta_m)$ for fixed $\eta_1,\dots,\eta_m$, then we obtain something strictly concave as sum of two concave functions, where one is strictly concave. Look at $(\eta_1,\dots,\eta_m) \mapsto \widetilde{L}(X,\eta_1,\dots,\eta_m)$ for fixed $X \in L^2(\Omega,\mathcal{F},P)$, then we obtain something affine.

4.6. Lemma. *Let u be a utility function and $(S_n^0, S_n^1, \dots, S_n^d)_{n=0,\dots,N}$ be a financial market, which is arbitrage-free, then*

$$\sup_{X\in L^2(\Omega,\mathcal{F},P)} \Phi(X) = U(x).$$

PROOF. We can easily prove the following facts:

$$\Phi(X) = -\infty \text{ if } E_Q(X) > x$$

for at least one $Q \in \mathcal{M}^a(\widetilde{S})$. Furthermore

$$\Phi(X) = E_P(u(X)) \text{ if } E_Q(X) \leq x$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$. Consequently

$$\sup_{X\in L^2(\Omega,\mathcal{F},P)} \Phi(X) = \sup_{\substack{X\in L^2(\Omega,\mathcal{F},P) \\ E_Q(X)\leq x \text{ for } Q\in\mathcal{M}^a(\widetilde{S})}} E_P(u(X)) = U(x)$$

since u is increasing. □

For the application of the minimax theorem we need to calculate ψ, which is done in the next Lemma. Therefore we assume the generic conditions for conjugation as stated in the mathematics sections.

4.7. Lemma. *Given an arbitrage-free financial market $(S^0, \dots, S^d)$, the function*

$$\psi(y,Q) = \sup_{X \in L^2(\Omega,\mathcal{F},P)} L(X,y,Q)$$

can be expressed by the conjugate function v of u,

$$\psi(y,Q) = E_P(v(y\frac{dQ}{dP})) + yx.$$

PROOF. By definition we have

$$\begin{aligned} L(X,y,Q) &= E_P(u(X)) - y(E_Q(X) - x) \\ &= E_P(u(X) - y\frac{dQ}{dP}X) + yx. \end{aligned}$$

If we fix $Q \in \mathcal{M}^a(\widetilde{S})$ and $y > 0$, then the calculation of the supremum over all random variables yields

$$\begin{aligned} & E_P(u(X) - y\frac{dQ}{dP}X) \\ &= E_P(\sup_{X \in L^2(\Omega,\mathcal{F},P)} u(X) - y\frac{dQ}{dP}X) \\ &= E_P(v(y\frac{dQ}{dP})) \end{aligned}$$

by definition of the conjugate function. □

4.8. Definition. *Given the above setting we call the optimization problem*

$$V(y) := \inf_{Q \in \mathcal{M}^a(\widetilde{S})} E_P(v(y\frac{dQ}{dP}))$$

the dual problem and V the dual value function for $y > 0$.

Next we formulate that the dual optimization problem has a solution.

4.9. Lemma. *Let u be a utility function under the above assumptions and assume $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$, then there is a unique optimizer $\widehat{Q}(y)$ such that*

$$V(y) = \inf_{Q \in \mathcal{M}^a(\widetilde{S})} E_P(v(y\frac{dQ}{dP})) = E_P(v(y\frac{d\widehat{Q}(y)}{dP})).$$

Furthermore

$$\inf_{y>0}(V(y) + xy) = \inf_{\substack{y>0 \\ Q \in \mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP})) + xy).$$

PROOF. Since v is strictly convex, C^2 on $]0,\infty[$ and $v'(0) = -\infty$ we obtain by compactness the existence of an optimizer $\widehat{Q}(y)$ and by $v'(0) = -\infty$ that the optimizer is an equivalent martingale measure (since one can decrease the value of $v(y\frac{dQ}{dP})$ by moving away from the boundary). By strict convexity the optimizer is also unique. The gradient condition for $\widehat{Q}(y)$ reads as follows

$$E_P(v'(\widehat{Q}(y))(\frac{d\widehat{Q}(y)}{dP} - \frac{dQ}{dP})) = 0$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$. The function V shares the same qualitative properties as v and therefore we can define the concave conjugate. Fix $x \in \operatorname{dom}(u)$ and take the optimizer $\widehat{y} = \widehat{y}(x) > 0$, then

$$\begin{aligned}\inf_{y>0}(V(y)+xy) = V(\widehat{y}) + x\widehat{y} &\leq \inf_{Q\in\mathcal{M}^a(\widetilde{S})} E_P(v(y\frac{dQ}{dP})) + xy \\ &\leq E_P(v(y\frac{dQ}{dP})) + xy\end{aligned}$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$ and $y > 0$, so

$$\inf_{y>0}(V(y)+xy) \leq \inf_{\substack{y>0 \\ Q\in\mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP})) + xy).$$

Take $y_1 > 0$ and $Q_1 \in \mathcal{M}^e(\widetilde{S})$ for some $\epsilon > 0$ and assume that

$$\begin{aligned}\inf_{y>0}(V(y)+xy) + 2\epsilon &\geq V(y_1) + xy_1 + \epsilon \\ &\geq E_P(v(y_1\frac{dQ_1}{dP})) + xy_1 \\ &\geq \inf_{\substack{y>0 \\ Q\in\mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP})) + xy).\end{aligned}$$

Since this holds for every $\epsilon > 0$ we can conclude. □

4.10. Theorem. *Let $(S^0, \dots, S^d)$ be an arbitrage-free market and u a utility function with the above properties, then*

$$U(x) = \inf_{\substack{y>0 \\ Q\in\mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP})) + xy)$$

and the mini-max assertion holds.

Proof. Fix $x \in \operatorname{dom}(u)$ and take an optimizer $\widehat{X}$, then there are Lagrangian multipliers $\widehat{\eta_1}, \dots, \widehat{\eta_m} \geq 0$ such that $\widehat{y} := \sum_{i=1}^m \widehat{\eta_i} > 0$ and

$$\widetilde{L}(\widehat{X}, \widehat{\eta_1}, \dots, \widehat{\eta_m}) = U(x),$$

and the constraints are satisfied so $E_{Q_i}(\widehat{X}) = x$ and $\widehat{X}$ is an optimizer. We define a measure $\widehat{Q}$ via

$$u'(\widehat{X}) = \widehat{y}\frac{d\widehat{Q}}{dP}.$$

Since

$$u'(\widehat{X}) - \widehat{y}\sum_{i=1}^m \frac{\widehat{\eta_i}}{\widehat{y}}\frac{dQ_i}{dP} = 0$$

by the Lagrangian multipliers method, we see that

$$\widehat{y}\frac{d\widehat{Q}}{dP} = \widehat{y}\sum_{i=1}^m \frac{\widehat{\eta_i}}{\widehat{y}}\frac{dQ_i}{dP}$$

and therefore $\widehat{Q} \in \mathcal{M}^e(\widetilde{S})$ (its Radon-Nikodym derivative is strictly positive). Furthermore

$$E_P(v(\widehat{y}\frac{d\widehat{Q}}{dP})) + x\widehat{y} = \inf_{Q\in\mathcal{M}^a(\widetilde{S})}(E_P(v(\widehat{y}\frac{dQ}{dP})) + x\widehat{y}),$$

since $v'(y) = -(u')^{-1}(y)$ and $Q_* \in \mathcal{M}^e(\widetilde{S})$ is a minimum if and only if

$$E_P(v'(y\frac{dQ_*}{dP})(\frac{dQ_*}{dP} - \frac{dQ}{dP})) = 0$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$. This is satisfied by $\widehat{Q}$. By definition of v we obtain

$$\begin{aligned} E_P(v(\widehat{y}\frac{d\widehat{Q}}{dP})) + x\widehat{y} &= \sup_{X \in L^2(\Omega,\mathcal{F},P)} L(X,\widehat{y},\widehat{Q}) \\ &= L(\widehat{X},\widehat{y},\widehat{Q}), \end{aligned}$$

since $u'(\widehat{X}) = \widehat{y}\frac{d\widehat{Q}}{dP}$, $v(y) = u((u')^{-1}(y) - y(u')^{-1}(y)$, so $v(\widehat{y}\frac{d\widehat{Q}}{dP}) = u(\widehat{X}) - \frac{d\widehat{Q}}{dP}\widehat{y}\widehat{X}$. However $L(\widehat{X},\widehat{y},\widehat{Q}) = U(x)$ by assumption on optimality of $\widehat{X}$. Therefore

$$E_P(v(\widehat{y}\frac{d\widehat{Q}}{dP})) + x\widehat{y} = U(x)$$

and $\widehat{y}$ is the minimizer since

$$E_P(v'(\widehat{y}\frac{d\widehat{Q}}{dP})\frac{d\widehat{Q}}{dP}) = -x$$

by assumption. Calculating with the formulas for v yields

$$\begin{aligned} \inf_{\substack{y>0 \\ Q \in \mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP}))) + xy) &= \inf_{y>0}(E_P(v(y\frac{d\widehat{Q}}{dP})) + xy) \\ &= U(x) \\ &= E_P(u(\widehat{X})) \end{aligned}$$

by definition. □

This Theorem enables us to formulate the following duality relation. Given a utility optimization problem for $x \in \operatorname{dom}(u)$

$$\sup_{Y \in \mathcal{K}} E_P(u(x+Y)) = U(x),$$

then we can associate a dual problem, namely

$$\inf_{Q \in \mathcal{M}^a(\widetilde{S})} E_P(v(y\frac{dQ}{dP})) = V(y)$$

for $y > 0$. The main assertion of the minimax considerations is that

$$\inf_{y>0}(V(y) + xy) = U(x),$$

so the concave conjugate of V is U and since V shares the same regularity as U, also U is the convex conjugate of V. First we solve the dual problem (which is much easier) and obtain $y \mapsto \widehat{Q}(y)$. For given $x \in \operatorname{dom}(u)$ we can calculate $\widehat{y}(x)$ and obtain

$$\begin{aligned} V(\widehat{y}(x)) + x\widehat{y}(x) &= U(x) \\ u'(\widehat{X}(x)) &= \widehat{y}(x)\frac{d\widehat{Q}(\widehat{y}(x))}{dP}. \end{aligned}$$

CHAPTER 2

Continuous time models

1. The Bachelier- and Black-Scholes-Model

In this chapter we are going to apply the intuition from discrete models for the pricing and hedging of contingent claims in continuous time models. We are finally going to prove the Black-Scholes formula and some hedging formulas.

The driving engine of many well-known continuous time models is Brownian motion. We shall provide the basic definition of it in dimension 1 and are already able to work out one basic example of continuous time models.

1.1. Definition. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_t)_{t\geq 0}$ a filtration of σ-algebras which satisfies the usual conditions, i.e.*

- *the σ-algebra $\mathcal{F}_t$ contains all P-nullsets.*
- *right continuity holds, $\cap_{t>s}\mathcal{F}_t = \mathcal{F}_s$ for $s \geq 0$.*

Brownian motion then is a stochastic process $(B_t)_{t\geq 0}$ such that

- *B_t is $\mathcal{F}_t$-measurable for $t \geq 0$ (the process is adapted to the filtration).*
- *$B_t - B_s$ is independent of $\mathcal{F}_s$ for $t \geq s \geq 0$.*
- *$B_t - B_s$ is normally distributed $N(0, t-s)$ for $t \geq s \geq 0$.*
- *$B_0 = 0$.*

Furthermore we assume already in the definition that the paths of Brownian motion are continuous, i.e. for all $\omega \in \Omega$ the curve

$$t \mapsto B_t(\omega)$$

is continuous. The same definition can be done on $[0,T]$ and yields a Brownian motion on $[0,T]$.

We can immediately draw some basic conclusions:

1.2. Lemma. *Let $(B_t)_{t\geq 0}$ be a Brownian motion on $(\Omega, \mathcal{F}, P)$, then*

1. *Brownian motion is a martingale, i.e. $E(B_t|\mathcal{F}_s) = B_s$ for $t \geq s$.*
2. *the random variables $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent for $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$ and $n \geq 1$.*

PROOF. We insert directly into the definition. □

We can now give the basic definition of a financial market with finite time horizon $T > 0$, such that second moments exist and interest rates are constant.

1.3. Definition. *Let $(\Omega, \mathcal{F}_T, P)$ be a probability space and $(\mathcal{F}_t)_{0\leq t\leq T}$ a filtration of σ-algebras which satisfies the usual conditions. A financial market is given by a bank account process $S_t^0 = \exp(rt)$, where $r \geq 0$ and $0 \leq t \leq T$ denotes the interest rate and an adapted process $(S_t^1)_{0\leq t\leq T}$ with continuous paths.*

We assume that $S_t^1 \in L^2(\Omega, \mathcal{F}_T, P)$ *and* $S_0 > 0$ *is a constant. A simple portfolio* $(\psi_t, \phi_t)_{0 \le t \le T}$ *is given by stochastic processes* $(\psi_t, \phi_t)_{0 \le t \le T}$ *such that there is* $0 = t_0 < t_1 < t_2 < \cdots < t_n$ *and* $F_i, G_i \in L^\infty(\Omega, \mathcal{F}_{t_i}, P)$ *for* $i = 0, \ldots, n-1$ *such that*

$$\psi_t = \sum_{i=0}^{n-1} G_i 1_{]t_i, t_{i+1}]}(t),$$
$$\phi_t = \sum_{i=0}^{n-1} F_i 1_{]t_i, t_{i+1}]}(t),$$

where $\psi_0 = G_0$ *and* $\phi_0 = F_0$ *by definition. The value process is given by*

$$V_t(\psi, \phi) = \psi_t S_t^0 + \phi_t S_t^1$$

for $0 \le t \le T$. *The discounted value process is given by*

$$\widetilde{V}_t(\psi, \phi) = \psi_t + \phi_t \widetilde{S}_t^1,$$

with $\widetilde{S}_t^1 = \exp(-rt) S_t^1$ *for* $0 \le t \le T$. *A simple portfolio is called self-financing if for* $i = 0, \ldots, n-1$ *we have*

$$\psi_{t_i} S_{t_i}^0 + \phi_{t_i} S_{t_i}^1 = \psi_{t_{i+1}} S_{t_{i+1}}^0 + \phi_{t_{i+1}} S_{t_{i+1}}^1.$$

We denote by $\mathcal{K}$ *the space of all discounted outcomes at initial investment* 0.

As in discrete time we can characterize the discounted outcomes by simple stochastic integrals.

1.4. Lemma. *Given a financial market, then for every self-financing portfolio* $(\psi_t, \phi_t)_{0 \le t \le T}$ *we obtain*

$$\widetilde{V}_t(\psi, \phi) = V_0(\psi, \phi) + \sum_{i=0}^{n-1} \phi_{t_i} (\widetilde{S}_{t_{i+1} \wedge t}^1 - \widetilde{S}_{t_i \wedge t}^1) = V_0(\psi, \phi) + (\phi \cdot \widetilde{S})_t,$$

hence

$$\mathcal{K} = \{(\phi \cdot \widetilde{S})_t \text{ for } \phi \text{ a simple, self-financing trading strategy}\}$$

We shall assume a complete framework for the sequel.

1.5. Condition. *We shall assume that the* L^2*-closure of* $\mathcal{K}$ *can be described by*

$$\overline{\mathcal{K}} = \{X \in L^2(\Omega, \mathcal{F}_T, P) \text{ such that } E_Q(X) = 0\}$$

for some equivalent measure $Q \sim P$. *We call this market complete.*

1.6. Lemma. *Given a complete financial market, the measure* Q *is the unique absolutely continuous martingale measure for the process* $(\widetilde{S}_t^1)_{0 \le t \le T}$. *Furthermore*

$$\overline{\mathcal{K}} \cap L^2_{\ge 0}(\Omega, \mathcal{F}_T, P) = \{0\}.$$

Proof. The proof is very simple. Since $1_A(\widetilde{S}_t^1 - \widetilde{S}_s^1) \in \mathcal{K}$ for $A \in \mathcal{F}_s$ and $t \ge s$, we have that

$$E(\widetilde{S}_t^1 - \widetilde{S}_s^1 | \mathcal{F}_s) = 0$$

for $t \ge s$, which yields the result. For uniqueness we apply the following argument: Given $X \in L^\infty(\Omega, \mathcal{F}_T, P)$, then $X - E_Q(X) \in \overline{\mathcal{K}}$. Given another, absolutely continuous martingale measure Q', we know that

$$E_Q(k \frac{dQ'}{dQ}) = 0$$

for all $k \in \mathcal{K}$. There is a sequence $k_n \in \mathcal{K}$ (which can be chosen uniformly bounded), which converges to $X - E_Q(X)$, hence

$$E_Q(k_n \frac{dQ'}{dQ}) \to E_Q((X - E_Q(X))\frac{dQ'}{dQ})$$

as $n \to \infty$, hence

$$E_Q((X - E_Q(X))\frac{dQ'}{dQ}) = 0$$

for all $X \in L^\infty(\Omega, \mathcal{F}_T, P)$. Consequently $Q' = Q$. For the second assertion we take $Y \in \overline{\mathcal{K}}$ such that $Y \in L^2_{\geq 0}(\Omega, \mathcal{F}_T, P)$, then

$$E_Q(Y) = 0,$$

hence by equivalence $Y = 0$. □

1.7. Remark. We see that the set of martingale measures is in fact determined by $\mathcal{K}$ in our setting. This is a starting point of a general analysis of no-arbitrage and no-free-lunch criteria.

We construct now the first main example of a continuous time model, known as Bachelier model. We assume zero interest rates $r = 0$ (or equally that the discounted price process equals S_t^B). Let $(B_t)_{0 \leq t \leq T}$ be a Brownian motion on $(\Omega, \mathcal{F}_T, P)$ and let $S_0 > 0$ and $\sigma > 0$ be constants, then

$$S_t^B := S_0 + \sigma_B B_t$$

for $0 \leq t \leq T$.

1.8. Theorem. *For the Bachelier model we have $\overline{\mathcal{K}} = \{X \in L^2(\Omega, \mathcal{F}_T, P)$ such that $E_P(X) = 0\}$, so in particular $(S_t^B)_{0 \leq t \leq T}$ is a martingale.*

PROOF. For the proof of this theorem we refer to any text book in stochastic analysis. The theorem is known as martingale representation theorem. □

Given a derivative $Y \in L^2(\Omega, \mathcal{F}_T, P)$, we know from finite dimensional theory that the only arbitrage-free prices are given through

$$E(Y|\mathcal{F}_t) = \pi(Y)_t$$

for $0 \leq t \leq T$. We shall see that in the Bachelier framework this can be easily calculated, which is the "main advantage" of continuous time models!

1.9. Theorem. *Let $S_0, \sigma > 0$ be given, then the price of a European call with strike price $K > 0$ and maturity T at time $t = 0$ is given through*

$$C(S_0, T, K) = (S_0 - K)\Phi(\frac{S_0 - K}{\sigma_B \sqrt{T}}) + \sigma_B \sqrt{T} \phi(\frac{S_0 - K}{S_0 \sigma \sqrt{T}})$$

with

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}),$$
$$\Phi(x) = \int_{-\infty}^{x} \phi(x) dx.$$

PROOF. The proof is a simple integration with respect to normal distribution. We calculate

$$E((S_T - K)_+) = \int_{\frac{K-S_0}{\sigma_B\sqrt{T}}}^{\infty} (S_0 + \sigma_B \sqrt{T}x) - K)\phi(x)dx$$
$$= (S_0 - K)\Phi(\frac{S_0 - K}{\sigma_B\sqrt{T}}) + \sigma_B\sqrt{T}\phi(\frac{S_0 - K}{\sigma_B\sqrt{T}}),$$

which is the result. □

The second important example is the Black-Scholes model. Given $\mu \geq 0$ and $S_0, \sigma > 0$, then

$$S_t^{BS} := S_0 \exp(\mu t - \frac{\sigma^2}{2}t + \sigma B_t)$$

for $0 \leq t \leq T$. The process is adapted and has continuous paths. Furthermore it is a martingale with respect to the following measure.

1.10. Proposition. *Given the Black-Scholes model S^{BS} on $[0, T]$, the measure Q on $(\Omega, \mathcal{F}_T, P)$ by*

$$\frac{dQ}{dP} = \exp(-\frac{\mu}{\sigma}B_T - \frac{\mu^2}{2\sigma^2}T)$$

is an equivalent martingale measure for S^{BS}.

PROOF. We prove first that for $a \in \mathbb{R}$ the stochastic process on $[0, T]$

$$(\exp(-aB_t - \frac{a^2}{2}t))_{0 \leq t \leq T}$$

is a martingale with respect to the filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$. Therefore we show for $t \geq s$

$$E(\exp(-aB_t - \frac{a^2}{2}t)|\mathcal{F}_s) = E(\exp(-a(B_t - B_s) - \frac{a^2}{2}(t-s))\exp(-aB_s - \frac{a^2}{2}s)|\mathcal{F}_s)$$
$$= \exp(-aB_s - \frac{a^2}{2}s)E(\exp(-a(B_t - B_s) - \frac{a^2}{2}(t-s)|\mathcal{F}_s)$$
$$= \exp(-aB_s - \frac{a^2}{2}s)\exp(\frac{a^2(t-s)}{2})\exp(-\frac{a^2}{2}(t-s))$$
$$= \exp(-aB_s - \frac{a^2}{2}s),$$

since $B_t - B_s$ is independent of $\mathcal{F}_s$ and is normally distributed $N(0, 1)$. Next we prove that the process

$$\widetilde{B}_t = B_t + at$$

is Brownian motion with respect to the measure Q_T on $[0, T]$. This means that we have to check all properties of arbitrage measures for the process $(\widetilde{B}_t)_{0 \leq t \leq T}$. Continuity of paths is clear, also adaptedness, furthermore $\widetilde{B}_0 = 0$, consequently we have to check independence and the Gaussian property. Therefore we show

$$E_{Q_T}(\exp(v(\widetilde{B}_t - \widetilde{B}_s))|\mathcal{F}_s) = \exp(\frac{\nu^2(t-s)}{2})$$

for complex ν. We apply formula 4.7 and obtain

$$\begin{aligned}
E_{Q_T}(\exp(v(\widetilde{B}_t - \widetilde{B}_s))|\mathcal{F}_s) &= \frac{1}{X_s} E_P(\exp(v(\widetilde{B}_t - \widetilde{B}_s))X_T|\mathcal{F}_s) \\
&= \frac{1}{X_s} E_P(\exp(v(\widetilde{B}_t - \widetilde{B}_s))E(X_T|\mathcal{F}_t)|\mathcal{F}_s) \\
&= \frac{1}{X_s} E_P(\exp(v(\widetilde{B}_t - \widetilde{B}_s))X_t|\mathcal{F}_s) \\
&= \exp(aB_s + \frac{a^2}{2}s)\times \\
&\times E(\exp(v(B_t - B_s) + va(t-s)) \exp(-aB_t - \frac{a^2}{2}t)|\mathcal{F}_s) \\
&= E(\exp((v-a)(B_t - B_s) + (va - \frac{a^2}{2})(t-s))|\mathcal{F}_s) \\
&= \exp(\frac{(v-a)^2}{2}(t-s) + (va - \frac{a^2}{2})(t-s)) \\
&= \exp(\frac{\nu^2(t-s)}{2})
\end{aligned}$$

for $t \geq s$ with the martingale

$$X_s = \exp(-aB_s - \frac{a^2}{2}s) = E(\exp(-aB_T - \frac{a^2}{2}T)|\mathcal{F}_s)$$

for $T \geq s \geq 0$.Hence we know for $a = \frac{\mu}{\sigma}$ we can write equivalently

$$S_t^{BS} = S_0 \exp(\sigma \widetilde{B}_t - \frac{\sigma^2}{2}t)$$

for $0 \leq t \leq T$ and therefore by the previous results, the stochastic process $(S_t)_{0 \leq t \leq T}$ is a Q_T-martingale. □

1.11. Theorem. *For the Black-Scholes model we have* $\overline{\mathcal{K}} = \{X \in L^2(\Omega, \mathcal{F}_T, P)$ *such that* $E_{Q_T}(X) = 0\}$*, so in particular* $(S_t^{BS})_{0 \leq t \leq T}$ *is a* Q_T*-martingale.*

PROOF. Again we refer to any textbook in stochastic analysis. □

Finally we can prove the Black-Scholes formula, which is pricing with respect to the unique equivalent martingale measure Q_T.

1.12. Theorem. *Given the Black-Scholes model* $(S_t^{BS})_{0 \leq t \leq T}$*, a maturity time* $T_0 \leq T$ *and a strike price* $K \geq 0$*, the unique price of the European call* $(S_{T_0} - K)_+$ *without interest rates is given through*

$$C(S_0, K, T_0) = S_0 \Phi(\frac{\ln \frac{S_0}{K} + \frac{1}{2}\sigma^2 T_0}{\sigma\sqrt{T_0}}) - K\Phi(\frac{\ln \frac{S_0}{K} - \frac{1}{2}\sigma^2 T_0}{\sigma\sqrt{T_0}}).$$

The price with interest rates r *is given through*

$$C(S_0, K, T_0, r) = S_0 \Phi(\frac{\ln \frac{S_0}{K} + (\frac{1}{2}\sigma^2 + r)T_0}{\sigma\sqrt{T_0}}) - e^{-rT_0} K\Phi(\frac{\ln \frac{S_0}{K} - (\frac{1}{2}\sigma^2 - r)T_0}{\sigma\sqrt{T_0}}).$$

PROOF. We have to calculate for $r = 0$ the following integral

$$\begin{aligned}
E_{Q_T}((S_{T_0} - K)_+) &= E_{Q_T}((S_0 \exp(\sigma \widetilde{B}_{T_0} - \frac{\sigma^2}{2} T_0) - K)_+) \\
&= \int_{-\infty}^{\infty} (S_0 \exp(\sigma \sqrt{T_0} x - \frac{\sigma^2}{2} T_0) - K)_+ \phi(x) dx \\
&= \int_{\frac{\ln \frac{K}{S_0} + \frac{\sigma^2}{2} T_0}{\sigma \sqrt{T_0}}}^{\infty} (S_0 \exp(\sigma \sqrt{T_0} x - \frac{\sigma^2}{2} T_0) - K) \phi(x) dx \\
&= S_0 \int_{\frac{\ln \frac{K}{S_0} + \frac{\sigma^2}{2} T_0}{\sigma \sqrt{T_0}}}^{\infty} \exp(\sigma \sqrt{T_0} x - \frac{\sigma^2}{2} T_0) \phi(x) dx - \\
&- K \int_{\frac{\ln \frac{K}{S_0} + \frac{\sigma^2}{2} T_0}{\sigma \sqrt{T_0}}}^{\infty} \phi(x) dx \\
&= S_0 \Phi(\frac{\ln \frac{S_0}{K} + \frac{1}{2} \sigma^2 T_0}{\sigma \sqrt{T_0}}) - K \Phi(\frac{\ln \frac{S_0}{K} - \frac{1}{2} \sigma^2 T_0}{\sigma \sqrt{T_0}}).
\end{aligned}$$

If interest rates are not 0 we have to calculate

$$E_{Q_T}(e^{-rT_0}(S_{T_0} - K)_+) = E_{Q_T}((e^{-rT_0} S_{T_0} - e^{-rT_0} K)_+),$$

where $e^{-rT} S_T$ is a martingale with respect to Q_T, hence replacing K by $e^{-rT_0} K$ leads to the Black-Scholes formula. □

2. Hedging

We finally address the question of hedging in the Black-Scholes and Bachelier model. Without going into detail related to stochastic analysis (where we refer to finance related text books), we state the main properties of the hedging portfolio. We state the main theorems for the application of the Ito formula.

2.1. Theorem. *Let $t \geq 0$ be a fixed point in time and $(B_s)_{s \geq 0}$ a Brownian motion, then*

$$\lim_{n \to \infty} \sum_{i=0}^{2^n - 1} (B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^2 = t$$

almost surely.

PROOF. We define for $n \geq 1$

$$S_n = \sum_{i=0}^{2^n - 1} (B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^2$$

and can immediately prove by the basic properties of Brownian motion (covariance and independence) that

$$\begin{aligned}
E(S_n) &= t \\
E(S_n^2) &= \sum_{i=0}^{2^n-1} E((B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^4) + \sum_{\substack{i,j=0\\ i\neq j}}^{2^n-1} E((B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^2(B_{\frac{j(i+1)}{2^n}} - B_{\frac{ji}{2^n}})^2) \\
&= t^2(3\frac{2^n}{2^{2n}} + \frac{2^n-1}{2^n}) \\
&= t^2(\frac{2}{2^n} + 1)
\end{aligned}$$

for $n \geq 1$. Therefore we can conclude

$$\begin{aligned}
E((S_n - t)^2) &= t^2(\frac{2}{2^n} + 1 - 2 + 1) \\
&= \frac{t^2}{2^{n-1}}
\end{aligned}$$

for $n \geq 1$. By Chebyshev's inequality we obtain finally

$$P((S_n - t)^2 \geq \frac{1}{2^{\frac{n}{2}}}) \leq 2^{\frac{n}{2}} \frac{t^2}{2^{n-1}} = \frac{1}{2^{\frac{n}{2}}} 2t^2,$$

which leads by the Borel-Cantelli Lemma to the assertion that the set of ω with $(S_n - t)^2 \geq \frac{1}{2^{\frac{n}{2}}}$ for infinitely many $n \geq 1$ is of measure 0. Hence on a set of measure 1 we have

$$\lim_{n\to\infty} S_n = t,$$

which is the desired assertion. □

Now we turn to the construction of the Ito-integral. Given a standard Brownian motion $(B_t)_{t\geq 0}$ on $\mathbb{R}^d$. We denote by $L^2(\mathbb{R}_{\geq 0} \times \Omega, \mathcal{F}_p, dt \otimes P)$ the set of all progressively measurable processes, i.e the set of

$$\phi : \mathbb{R}_{\geq 0} \times \Omega \to \mathbb{R},$$

which are measurable with respect to the σ-algebra $\mathcal{F}_p$, i.e. the σ-algebra generated by $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$ for $t \geq 0$ and square-integrable thereon. These are all maps such that the restriction $\phi 1_{[0,t]}$ lies in $L^2([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t, dt \otimes P)$ and

$$E(\int_0^\infty \phi(s)^2) = \int_\Omega \int_0^\infty \phi(s,\omega)^2 ds P(d\omega) < \infty.$$

The subspace of simple predictable processes, i.e.

$$u(t) = \sum_{i=0}^{n-1} F_i 1_{]t_i,t_{i+1}]}(t)$$

with F_i a $\mathcal{F}_{t_i}$-measurable and $E(F_i^2) < \infty$ (hence $F_i \in L^2(\Omega, \mathcal{F}_{t_i}, P)$, $n \geq 0$ and $0 = t_0 \leq t_1 \leq ... \leq t_n$, is denoted by $\mathcal{E}$. On $\mathcal{E}$ we define the Ito-integral by

$$I(u) = \int_0^\infty u(t) dB_t := \sum_{i=0}^{n-1} F_i(B_{t_{i+1}} - B_{t_i})$$

2.2. Theorem. *The mapping* $I : \mathcal{E} \to L^2(\Omega, \mathcal{F}, P)$ *is a well defined isometry and* $E(I(u)) = 0$ *for all* $u \in \mathcal{E}$, *i.e.*

$$E(I(u)I(v)) = E(\int_0^\infty u(t)v(t)dt).$$

Proof. The proof follows from covariance properties of Brownian motion. For the first property we simply observe that

$$\begin{aligned} E(I(u)) &= E(\sum_{i=0}^{n-1} F_i(B_{t_{i+1}} - B_{t_i})) \\ &= \sum_{i=0}^{n-1} E(F_i E((B_{t_{i+1}} - B_{t_i})|\mathcal{F}_{t_i})) \\ &= 0 \end{aligned}$$

for all $u \in \mathcal{E}$. For the second property we observe – due to bilinearity - that it is sufficient to show $E(I(u)^2) = E(\int_0^\infty u(t)^2 dt)$ for all $u \in \mathcal{E}$. Again we observe directly

$$\begin{aligned} E(I(u)^2) &= E(\sum_{i,j=0}^{n-1} F_i F_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})) \\ &= E(\sum_{i=0}^{n-1} F_i^2 (B_{t_{i+1}} - B_{t_i})^2) + 2E(\sum_{i<j=0}^{n-1} F_i F_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})) \\ &= \sum_{i=0}^{n-1} E(F_i^2 E((B_{t_{i+1}} - B_{t_i})^2|\mathcal{F}_{t_i}) + \\ &+ 2\sum_{i<j=0}^{n-1} E(F_i F_j (B_{t_{i+1}} - B_{t_i}) E((B_{t_{j+1}} - B_{t_j})|\mathcal{F}_{t_j})) \\ &= \sum_{i=0}^{n-1} E(F_i^2)(t_{i+1} - t_i) \\ &= E(\int_0^\infty u(t)^2 dt) \end{aligned}$$

for $u \in \mathcal{E}$. □

2.3. Definition. *The closure of* $\mathcal{E}$ *in* $L^2(\mathbb{R}_{\geq 0} \times \Omega, \mathcal{F}_p, dt \otimes P)$ *is denoted by* $L^2(B)$. *The unique continuous extension* $I : L^2(B) \to L^2(\Omega)$ *is called the stochastic integral with respect to Brownian motion or the Ito integral, we denote*

$$\int_0^\infty u(t)dB_t := I(u).$$

In particular we have for all $u, v \in L^2(B)$

$$E(\int_0^\infty u(t)dB_t) = 0$$

$$E(\int_0^\infty u(t)dB_t \int_0^\infty v(t)dB_t) = E(\int_0^\infty u(t)v(t)dt)$$

The definite integral is defined in the following way

$$\int_0^t u(s)dB_s := \int_0^t u(s)1_{[0,t]}(s)dB_s$$

for $t \geq 0$, which is well defined since the processes u are progressively measurable.

2.4. Theorem. *The vector space $\mathcal{E}$ is dense in $L^2(\mathbb{R}_{\geq 0} \times \Omega, \mathcal{F}_p, dt \otimes P)$.*

PROOF. see any book on stochastic analysis. □

2.5. Corollary. *The process $M_t := \int_0^t u(s)dB_s$ has a version with continuous paths.*

2.6. Remark. All simple processes $u \in \mathcal{E}$ are progressively measurable by definition. Given $u \in L^2(\mathbb{R}_{\geq 0} \times \Omega, \mathcal{F}_p, dt \otimes P)$ with continuous paths, then we can approximate the process $(u_s 1_{[0,t]})_{s\geq 0}$ by elements in $\mathcal{E}$ of the form

$$u_s^n := \sum_{i=0}^{2^n-1} u_{\frac{ti}{2^n}} 1_{]\frac{ti}{2^n}, \frac{t(i+1)}{2^n}]}(s),$$

which have the property

$$u_s^n \to u_s$$

surely by continuity and $u^n \to u1_{[0,t]}$ in $L^2(\mathbb{R}_{\geq 0} \times \Omega, \mathcal{F}_p, dt \otimes P)$ by dominated convergence. Therefore we can calculate the Ito integral for processes with continuous paths by

$$\lim_{n\to\infty} \int_0^t u_s^n dB_s = \lim_{n\to\infty} \sum_{i=0}^{2^n-1} u_{\frac{ti}{2^n}} (B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}}).$$

As limit we shall always take a version with continuous paths. We could equally take any other refining sequence of partitions instead of the dyadic one.

2.7. Exercise. As an easy exercise one can prove

$$\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t).$$

We simply take the limit of

$$\sum_{i=0}^{2^n-1} B_{\frac{ti}{2^n}} (B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}}) = \frac{1}{2} \sum_{i=0}^{2^n-1} (B^2_{\frac{t(i+1)}{2^n}} - B^2_{\frac{ti}{2^n}}) - \frac{1}{2} \sum_{i=0}^{2^n-1} (B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^2$$

$$= \frac{B_t^2}{2} - \frac{1}{2} \sum_{i=0}^{2^n-1} (B_{\frac{t(i+1)}{2^n}} - B_{\frac{ti}{2^n}})^2$$

applying the result on almost sure convergence of the quadratic variation.

2.8. Theorem. *Let $f \in C_b^2([0,T] \times \mathbb{R}, \mathbb{R})$ (bounded with bounded derivatives) be given. Suppose $u, v \in L^2(\mathbb{R}_{\geq 0} \times \Omega, \mathcal{F}_p, dt \otimes P)$. Let X be a the continuous process*

$$X_t := X_0 + \int_0^t u(s)dB_s + \int_0^t v(s)ds,$$

then

$$f(t,X_t) = f(0,X_0) + \int_0^t \frac{\partial}{\partial s} f(s,X_s)ds + \int_0^t \frac{\partial}{\partial x} f(s,X_s)u(s)dB_s + \\ + \int_0^t \frac{\partial}{\partial x} f(s,X_s)v(s)ds + \frac{1}{2}\int_0^t \frac{\partial^2}{\partial x^2} f(s,X_s)u^2(s)ds.$$

In short notation this is written as,

$$dX_t := u(t)dB_t + v(t)dt,$$
$$df(t,X_t) = \frac{\partial}{\partial t} f(t,X_t)dt + \frac{\partial}{\partial x} f(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2}{\partial x^2} f(t,X_t)d,$$

where the process is given through $d\langle X\rangle_t := u^2(t)dt$.

2.1. Bachelier Hedging. In order to come up with a Hedging formula we need to redefine our model. From now on we call – given a Brownian motion $(B_t)_{0\leq t\leq T}$ – the (discounted) price process

$$S_t = S_0 + \sigma_B B_t,$$
$$dS_t = \sigma_B dB_t$$

for $0 \leq t \leq T$, where we call σ_B the absolute Bachelier volatility. We can calculate – by the previous methods – the price of a European Call Option in this model

$$\begin{aligned} C^B(S_0,T) &:= E((S_T - K)_+) \\ &= \int_{\frac{K-S_0}{\sigma_B\sqrt{T}}}^{\infty} (S_0 + \sigma_B\sqrt{T}x - K)\phi(x)dx \\ &= (S_0 - K)\Phi(\frac{S_0 - K}{\sigma_B\sqrt{T}}) + \sigma_B\sqrt{T}\phi(\frac{S_0 - K}{\sigma_B\sqrt{T}}). \end{aligned}$$

By simple differentiation we check that

$$\frac{\partial}{\partial T} C^B(S_0,T) = \frac{(\sigma_B)^2}{2}\frac{\partial^2}{\partial S_0^2} C^B(S_0,T)$$

for $T > 0$ and $S_0 \in \mathbb{R}$. Ito's Formula for the stochastic process $(C^B(S_t, T-t))_{0\leq t\leq T}$ then yields the following result:

$$\begin{aligned} C^B(S_T,0) &= C^B(S_0,T) - \int_0^T \frac{\partial}{\partial T} C^B(S_t,T-t)dt + \\ &+ \int_0^T \frac{\partial}{\partial S_0} C^B(S_t,T-t)dS_t + \\ &+ \frac{1}{2}\int_0^T \frac{(\sigma_B)^2}{2}\frac{\partial^2}{\partial S_0^2} C^B(S_t,T-t)dt \\ &= C^B(S_0,T) + \int_0^T \frac{\partial}{\partial S_0} C^B(S_t,T-t)dS_t. \end{aligned}$$

Consequently we can build a self-financing portfolio at initial wealth $C^B(S_0,T)$, which replicates the European Call.

2.2. Black-Scholes Hedging. We take a Black-Scholes model with volatility $\sigma > 0$, drift μ and today's price S_0,

$$S_t = S_0 \exp(\mu t - \frac{\sigma^2}{2} t + \sigma B_t)$$

for $0 \leq t \leq T$. Furthermore we assume an interest rate $r \geq 0$, we obtain the discounted price process

$$\begin{aligned} \widetilde{S}_t &= S_0 \exp(\mu t - rt - \frac{\sigma^2}{2} t + \sigma B_t), \\ d\widetilde{S}_t &= \widetilde{S}_t(\mu - r)dt + \widetilde{S}_t \sigma dB_t. \end{aligned}$$

We calculate – like in the Bachelier model – the price of a European Call Option, hence

$$C(S_0, T, r) = S_0 \Phi(\frac{\ln \frac{S_0}{K} + (\frac{1}{2}\sigma^2 + r)T}{\sigma\sqrt{T}}) - e^{-rT_0} K \Phi(\frac{\ln \frac{S_0}{K} - (\frac{1}{2}\sigma^2 - r)T}{\sigma\sqrt{T}}).$$

As before we see that

$$\frac{\partial}{\partial T} C(S_0, T, r) = \frac{\sigma^2 S_0^2}{2} \frac{\partial^2}{\partial S_0^2} C^B(S_0, T, r)$$

In order to calculate the Hedging Portfolio, we apply Ito's Formula to the process $(C(\widetilde{S}_t, T-t))_{0 \leq t \leq T}$,

$$\begin{aligned} C(S_T, 0, r) &= C(S_0, T, r) - \int_0^T \frac{\partial}{\partial T} C(\widetilde{S}_t, T-t, r) dt + \\ &+ \int_0^T \frac{\partial}{\partial S_0} C(\widetilde{S}_t, T-t, r) d\widetilde{S}_t + \\ &+ \frac{1}{2} \int_0^T \frac{\sigma^2 S_t^2}{2} \frac{\partial^2}{\partial S_0^2} C(\widetilde{S}_t, T-t, \widetilde{S}_t) dt \\ &= C^B(S_0, T) + \int_0^T \frac{\partial}{\partial S_0} C(\widetilde{S}_t, T-t) d\widetilde{S}_t. \end{aligned}$$

CHAPTER 3

Mathematical Preliminaries

1. Methods from convex analysis

In this chapter basic duality methods from convex analysis are discussed. We shall also apply the notions of dual normed vector spaces in finite dimensions. Let V be a real vector space with norm and real dimension $\dim V < \infty$, then we can define the *pairing*

$$\langle .,. \rangle : V \times V' \to \mathbb{R}$$
$$(v,l) \mapsto l(v)$$

where V' denotes the dual vector space, i.e. the space of continuous linear functionals $l : V \to \mathbb{R}$. The dual space carries a natural dual norm namely

$$||l|| := \sup_{||v|| \leq 1} |l(v)|.$$

We obtain the following duality relations:

- If for some $v \in V$ it holds that $\langle v,l \rangle = 0$ for all $l \in V'$, then $v = 0$.
- If for some $l \in V'$ it holds that $\langle v,l \rangle = 0$ for all $v \in V$, then $l = 0$.
- There is a natural isomorphism $V \to V''$ and the norms on V and V'' coincide (with respect to the previous definition).

If V is an euclidean vector space, i.e. there is a scalar product $\langle .,. \rangle : V \times V \to \mathbb{R}$, which is symmetric and positive definite, then we can identify V' with V and every linear functional $l \in V'$ can be uniquely represented $l = \langle ., x \rangle$ for some $x \in V$.

1.1. Definition. *Let V be a finite dimensional vector space. A subset $C \subset V$ is called* convex *if for all $v_1, v_2 \in C$ also $tv_1 + (1-t)v_2 \in C$ for $t \in [0,1]$.*

Since the intersection of convex sets is convex, we can define the convex hull of any subset $M \subset V$, which is denoted by $\langle M \rangle_{conv}$. We also define the closed convex hull $\overline{\langle M \rangle_{conv}}$, which is the smallest closed, convex subset of V containing M. If M is compact the convex hull $\langle M \rangle_{conv}$ is already closed and therefore compact.

1.2. Definition. *Let C be a closed convex set, then $x \in C$ is called* extreme point *of C if for all $y, z \in C$ with $x = ty + (1-t)z$ and $t \in [0,1]$, we have either $t = 0$ or $t = 1$. This is equivalent to saying that there are no two different points $x_1 \neq x_2$ such that $x = \frac{1}{2}(x_1 + x_2)$.*

First we treat a separation theorem, which is valid in a fairly general context and known as Hahn-Banach Theorem.

1.3. Theorem. *Let C be a closed convex set in an euclidean vector space V, which does not contain the origin, i.e. $0 \notin C$. Then there exists a linear functional $\xi \in V'$ and $\alpha > 0$ such that for all $x \in C$ we have $\xi(x) \geq \alpha$.*

Proof. Let r be a radius such that the closed ball $B(r)$ intersects C. The continuous map $x \mapsto ||x||$ achieves a minimum $x_0 \neq 0$ on $B(r) \cap C$, which we denote by x_0, since $B(r) \cap C$ is compact. We certainly have for all $x \in C$ the relation $||x|| \geq ||x_0||$. By convexity we obtain that $x_0 + t(x - x_0) \in C$ for $t \in [0,1]$ and hence

$$||x_0 + t(x - x_0)||^2 \geq ||x_0||^2.$$

This equation can be expanded for $t \in [0,1]$,

$$\begin{aligned} ||x_0||^2 + 2t\,\langle x_0, x - x_0\rangle + t^2||(x - x_0)||^2 &\geq ||x_0||^2, \\ 2t\,\langle x_0, x - x_0\rangle + t^2||(x - x_0)||^2 &\geq 0. \end{aligned}$$

Take now small t and assume $\langle x_0, x - x_0\rangle < 0$ for some $x \in C$, then there appears a contradiction in the previous inequality, hence we obtain

$$\langle x_0, x - x_0\rangle \geq 0$$

and consequently $\langle x, x_0\rangle \geq ||x_0||^2$ for $x \in C$, so we can choose $\xi = \langle ., x_0\rangle$. □

As a corollary we have that each subspace $V_1 \subset V$, which does not intersect with a convex, compact and non-empty subset $K \subset V$ can be separated from K, i.e. there is $\xi \in V'$ such that $\xi(V_1) = 0$ and $\xi(x) > 0$ for $x \in K$. This is proved by considering the set

$$C := K - V := \{w - v \text{ for } v \in V \text{ and } w \in K\},$$

which is convex and closed, since V, K are convex and K is compact, and which does not contain the origin. By the above theorem we can find a separating linear functional $\xi \in V'$ such that $\xi(w - v) \geq \alpha$ for all $w \in K$ and $v \in V$, which means in particular that $\xi(w) > 0$ for all $w \in K$. Furthermore we obtain from $\xi(w) - \xi(v) \geq \alpha$ for all $v \in V$ that $\xi(v) = 0$ for all $v \in V$ (replace v by λv, which is possible since V is a vector space, and lead the assertion to a contradiction in case that $\xi(v) \neq 0$).

1.4. Theorem. *Let C be a compact convex non-empty set, then C is the convex hull of all its extreme points.*

Proof. We have to show that there is an extreme point. We take a point $x \in C$ such that the distance $||x||^2$ is maximal, then x is an extreme point. Assume that there are two different points x_1, x_2 such that $x = \frac{1}{2}(x_1 + x_2)$, then

$$\begin{aligned} ||x||^2 = ||\frac{1}{2}(x_1 + x_2)||^2 &< \frac{1}{2}(||x_1||^2 + ||x_2||^2) \\ &\leq \frac{1}{2}(||x||^2 + ||x||^2) = ||x||^2, \end{aligned}$$

by the parallelogram law $\frac{1}{2}(||y||^2 + ||z||^2) = ||\frac{1}{2}(y+z)||^2 + ||\frac{1}{2}(y-z)||^2$ for all $y, z \in V$ and the maximality of $||x||^2$. This is a contradiction. Therefore we obtain at least one extreme point.

The set of all extreme points is a compact set, since it lies in C and is closed. Take now the convex hull of all extreme points, which is a closed convex subset S of C and hence compact. If there is $x \in C \setminus S$, then we can separate by a hyperplane l the point x and S such that $l(x) \geq \alpha > l(y)$ for $y \in S$. The set $\{l \geq \alpha\} \cap C$ is compact, convex, nonempty and has therefore an extreme point z, which is also an extreme point of C. So $z \in S$, which is a contradiction. □

Next we treat basic duality theory in the finite dimensional vector space V with euclidean structure. We identify the dual space V' with V by the above representation.

1.5. Definition. *A subset $C \subset V$ is called convex cone if for all $v_1, v_2 \in C$ the sum $v_1 + v_2 \in C$ and $\lambda v_1 \in C$ for $\lambda \geq 0$. Given a cone C we define the polar C^0*

$$C^0 := \{l \in V \text{ such that } \langle l, v\rangle \leq 0 \text{ for all } v \in C\}.$$

The intersection of convex cones is a convex cone and therefore we can speak of the smallest convex cone containing an arbitrary set $M \subset V$, which is denoted by $\langle M\rangle_{cone}$. We want to prove the bipolar theorem for convex cones.

1.6. Theorem (Bipolar Theorem)**.** *Let $C \subset V$ be a convex cone, then $C^{00} \subset V$ is the closure of C.*

Proof. We show both inclusions. Take $v \in \overline{C}$, then $\langle l, v\rangle \leq 0$ for all $l \in C^0$ by definition of C^0 and therefore $v \in C^{00}$. If there were $v \in C^{00} \setminus \overline{C}$, where $\overline{C}$ denotes the closure of C, then for all $l \in C^0$ we have that $\langle l, v\rangle \leq 0$ by definition. On the other hand we can find $l \in V$ such that $\langle l, \overline{C}\rangle \leq 0$ and $\langle l, v\rangle > 0$ by the separation theorem since $\overline{C}$ is a closed cone. Take therefore l and α such that $\langle l, \overline{C}\rangle \leq \alpha$ and $\langle l, v\rangle > \alpha$. Since $0 \in \overline{C}$ we get $\alpha \geq 0$ and if there were $x \in \overline{C}$ with $\langle l, x\rangle > 0$, then for all $\lambda \geq 0$ we have $\langle l, \lambda x\rangle = \lambda \langle l, x\rangle \leq \alpha$, which is a contradiction, so $\langle l, x\rangle \leq 0$. By assumption we have $l \in C^0$, however this yields a contradiction since $\langle l, v\rangle > 0$ and $v \in C^{00}$. □

1.7. Definition. *A convex cone C is called* polyhedral *if there is a finite number of linear functionals $l_1, \ldots, l_m$ such that*

$$C := \cap_{i=1}^{n} \{v \in V | \langle l_i, v\rangle \leq 0\}.$$

In particular a polyhedral cone is closed as intersection of closed sets.

1.8. Lemma. *Given $e_1, \ldots, e_n \in V$. For the cone $C = \langle e_1, \ldots, e_n\rangle_{con}$ the polar can be calculated as*

$$C^0 = \{l \in V \text{ such that } \langle l, e_i\rangle \leq 0 \text{ for all } i = 1, \ldots, n\}.$$

Proof. The convex cone $C = \langle e_1, \ldots, e_n\rangle_{cone}$ is given by

$$C = \{\sum_{i=1}^{n} \alpha_i e_i \text{ for } \alpha_i \geq 0 \text{ and } i = 1, \ldots, n\}.$$

Given $l \in C^0$, the equation $\langle l, e_i\rangle \leq 0$ necessarily holds and we have the inclusion $\subset$. Given $l \in V$ such that $\langle l, e_i\rangle \leq 0$ for $i = 1, \ldots, n$, then for $\alpha_i \geq 0$ the equation $\sum_{i=1}^{n} \alpha_i \langle l, e_i\rangle \leq 0$ holds and therefore $l \in C^0$ by the explicit description of C as $\sum_{i=1}^{n} \alpha_i e_i$ for $\alpha_i \geq 0$. □

1.9. Corollary. *Given $e_1, \ldots, e_n \in V$, the cone $C = \langle e_1, \ldots, e_n\rangle_{con}$ has a polar which is polyhedral and therefore closed.*

Proof. The polyhedral cone is given through

$$\begin{aligned} C^0 &= \{l \in V \text{such that } \langle l, e_i\rangle \leq 0 \text{ for all } i = 1, \ldots, n\} \\ &= \cap_{i=1}^{n} \{l \in V | \langle l, e_i\rangle \leq 0\}. \end{aligned}$$

□

1.10. Lemma. *Given a finite set of vectors $e_1, \dots, e_n \in V$ and the convex cone $C = \langle e_1, \dots, e_n \rangle_{con}$, then C is closed.*

Proof. Assume that $C = \langle e_1, \dots, e_n \rangle_{con}$ for vectors $e_i \in V$. If the e_i are linearly independent, then C is closed by the argument, that any $x \in C$ can be uniquely written as $x = \sum_{i=1}^n \alpha_i e_i$. Suppose next that there is a non-trivial linear combination $\sum_{i=1}^n \beta_i e_i = 0$ with $\beta \in \mathbb{R}^n$ non-zero and some $\beta_i < 0$. We can write $x \in C$ as

$$x = \sum_{i=1}^n \alpha_i e_i = \sum_{i=1}^n (\alpha_i + t(x)\beta_i) e_i = \sum_{j \neq i(x)} \alpha'_i e_i$$

with

$$i(x) \in \{i \text{ such that } |\frac{\alpha_i}{\beta_i}| = \max_{\beta_j < 0} |\frac{\alpha_j}{\beta_j}|\},$$
$$t(x) = -\frac{\alpha_{i(x)}}{\beta_{i(x)}}$$

Then $\alpha'_j \geq 0$ by definition. Consequently we can construct by variation of x a decomposition

$$C = \cup_{i=1}^{n'} C_i$$

where C_i are cones generated by $n-1$ vectors from the set $e_1, \dots, e_n$. By induction on the number of generators n we can conclude, since the cone generated by one element e_1 is obviously closed. □

1.11. Proposition. *Let $C \subset V$ be a convex cone generated by $e_1, \dots, e_n$ and $\mathcal{K}$ a subspace, then $\mathcal{K} - C$ is closed convex.*

Proof. First we prove that $\mathcal{K} - C$ is a convex cone. Taking $v_1, v_2 \in \mathcal{K} - C$, then $v_1 = k_1 - c_1$ and $v_2 = k_2 - c_2$, therefore

$$v_1 + v_2 = k_1 + k_2 - (c_1 + c_2) \in \mathcal{K} - C,$$
$$\lambda v_1 = \lambda k_1 - \lambda c_1 \in \mathcal{K} - C.$$

In particular $0 \in \mathcal{K} - C$. The convex cone is generated by a generating set $e_1, \dots, e_n$ for C and a basis $f_1, \dots, f_p$ for $\mathcal{K}$, which has to be taken with $-$ sign, too. So

$$\mathcal{K} - C = \langle -e_1, \dots, -e_n, f_1, \dots, f_p, -f_1, \dots, -f_p \rangle_{con}$$

and therefore $\mathcal{K} - C$ is closed by Lemma 1.10. □

1.12. Theorem (Farkas Lemma)**.** *Let $e_1, \dots, e_n \in V$ be given, then the cone $C = \langle e_1, \dots, e_n \rangle_{con} = C^{00}$. Another formulation is that $b \in C$ if and only if $\langle b, x \rangle \leq 0$ for all $x \in C^0$ (which means $b \in C^{00}$).*

Proof. The cone C is closed and therefore $C = C^{00}$ by the Bipolar Theorem 1.6. □

1.13. Lemma. *Let C be a polyhedral cone, then there are finitely many vectors $e_1, \dots, e_n \in V$ such that*

$$C = \langle e_1, \dots, e_n \rangle_{con} .$$

PROOF. By assumption $C = \cap_{i=1}^{p}\{v \in V \,|\, \langle l_i, v\rangle \le 0\}$ for some vectors $l_i \in V$. We intersect C with $[-1,1]^m$ and obtain a convex, compact set. This set is generated by its extreme points. We have to show that there are only finitely many extreme points. Assume that there are infinitely many extreme points, then there is also an adherence point $x \in C$. Take a sequence of extreme points $(x_n)_{n\ge 0}$ such that $x_n \to x$ as $n \to \infty$ with $x_n \neq x$. We can write the defining inequalities for $C \cap [-1,1]^m$ by

$$\langle k_j, v\rangle \le a_j$$

for $j = 1, \dots, r$ and we obtain $\lim_{n\to\infty} \langle k_j, x_n\rangle = \langle k_j, x\rangle$. Define

$$\epsilon := \min_{\langle k_j, x\rangle < a_j} a_j - \langle k_j, x\rangle > 0.$$

Take n_0 large enough such that $|\langle k_j, x_{n_0}\rangle - \langle k_j, x\rangle| \le \frac{\epsilon}{2}$, which is possible due to convergence. Then we can look at $x_{n_0} + t(x - x_{n_0}) \in C$ for $t \in [0,1]$. We want to find a continuation of this segment for some $\delta > 0$ such that $x_{n_0} + t(x - x_{n_0}) \in C$ for $[-\delta, 1]$. Therefore we have to check three cases:

- If $\langle k_j, x_{n_0}\rangle = \langle k_j, x\rangle = a_j$, then we can continue for all $t \le 0$ and the inequality $\langle k_j, x_{n_0} + t(x - x_{n_0})\rangle = a_j$ remains valid.
- If $\langle k_j, x\rangle = a_j$ and $\langle k_j, x_{n_0}\rangle < a_j$, we can continue for all $t \le 0$ and the inequality $\langle k_j, x_{n_0} + t(x - x_{n_0})\rangle \le a_j$ remains valid.
- If $\langle k_j, x\rangle < a_j$, then we define $\delta = 1$ and obtain that for $-1 \le t \le 1$ the inequality $\langle k_j, x_{n_0} + t(x - x_{n_0})\rangle \le a_j$ remains valid.

Therefore we can find δ and continue the segment for small times. Hence x_n cannot be an extreme point, since it is a nontrivial convex combination of $x_{n_0} - \delta(x - x_{n_0})$ and x, which is a contradiction. Therefore $C \cap [-1,1]^m$ is generated by finitely many extreme points $e_1, \dots, e_n$and so

$$C = \langle e_1, \dots, e_n\rangle_{con}$$

by dilatation. □

2. Optimization Theory

We shall first consider general principles in optimization theory related to analysis and proceed to special functionals.

2.1. Definition. *Let $U \subset \mathbb{R}^m$ be a subset with $U \subset V$, where V is open in $\mathbb{R}^m$. Let $F : V \to \mathbb{R}$ be a C^2-function. A point $x \in U$ is called local maximum (local minimum) of F on U if there is a neighborhood W_x of x in V such that for all $y \in U \cap W_x$*

$$F(y) \le F(x)$$

or respectively $F(y) \ge F(x)$.

2.2. Lemma. *Let $U \subset \mathbb{R}^m$ be a subset with $U \subset V$, where V is open in $\mathbb{R}^m$ and let $F : V \to \mathbb{R}$ be a C^2-function. Given a local maximum (or local minimum) $x \in U$ of F on U and a C^2-curve $c :]-1,1[\to V$ such that $c(0) = x$ and $c(t) \in U$ for $t \in]-1,1[$, the following necessary condition holds true,*

$$\frac{d}{dt}|_{t=0} F(c(t)) = \langle \operatorname{grad} F(x), c'(0)\rangle = 0.$$

PROOF. The function $t \mapsto F(c(t))$ has a local extremum at $t = 0$ and therefore the first derivative at $t = 0$ must vanish. □

We shall now prove a version of the Lagrangian multiplier theorem for affine subspaces $U \subset \mathbb{R}^m$. We take a affine subspace $U \subset \mathbb{R}^m$ and an open neighborhood $V \subset \mathbb{R}^m$ such that $U \cap V \neq \emptyset$, where a C^2-function $F : V \to \mathbb{R}$ is defined.

2.3. Theorem. *Let x be a local maximum (local minimum) of F on $U \cap V$ and assume that there are $k := m - \dim U$ vectors $l_1, \dots, l_k \in \mathbb{R}^m$ and real numbers $a_1, \dots, a_k \in \mathbb{R}$ such that*

$$U = \{x \in V \text{ with } \langle l_i, x\rangle = a_i \text{ for } i = 1, \dots, k\}.$$

Then

$$\operatorname{grad} F(x) \in \langle l_1, \dots, l_k\rangle$$

or in other words there are real numbers $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ such that

$$\operatorname{grad} F(x) = \lambda_1 l_1 + \cdots + \lambda_k l_k.$$

Proof. Take a C^2-curve $c :]-1, +1[\to V$, then c takes values in U if and only if

$$c(0) \in U$$

and

$$\langle l_i, c'(t)\rangle = 0$$

for $i = 1, \dots, k$ and $t \in]-1, 1[$. The proof is simply done by Taylor's formula. Fix $t \in]-1, 1[$ and take

$$c(t) = c(0) + \int_0^t c'(s)ds.$$

By definition $c(t) \in U$ if and only if $\langle l_i, c(t)\rangle = a_i$, but

$$\begin{aligned}\langle l_i, c(t)\rangle &= \langle l_i, c(0)\rangle + \int_0^t \langle l_i, c'(s)\rangle\, ds \\ &= a_i\end{aligned}$$

by assumption for $i = 1, \dots, k$. We denote the span of $l_1, \dots, l_k$ by T and can consequently state that a C^2-curve $c :]-1, +1[\to V$ takes values in U if and only if $c(0) \in U$ and $c'(t) \in T^0$ for all $t \in]-1, 1[$. Furthermore we can say that T^0 is generated by all derivatives of C^2-curves $c :]-1, +1[\to V$ taking values in U at time $t = 0$ (simply take a line with direction a vector in T^0 through some point of U).

By the previous lemma we know that for all C^2-curves $c :]-1, +1[\to V$ with $c(0) = x$ the relation

$$\langle \operatorname{grad} F(x), c'(0)\rangle = 0$$

holds. Therefore $\operatorname{grad} F(x) \in T^{00}$. By the bipolar theorem we know that $T^{00} = T = \langle l_1, \dots, l_k\rangle$, which proves the result. □

2.4. Remark. This leads immediately to the receipt of Lagrangian multipliers as it is well known from basic calculus: a necessary condition for an extremal point of $F : V \to \mathbb{R}$ subject to the conditions $\langle l_i, x\rangle = a_i$ for $i = 1, \dots, k$ is to solve the extended problem with the Lagrangian L

$$L(x, \lambda_1, \dots, \lambda_k) = F(x) - \sum_{i=1}^{k} \lambda_i(\langle l_i, x\rangle - a_i).$$

Taking the gradients leads to the system of equations

$$\operatorname{grad} F(x) - \sum_{i=1}^{k} \lambda_i l_i = 0$$
$$\langle l_i, x \rangle = a_i$$

for $i = 1, \dots, k$, which necessarily has a solution if there is an extremal point at x.

2.5. Remark. How to calculate a gradient? The gradient of a C^1-function $F : V \to \mathbb{R}$ on a finite dimensional vector space V is defined through

$$\langle \operatorname{grad} F(x), w \rangle = \frac{d}{ds}|_{s=0} F(x + sw),$$

for $x \in V$ and $w \in \mathbb{R}^n$ (and a scalar product!). This can be calculated with respect to any basis and gives a coordinate representation. The derivative of F is understood as element of the dual space

$$dF(x)(w) := \frac{d}{ds}|_{s=0} F(x + sw)$$

for $x \in V$ and $w \in \mathbb{R}^n$ (even without scalar product!). The derivative can be calculated with respect to a basis $(e_i)_{i=1,\dots,\dim V}$. That means that it simply represents a collection of directional derivatives of a function, i.e.

$$\operatorname{grad}_{(e_i)} F(x) := (\frac{d}{ds}|_{s=0} F(x + se_i))_{i=1,\dots,\dim V}$$

for $x \in V$.

3. Conjugate Functions

Given a concave, increasing function $u : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$, which usual conventions for the calculus with $-\infty$. We denote by $\operatorname{dom}(u)$ the set $\{u > -\infty\}$ and assume that $\operatorname{dom}(u)$ is either $]0, \infty[$ or $\mathbb{R}$. We shall always assume that u is strictly concave and C^2 on $\operatorname{dom}(u)$ and that

$$\lim_{x \downarrow 0} u(x) = -\infty$$

in the case $\operatorname{dom}(u) =]0, \infty[$ and

$$\lim_{x \to -\infty} u(x) = -\infty$$

in the case $\operatorname{dom}(u) = \mathbb{R}$.

In this and more general cases we can define the conjugate function

$$v(y) := \sup_{x \in \mathbb{R}} (u(x) - yx)$$

for $y > 0$.

Since the function $x \mapsto u(x) - yx$ is strictly concave for every $y > 0$, there is some hope for a maximum. If there is one, let's say $\widehat{x}$, then it satisfies

$$u'(\widehat{x}) = y. \tag{3.1}$$

Since the second derivative exists and is strictly negative, $\widehat{x}$ is a local maximum if the above equation is satisfied. By strict concavity the local maximum is unique and global, too.

We need basic assumptions for the existence and regularity of the conjugate function:

(1) If $\mathrm{dom}(u) =]0,\infty[$ (negative wealth not allowed), then we assume

$$\lim_{x\downarrow 0} u'(x) = \infty,$$
$$\lim_{x\to\infty} u'(x) = 0 \text{ (marginal utility tends to 0).}$$

(2) If $\mathrm{dom}(u) = \mathbb{R}$ (negative wealth allowed), then we assume

$$\lim_{x\downarrow -\infty} u'(x) = \infty,$$
$$\lim_{x\to\infty} u'(x) = 0 \text{ (marginal utility tends to 0).}$$

Under these assumptions we can state the following theorem on existence and convexity of v.

3.1. Theorem. *Let $u : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ be a concave function satisfying the above assumptions, then the conjugate function is strictly convex and C^2 on* $\mathrm{dom}(v) =]0,\infty[$. *Additionally for* $\mathrm{dom}(u) =]0,\infty[$ *we have*

$$v'(0) := \lim_{y\downarrow 0} v'(y) = -\infty,$$
$$\lim_{y\to\infty} v'(y) = 0$$

and for $\mathrm{dom}(u) = \mathbb{R}$

$$v'(0) := \lim_{y\downarrow 0} v'(y) = -\infty,$$
$$\lim_{y\to\infty} v'(y) = \infty$$

Furthermore the inversion formula

$$u(x) = \inf_{y>0}(v(y) + xy)$$

holds true.

Proof. By formula 3.1 and our assumptions we see that for every $y > 0$ there is exactly one $\widehat{x}$, since u' is strictly decreasing and C^1. We denote the inverse of u' by $(u')^{-1}$. Therefore v is well-defined and at least C^1, since the inverse is C^1. Furthermore

$$\begin{aligned}
v(y) &= u((u')^{-1}(y)) - y \cdot (u')^{-1}(y) \\
v'(y) &= u'((u')^{-1}(y))((u')^{-1})'(y) - (u')^{-1}(y) - y((u')^{-1})'(y) \\
&= -(u')^{-1}(y) \\
v''(y) &= -((u')^{-1})'(y) = -\frac{1}{u''((u')^{-1}(y))} > 0
\end{aligned}$$

Hence v is C^2 on $]0,\infty[$ and a fortiori, by $v'' > 0$, strictly convex.

We know that u' is positive and strictly decreasing from ∞ to 0 by the previous assumptions, hence the two limiting properties for v, since $v'(y) = -(u')^{-1}(y)$.

Replacing v by $-v$, we can apply the same reasoning for existence of the concave conjugate of v. Take $\widehat{y} > 0$ such that $\inf_{y>0}(v(y) + xy)$ takes the infimum, then necessarily

$$v'(\widehat{y}) = -x,$$

hence $-(u')^{-1}(\widehat{y}) = -x$ and therefore $\widehat{y}(x) = u'(x)$. Inserting yields

$$\begin{aligned} v(u'(x)) + x\widehat{y}(x) &= u((u')^{-1}(u'(x))) - u'(x)(u')^{-1}(u'(x)) + xu'(x) \\ &= u(x), \end{aligned}$$

which is the desired relation. □

4. Methods from Probability Theory

In this section we shall fix notations and introduce stochastic processes on finite probability spaces. Even though all spaces which are going to appear are finite dimensional spaces, we shall introduce different norms or even metrics on them to focus on the correct functional analytic background. This way one can easily generalize the results to the continuous time setting.

In the sequel we denote by Ω a finite, non-empty set. A subset $\mathcal{F} \subset 2^\Omega$ of the power set is called a *σ-algebra* if it is closed under countable unions, closed under taking complements and contains Ω. A *probability measure* is a map

$$P : \mathcal{F} \to \mathbb{R}$$

such that

- for all mutually disjoint sequences $(A_n)_{n\geq 0} \in \mathcal{F}$ we have $P(\cup_{n\geq 0} A_n) = \sum_{n\geq 0} P(A_n)$.
- $P(\Omega) = 1$.

In the case of finite probability spaces a measure is given by its values on the *atoms* of the σ-algebra, i.e. the non-empty sets $A \in \mathcal{F}$ such that any subset $B \subset A$ with $B \in \mathcal{F}$ we have either $B = \emptyset$ or $B = A$. Any set $C \in \mathcal{F}$ can be decomposed uniquely into atoms, i.e.

$$C = \cup_{\substack{A \text{ is atom} \\ A\subset C}} A.$$

We denote the set of atoms by $\mathcal{A}(\mathcal{F})$. We denote the set of all probability measures on $(\Omega, \mathcal{F})$ by $\mathbb{P}(\Omega)$ and can characterize these measures as maps from the atoms of $\mathcal{F}$ to the non-negative real numbers such that sum over all atoms equals 1.

When we speak of a probability space $(\Omega, \mathcal{F}, P)$ we shall always assume that $\mathcal{F}$ is complete with respect to P, i.e. for every set $B \subset \Omega$, such that $B \subset A$ with $A \in \mathcal{F}$ and $P(A) = 0$, we have $B \in \mathcal{F}$.

We call such sets *P-nullsets*. The *P-completeness* assumption allows to deal with maps, which are defined up to sets of probability 0.

A random variable $X : (\Omega, \mathcal{F}) \to \mathbb{R}^d$ is a measurable map, i.e. the inverse image of Borel measurable sets is measurable in $\mathcal{F}$. The set of measurable maps is denoted by $L^0(\Omega, \mathcal{F}, P; \mathbb{R}^d)$, a measurable map takes constant values on each atom of the measurable space and we denote these values by $X(A)$ for A an atom in $\mathcal{F}$.

Given a set $M \subset 2^\Omega$, there is a smallest σ-algebra containing M denoted by $\sigma(M)$. If the set M is given as inverse image of Borel subsets from $\mathbb{R}^d$ via a map $X : \Omega \to \mathbb{R}^d$, then we write for the generated σ-algebra $\sigma(X)$. This is the smallest σ-algebra such that X is measurable $X : (\Omega, \sigma(X)) \to \mathbb{R}^d$.

Given a probability space $(\Omega, \mathcal{F}, P)$ we can define the *expectation* $E(X)$ of a random variable via

$$E(X) := \sum_{A \text{ is atom}} P(A)X(A)$$

if X is finitely valued. The *p-th moment* of X is given by $E(||X||^p)$ for $p \geq 1$, the *variance* $\mathrm{var}(X)$ of X is given by $E(||X - E(X)||^2)$ and the *covariance* of two random variables $X, Y \in L^0(\Omega, \mathcal{F}, P; \mathbb{R}^d)$ through $\mathrm{cov}(X,Y) = E((X - E(X))(Y - E(Y))^T)$.

We shall at least formally make a difference between the following spaces (with respect to their topologies). On $L^0(\Omega, \mathcal{F}, P; \mathbb{R}^d)$ we consider *convergence in probability* which means $X_n \to X$ if $P(||X_n - X|| \geq \epsilon) \to 0$ as $n \to \infty$ for each $\epsilon > 0$. This means that each sequence converges on atoms pointwisely. On

$$L^p(\Omega, \mathcal{F}, P; \mathbb{R}^d) = \{X \in L^0 \text{ such that } E(||X||^p) < \infty\}$$

we consider *L^p-convergence* due to $X_n \to X$ if $E(||X_n - X||^p) \to 0$ as $n \to \infty$ for each $p \geq 1$, which coincides with L^0 on finite probability spaces.

Additionally, $L^2(\Omega, \mathcal{F}, P; \mathbb{R}^d)$ is an euclidean vector space with scalar product

$$\langle X, Y \rangle = E(XY)$$

for $X, Y \in L^2(\Omega, \mathcal{F}, P; \mathbb{R}^d)$.

$L^\infty(\Omega, \mathcal{F}, P; \mathbb{R}^d)$ is the set of bounded random variables with the *supremum norm*, which is also equal to L^0. If there is no doubt about the image space of the random variables, we leave the set after the semi-colon away.

A sequence $(X_n)_{n \geq 0}$ is said to converge *P-almost surely* to X if $X_n \to X$ outside a null set as $n \to \infty$. Notice that all these different topologies coincide even though the metrics or norms are different, since all the spaces are finite dimensional. In particular we can identify the probability measures on $(\Omega, \mathcal{F})$ with some linear functionals on $L^\infty(\Omega, \mathcal{F}, P)$, namely those positive linear functionals $l \in (L^\infty)'$ such that $l(1_\Omega) = 1$. The space $(L^\infty)'$ can be naturally identified with L^1.

Two sets $A, B \in \mathcal{F}$ are called *independent* if $P(A \cap B) = P(A)P(B)$. For more than two sets we have the appropriate, generalized notion. Two σ-algebras $\mathcal{G}_1, \mathcal{G}_2$ are called independent if for all $A_i \in \mathcal{G}_i, i = 1, 2$ the sets A_1 and A_2 are independent. A random variable X is called independent of $\mathcal{G}$ if $\mathcal{G}$ and $\sigma(X)$ are independent.

Consider $(\Omega, \mathcal{F}, P)$ a probability space, $X \in L^1(\Omega, \mathcal{F}, P)$ and a P-complete σ-algebra $\mathcal{G} \subset \mathcal{F}$, which contains all P-nullsets, then we can define the *conditional expectation* $E(X|\mathcal{G})$ via the property

- $E(X|\mathcal{G})$ is a $\mathcal{G}$-measurable random variable,
- for all $Y \in L^\infty(\Omega, \mathcal{G}, P)$ we have $E(XY) = E(E(X|\mathcal{G})Y)$.

Notice that $L^p(\Omega, \mathcal{G}, P)$ is a closed subspace of $L^p(\Omega, \mathcal{F}, P)$ for $1 \leq p \leq \infty$. The conditional expectation is well defined since $E(X.) : L^\infty(\Omega, \mathcal{G}, P) \to \mathbb{R}$ is a well-defined, continuous (absolutely continuous) linear functional and defines therefore an element of $L^1(\Omega, \mathcal{G}, P)$ by duality.

We can immediately write down the following Lemma on conditional expectations:

4.1. Lemma. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ be complete sub-σ-algebras, then*

- *for all $X \in L^1(\Omega, \mathcal{G}, P)$ we have $E(X|\mathcal{G}) = X$.*
- *the conditional expectation $E(.|\mathcal{G})$ is a linear map on $L^p(\Omega, \mathcal{F}, P)$ and an orthogonal projection as map from $L^2($ to $L^2(\Omega, \mathcal{F}, P)$.*
- *the conditional expectation is a positive map, i.e. $E(X|\mathcal{G}) \geq 0$ if $X \geq 0$.*
- *the tower law holds, $E(E(X|\mathcal{G})|\mathcal{H}) = E(X|\mathcal{H})$ for all $X \in L^1(\Omega, \mathcal{F}, P)$.*
- *Jensen's inequality holds, i.e. for convex $\phi : \mathbb{R} \to \mathbb{R}$ we have $\phi(E(X|\mathcal{G})) \leq E(\phi(X)|\mathcal{G})$ for $X \in L^1(\Omega, \mathcal{F}, P)$.*

- *for all* $Z \in L^1((\Omega, \mathcal{G}, P)$ *we have*
$$E(ZX|\mathcal{G}) = ZE(X|\mathcal{G})$$
for $X \in L^1(\Omega, \mathcal{F}, P)$.
- *If* X *is independent of* $\mathcal{G}$ *then* $E(X|\mathcal{G}) = E(X)$.
- *Let* $X, Y \in L^1(\Omega, \mathcal{F}, P)$ *be given and take* σ*-algebras* $\mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{F}$. *Assume* $A \in \mathcal{G}_1 \cap \mathcal{G}_2$ *such that* $X = Y$ *on* A *and* $A \cap \mathcal{G}_1 = A \cap \mathcal{G}_2$ *(in this case the* σ*-algebras* $\mathcal{G}_1, \mathcal{G}_2$ *are called locally on* A *equal* σ*-algebras). Then* $E(X|\mathcal{G}_1) = E(Y|\mathcal{G}_2)$ *on* A.
- *We denote the atoms of* $\mathcal{G}$ *by* $\mathcal{A}(\mathcal{G})$, *then we have*
$$E(X|\mathcal{G}) = \sum_{\substack{A \in \mathcal{A}(\mathcal{G}) \\ P(A) \neq 0}} \frac{E(1_A X)}{P(A)} 1_A.$$
Consequently the conditional expectation is well-defined up to sets of probability 0.

Proof. We prove Jensen's inequality, which follows directly from the fact that
$$\phi(x) = \sup_{\substack{ay+b \leq \phi(y) \\ \text{for all } y \in \mathbb{R}}} (ax + b).$$
This yields by linearity
$$\phi(E(X|\mathcal{G})) = \sup_{\substack{ay+b \leq \phi(y) \\ \text{for all } y \in \mathbb{R}}} E(aX + b)|\mathcal{G}) \leq E(\phi(X)|\mathcal{G}).$$
The assertion on independent random variables follows from
$$E(XY) = E(X)E(Y) = E(E(X)Y)$$
by independence for $X \in L^1(\Omega, \mathcal{F}, P)$, $Y \in L^\infty(\Omega, \mathcal{G}, P)$ independent of $\mathcal{G}$.

For the assertion on locally on A equal σ-algebras we take $1_A E(X|\mathcal{G}_1)$ and $1_A E(Y|\mathcal{G}_2)$, which are $\mathcal{G}_1 \cap \mathcal{G}_2$-measurable by locality. Define $B := A \cap \{E(X|\mathcal{G}_1) \geq E(X|\mathcal{G}_2)\} \in \mathcal{G}_1 \cap \mathcal{G}_2$, then
$$E(E(X|\mathcal{G}_1)B) = E(XB) = E(YB) = E(E(X|\mathcal{G}_2)B),$$
hence $E(X|\mathcal{G}_1) \leq E(X|\mathcal{G}_2)$ on A. Take the other direction and conclude the result.

For the last formula we take $Y \in L^0(\Omega, \mathcal{G}, P)$, which is constant on atoms of $\mathcal{G}$, hence
$$\begin{aligned} E(XY) &= \sum_{A \in \mathcal{A}(\mathcal{G})} E(X1_A)Y(A) \\ &= E(\sum_{A \in \mathcal{A}(\mathcal{G})} \frac{E(X1_A)}{P(A)} Y(A)), \end{aligned}$$
which proves the result. □

4.2. Remark. The interpretation of the conditional expectation is the following. Given the information of a σ-algebra $\mathcal{G}$, i.e. all random variables generating the σ-algebra $\mathcal{G}$, then one can calculate $E(X|\mathcal{G})$ as the best L^2-approximation of X in the set of random variables generating $\mathcal{G}$.

A *filtration* on $(\Omega, \mathcal{F}, P)$ is a finite sequence of σ-algebras

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_N \subset 2^\Omega,$$

where $\mathcal{F} = \mathcal{F}_N$ for $N \geq 1$. Filtrations represent increasing degrees of information on probability space. With these preparations we can formulate basic ideas of the theory of martingales. We shall always assume that $\mathcal{F}_0$ (and hence all contains all$\mathcal{F}_n$) contains all P-nullsets.

- A *stochastic process* on $(\Omega, \mathcal{F}, P)$ is a sequence of $\mathbb{R}^d$-valued random variables $(X_n)_{0 \leq n \leq N}$.
- A stochastic process $(X_n)_{0 \leq n \leq N}$ is called *adapted* to a filtration $(\mathcal{F}_n)_{0 \leq n \leq N}$ if X_n is $\mathcal{F}_n$-measurable for $0 \leq n \leq N$. In this case we shall often speak of an adapted process if there is no doubt about the filtration.
- A stochastic process $(H_n)_{0 \leq n \leq N}$ is called *predictable* if H_0 is constant and H_n is $\mathcal{F}_{n-1}$-measurable for $1 \leq n \leq N$. A predictable process is certainly adapted. It appears often that H_0 is redundant, however for our applications there is some use.
- Let $(H_n)_{0 \leq n \leq N}, (X_n)_{0 \leq n \leq N}$ be stochastic processes, then we define the Riemannian sum for $0 \leq n \leq N$

  $$(H \cdot X)_n := \sum_{i=1}^{n} H_i (X_i - X_{i-1}),$$

  where we take the scalar product of vectors in $\mathbb{R}^d$ in the sum. We can write down the basic partial integration relation

  $$(H \cdot X)_n = H_N X_N - H_0 X_0 - (X_{*-1} \cdot H)_n,$$

  where $(X_{*-1})_n := X_{n-1}$ for $1 \leq n \leq N$ and $(X_{*-1})_0 = X_0$.

4.3. Definition. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_n)_{0 \leq n \leq N}$ a filtration, then a sequence of $\mathbb{R}^d$-valued random variables $(M_n)_{0 \leq n \leq N}$ is called a* martingale *if*

$$E(M_n | \mathcal{F}_m) = M_m$$

for $0 \leq m \leq n \leq N$. The sequence is called a submartingale *(*supermartingale*) if $E(M_n | \mathcal{F}_m) \geq M_m$ ($E(M_n | \mathcal{F}_m) \leq M_m$ respectively) for $0 \leq m \leq n \leq N$.*

4.4. Definition. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_n)_{0 \leq n \leq N}$ a filtration, then a random variable $\tau : \Omega \to \mathbb{N}_{\geq 0}$ is called a stopping time if*

$$\{\tau \leq n\} \in \mathcal{F}_n$$

for $0 \leq n \leq N$. Let M be an adapted process and τ a stopping time with $\tau \leq N$ almost surely, then we can define

$$M_\tau(\omega) := M_{\tau(\omega)}(\omega)$$

for $\omega \in \Omega$. The stopped process M^τ is defined for any stopping time τ

$$M_n^\tau := M_{\tau \wedge n}$$

for $0 \leq n \leq N$. The stopped σ-algebra

$$\mathcal{F}_\tau := \{A \in \mathcal{F} \text{ such that } A \cap \{\tau \leq n\} \in \mathcal{F}_n \text{ for } 0 \leq n \leq N\}$$

contains all informations from the stopping time τ.

Indeed one can easily prove that $\mathcal{F}_\tau$ is a σ-algebra and that τ is $\mathcal{F}_\tau$-measurable, we can state the following basic Lemma:

4.5. Lemma. *Let $\tau, \eta, \eta_1, \eta_2, \ldots$ be stopping times, then*

- $\sum_{i=1}^k \eta_k$, $\inf \eta_i$, $\sup \eta_i$, $\limsup \eta_i$, $\liminf \eta_i$ *are stopping times.*
- *If $\tau \leq \eta$ bounded by N then $\mathcal{F}_\tau \subset \mathcal{F}_\eta$ and the sets $\{\tau \leq \eta\}$ and $\{\eta \leq \tau\}$ lie in $\mathcal{F}_{\tau\wedge\eta} = \mathcal{F}_\tau \cap \mathcal{F}_\eta$.*
- *If τ, η bounded by N, then $\{\tau \leq \eta\} \cap \mathcal{F}_\tau \subset \mathcal{F}_{\tau\wedge\eta}$.*
- *If τ bounded by N, then $\mathcal{F}_\tau = \mathcal{F}_n$ on $\{\tau = n\}$, i.e. $\{\tau = n\} \cap \mathcal{F}_\tau = \{\tau = n\} \cap \mathcal{F}_n$.*
- *Let τ be bounded by N. If $A \in \mathcal{F}_\tau$, then $\tau_A = \tau 1_A + N 1_{A^c}$ is a stopping time.*
- *Given an adapted sequence of random variables M and τ,η stopping times bounded by N, M_τ is $\mathcal{F}_\tau$-measurable and $E(M_\tau|\mathcal{F}_\eta)$ is $\mathcal{F}_{\tau\wedge\eta}$-measurable.*

PROOF. The proofs follow directly from the definition. We have that τ is a stopping time if and only if $\{\tau \leq n\} \in \mathcal{F}_n$ if and only if $\{\tau = n\} \in \mathcal{F}_n$ if and only if $\{\tau > n\} \in \mathcal{F}_n$ for all $0 \leq n \leq N$ respectively. We have

$$\begin{aligned}\{\inf \eta_i = n\} &= \{\omega, \eta_i(\omega) \geq n \text{ and } i_0 \text{ such that } \eta_{i_0}(\omega) = n\} \\ &= \cap_i \{\eta_i \geq n\} \cap (\cup_i \{\eta_i = n\}) \in \mathcal{F}_n.\end{aligned}$$

For the supremum we have

$$\begin{aligned}\{\sup \eta_i = n\} &= \{\omega, \eta_i(\omega) \leq n \text{ and } i_0 \text{ such that } \eta_{i_0}(\omega) = n\} \\ &= \cap_i \{\eta_i \leq n\} \cap (\cup_i \{\eta_i = n\}) \in \mathcal{F}_n.\end{aligned}$$

For the limits we can proceed in the same way. By positivity of stopping times we see that the sum is smaller than n if all entries are, hence the result. For the second assertion we know that

$$\{\tau \leq \eta\} \cap \{\tau \wedge \eta \leq n\} = \cup_{i=0}^n \{\eta \geq i\} \cap \{\tau = i\} \in \mathcal{F}_n.$$

Furthermore for $A \in \mathcal{F}_\tau \cap \mathcal{F}_\eta$ we see

$$A \cap \{\tau \wedge \eta \leq n\} = A \cap \{\tau \leq \eta\} \cap \{\tau \leq n\} \cup A \cap \{\eta \leq \tau\} \cap \{\eta \leq n\} \in \mathcal{F}_n.$$

Next we prove a locality statement

$$\{\tau \leq \eta\} \cap \mathcal{F}_\tau \subset \mathcal{F}_{\tau\wedge\eta},$$

which is clear by the assertion that for $A \in \mathcal{F}_\tau$

$$A \cap \{\tau \leq \eta\} \cap \{\eta \leq n\} = (A \cap \{\eta \leq n\}) \cap \{\tau \leq n\} \cap \{\tau \wedge n \leq \eta \wedge n\} \in \mathcal{F}_n,$$

hence $A \cap \{\tau \leq \eta\} \in \mathcal{F}_\eta$, but $A \cap \{\tau \leq \eta\} \in \mathcal{F}_\tau$ anyway. If $\tau = n$, then clearly $\mathcal{F}_\tau = \mathcal{F}_n$ by definition. If τ is general, then the previous assertion tells $\{\tau \leq n\} \cap \mathcal{F}_\tau \subset \mathcal{F}_n$, hence $\{\tau = n\} \cap \mathcal{F}_\tau \subset \mathcal{F}_n$. Therefore $\{\tau = n\} \cap \mathcal{F}_\tau \subset \{\tau = n\} \cap \mathcal{F}_n$. Interchanging the roles of n and τ yields the result, namely $\{n \leq \tau\} \cap \mathcal{F}_n \subset \mathcal{F}_\tau$, so $\{n = \tau\} \cap \mathcal{F}_n \subset \mathcal{F}_\tau$ and whence the assertion.

For the next assertion we know that

$$\{\tau_A = n\} = A \cap \{\tau = n\}$$

for $n < N$ and $\{\tau = N\} = A^c \cup \{\tau = n\} \in \mathcal{F}_N$. For the last assertion we conclude by

$$\{M_\tau \in A\} \cap \{\tau \leq n\} = \cup_{i=0}^n \{M_i \in A\} \cap \{\tau = i\} \in \mathcal{F}_n$$

by adaptedness. From the last assertions we know that locally on $\{\tau \leq \eta\}$ the σ-algebras $\mathcal{F}_\eta$ and $\mathcal{F}_{\tau\wedge\eta}$ agree and on $\{\tau \leq \eta\}$ the random variables M_τ and $E(M_\tau|\mathcal{F}_\eta)$ agree, hence

$$E(M_\tau|\mathcal{F}_{\tau\wedge\eta}) = E(M_\tau|\mathcal{F}_\eta)$$

on $\{\tau \leq \eta\}$ by Lemma 4.1. On $\{\tau \geq \eta\}$, where locally the σ-algebras $\mathcal{F}_\eta$ and $\mathcal{F}_{\tau\wedge\eta}$ agree, the random variables M_τ and M_τ agree and hence

$$E(M_\tau|\mathcal{F}_{\tau\wedge\eta}) = E(M_\tau|\mathcal{F}_\eta)$$

by Lemma 4.1. Consequently $E(M_\tau|\mathcal{F}_\eta)$ is $\mathcal{F}_{\tau\wedge\eta}$-measurable. □

4.6. Theorem (Doob's optional sampling)**.** *Let $(\Omega, \mathcal{F}, P)$ be a finite probability space and $(\mathcal{F}_n)_{0\leq n\leq N}$ a filtration. Let $(M_n)_{0\leq n\leq N}$ be an adapted process.*

(1) *If M is a martingale, then for every predictable process $(H_n)_{0\leq n\leq N}$ the stochastic integral $(H\cdot M)$ is a martingale. In particular $E((H\cdot M)_N) = 0$ and $E(M_\tau) = E(M_0)$ for all stopping times $\tau \leq N$.*

(2) *If the stochastic integral $(H \cdot M)$ satisfies*

$$E((H \cdot M)_N) = 0$$

for every predictable process H, then M is a martingale.

(3) *If for all stopping times $\tau \leq N$*

$$E(M_\tau) = E(M_0)$$

holds, then M is a martingale.

(4) *If M is martingale, then for all stopping times $\eta \leq \tau \leq N$ almost surely we have*

$$E(M_\tau|\mathcal{F}_\eta) = M_\eta.$$

More generally we have that for any two stopping times $\tau, \eta \leq N$

$$E(M_\tau|\mathcal{F}_\eta) = M_{\tau\wedge\eta}.$$

Proof. We prove the four assertions step-by-step:

- Let M be a martingale, then for $n \geq m$

$$\begin{aligned} E(\sum_{i=1}^{n} H_i(M_i - M_{i-1})|\mathcal{F}_m) &= E(\sum_{i=m+1}^{n} H_i E(M_i - M_{i-1}|\mathcal{F}_{i-1})|\mathcal{F}_m) + \\ &\quad + (H \cdot M)_m \\ &= (H \cdot M)_m \end{aligned}$$

  by the martingale property, the predictability of H and Lemma 4.1 . Since $(H \cdot M)_0 = 0$ we obtain $E((H \cdot M)_N) = 0$. We define the predictable (!) process

$$H_n := 1_{\{\tau > n-1\}} = 1 - 1_{\{\tau \leq n-1\}}$$

  for $1 \leq n \leq N$ with $H_0 = 0$ we obtain

$$(H \cdot M)_N = M_\tau - M_0.$$

- We construct several predictable processes H, namely fix $1 \leq j \leq N$ and $A \in \mathcal{F}_j$, then we define

$$\begin{aligned} H_n &= 0 \text{ for } n \neq j+1 \\ H_{j+1} &= 1_A \end{aligned}$$

and the hypothesis leads to $E(1_A(M_{j+1}-M_j)) = 0$, hence we can conclude $E(M_{j+1}|\mathcal{F}_j) = M_j$.

- For the constant stopping time $1 \le n \le N$ we know that
$$\tau = 1_A n + N 1_{A^c}$$
is a stopping time for $A \in \mathcal{F}_n$, furthermore $E(M_n) = E(M_0)$ But then
$$\begin{aligned} E(M_N 1_{A^c} + M_n 1_A) &= E(M_0), \\ E((M_N - M_n)1_{A^c} + M_n) &= E(M_0), \\ E((M_N - M_n)1_{A^c}) &= 0. \end{aligned}$$
Consequently $E(M_N|\mathcal{F}_n) = M_n$ which yields the martingale property.
- This is the main assertion of Doob's optional sampling theorem. Assume that M is a martingale, then we know for $n \le N$ that
$$E(M_N|\mathcal{F}_\tau) = E(M_N|\mathcal{F}_n) = M_n = M_\tau$$
on $\{\tau = n\}$ by Lemma 4.5. So $E(M_N|\mathcal{F}_\tau) = M_\tau$ on $\{\tau \le N\}$. If $\eta \le \tau \le N$ then
$$\begin{aligned} E(M_\tau|\mathcal{F}_\eta) &= E(E(M_N|\mathcal{F}_\tau)|\mathcal{F}_\eta) \\ &= E(M_N|\mathcal{F}_\eta) = M_\eta \end{aligned}$$
by the tower law, which proves the result. Now the general case for two stopping times η, τ
$$E(M_\tau|\mathcal{F}_\eta) = E(M_\tau|\mathcal{F}_{\tau\wedge\eta}) = M_{\tau\wedge\eta} \text{ on } \{\eta \le \tau\}$$
since the σ-algebras $\mathcal{F}_\eta$ and $\mathcal{F}_{\tau\wedge\eta}$ agree on $\{\eta \le \tau\}$. Furthermore
$$E(M_\tau|\mathcal{F}_\eta) = E(M_{\tau\wedge\eta}|\mathcal{F}_\eta) = M_{\tau\wedge\eta} \text{ on } \{\eta \ge \tau\}$$
since the random variables M_τ and $M_{\tau\wedge\eta}$ agree on $\{\eta \ge \tau\}$.

□

A particular application for martingales is the following Lemma. Therefore we need the notion of an equivalent measure $Q \sim P$, i.e. a measure such that for all $A \in \mathcal{F}$ we have $P(A) = 0$ if and only if $Q(A) = 0$. Given any measure Q on Ω we define the Radon-Nikodym derivative $\frac{dQ}{dP}$ as random variable, such that for all $Z \in L^0(\Omega, \mathcal{F}, P)$,
$$E_Q(Z) = E_P(Z\frac{dQ}{dP}).$$
Hence we obtain
$$\frac{dQ}{dP}(A) = \frac{Q(A)}{P(A)}$$
for all atoms $A \in \mathcal{A}(\mathcal{F})$ with $P(A) > 0$. A measure Q is called absolutely continuous with respect to P if for all $A \in \mathcal{F}$ with $P(A) = 0$ we have that $Q(A) = 0$. In the generic case of $P(\omega_i) > 0$ for all $i = 1, \ldots, |\Omega|$ every measure Q is absolutely continuous with respect to P.

4.7. Lemma (change of measure)**.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space with filtration $(\mathcal{F}_n)_{0\le n\le N}$ and Q be an equivalent probability measure such that*
$$\frac{dQ}{dP} = X$$

for some $X \in L^1(\Omega, \mathcal{F}, P)$. *Then* $Q|_{\mathcal{F}_n}$ *are equivalent probability measures on* $(\Omega, \mathcal{F}_n, P|_{\mathcal{F}_n})$ *for* $n = 0, \ldots, N$ *and*

$$\frac{dQ_n}{dP_n} =: X_n$$

is a P*-martingale. Here* P_n *denotes the restriction of* P *to* $\mathcal{F}_n$. *Furthermore we have the following formulas*

$$E_P(X|\mathcal{F}_n) = X_n$$

and

$$E_Q(Y|\mathcal{F}_n) = \frac{1}{X_n} E_P(YX|\mathcal{F}_n)$$

for all $Y \in L^1(\Omega, \mathcal{F}, Q)$. *In particular* $X_n > 0$ *almost surely with respect to* P.

PROOF. We know that X is strictly positive and $E_P(X) = 1$. The measures Q_n are certainly equivalent probability measures and we have

$$\begin{aligned} E_{Q_n}(Y) &= E_{P_n}(YX_n) \\ &= E_P(YX_n) \end{aligned}$$

for all $Y \in L^1(\Omega, \mathcal{F}_n, P)$, but also

$$\begin{aligned} E_{Q_n}(Y) &= E_Q(Y) \\ &= E_P(XY), \end{aligned}$$

which yields by definition of conditional expectations that $E_P(X|\mathcal{F}_n) = X_n$. In turn X_n is a martingale. Calculating now the conditional expectation with respect to Q yields to do

$$\begin{aligned} E_Q(YZ) &= E_P(YZX) \\ &= E_P(E_P(YX|\mathcal{F}_n)Z) \\ &= E_Q(\frac{1}{X_n} E_P(YX|\mathcal{F}_n)Z) \end{aligned}$$

for $Y \in L^1(\Omega, \mathcal{F}, Q)$ and $Z \in L^1(\Omega, \mathcal{F}_n, Q)$, which gives the desired relation. □

5. The central limit theorem

For purposes of comparison of discrete and continuous models, we shall apply the central limit theorem. Therefore we need the notion of a characteristic function:

5.1. Definition. *Let* X *be a real valued random variable on* $(\Omega, \mathcal{F}, P)$, *then the characteristic function*

$$\phi_X(u) := E(\exp(iuX))$$

is a well defined function for $u \in \mathbb{R}$.

5.2. Theorem. *Let* X *be a real valued random variable on* $(\Omega, \mathcal{F}, P)$, *then for the characteristic function the following properties hold:*

(1) *The function* ϕ_X *is continuous and* $\phi_X(0) = 1$.
(2) *For all* $u \in \mathbb{R}$ *we have that* $|\phi_X(u)| \leq 1$.
(3) *Given a real random variable* Y *independent of* X, *then*

$$\phi_{X+Y}(u) = \phi_X(u)\phi_Y(u)$$

for $u \in \mathbb{R}$.

(4) *Given a real valued random variable Y on $(\Omega, \mathcal{F}, P)$ such that*

$$\phi_X(u) = \phi_Y(u),$$

for $u \in \mathbb{R}$, then the distributions of X and Y are identical.

Proof. The first and second property follow from $\phi_X(0) = E(1) = 1$ and $|\phi_X(u)| \leq E(|\exp(iuX)|) = E(1) = 1$. The third property follows from independence since

$$\begin{aligned}\phi_{X+Y}(u) &= E(\exp(iuX)\exp(iuY)) = E(\exp(iuX))E(\exp(iuY)) \\ &= \phi_X(u)\phi_Y(u)\end{aligned}$$

for $u \in \mathbb{R}$.

For the forth property we know that the distribution of a random variable X is given by the distribution function

$$F_X(z) := P(X \leq z)$$

for $z \in \mathbb{R}$. The distribution function F_X can be calculated directly from the characteristic function via

$$F_X(z) = \frac{1}{2\pi} \int_{-\infty}^{z} \int_{-\infty}^{\infty} \exp(-iuv)\phi_X(u)dudv.$$

Hence if the characteristic functions coincide, the distribution functions coincide. □

For the proof of a simple version of the central limit theorem we shall apply Paul Chernoff's Theorem in an elementary form:

5.3. Theorem. *Let $c : \mathbb{R}_{\geq 0} \to \mathbb{C}$ be differentiable at 0 with the following properties:*

- $c(0) = 1$ *and* $c'(0) = k$,
- *there is $b > 0$ such that $|c(t)^m| \leq M$ for $0 \leq t \leq b$, $m \geq 1$ and some fixed constant M.*

Then the limit

$$\lim_{n \to \infty} c(\frac{t}{n})^n = e^{kt}$$

exist uniformly on compact intervals for $t \geq 0$.

5.4. Remark. The following proof is more complicated than the one of the lecture course, but it also works for in much more general situations. In the lecture we simply evaluated the inequality following from differentiability

$$(1 + s(k - \epsilon)k) \leq c(s) \leq (1 + s(k + \epsilon))$$

for $|s| < \delta$.

Proof. Take a complex number q with $|q| \leq 1$, then

$$|\exp(n(q-1)) - q^n| \leq \sqrt{n}|q - 1|$$

for $n \geq 0$. In fact we can write

$$
\begin{aligned}
|\exp(n(q-1)) - q^n| &= |e^{-n} \sum_{k=0}^{\infty} \frac{n^k}{k!}(q^k - q^n)| \\
&\leq e^{-n} \sum_{k=0}^{\infty} \frac{n^k}{k!} |q|^{k \wedge n} |1 - q^{|n-k|}| \\
&\leq M|1-q|e^{-n} \sum_{k=0}^{\infty} \frac{n^k}{k!} |n-k| \\
&\leq M|1-q|e^{-n} \sum_{k=0}^{\infty} \sqrt{\frac{n^k}{k!}} \sqrt{\frac{n^k}{k!}} |n-k| \\
&\leq M|1-q|e^{-n} \sqrt{\sum_{k=0}^{\infty} \frac{n^k}{k!}} \sqrt{\sum_{k=0}^{\infty} \frac{n^k}{k!}(n-k)^2} \\
&\leq M|1-q|e^{-n} e^{-\frac{n}{2}} \sqrt{n} e^{-\frac{n}{2}} = M\sqrt{n}|q-1|
\end{aligned}
$$

by the Cauchy-Schwartz inequality for infinite sums and the equality

$$
\sum_{k=0}^{\infty} \frac{n^k}{k!}(n-k)^2 = ne^n,
$$

which follows from direct calculation. We define now $q_t := \frac{c(t)-1}{t}$ for $t \geq 0$ with $q_0 = k$, the first derivative at $t = 0$. Then certainly by continuity

$$
\lim_{n \to \infty} e^{q_{\frac{t}{n}} t} = e^{kt}
$$

and

$$
\begin{aligned}
|e^{q_{\frac{t}{n}}} - c(\frac{t}{n})^n| &= |e^{n(c(\frac{t}{n})-1)} - c(\frac{t}{n})^n| \leq M\sqrt{n}|c(\frac{t}{n}) - 1| \\
&= \frac{Mt}{\sqrt{n}} |\frac{c(\frac{t}{n}) - 1}{\frac{t}{n}}| \to 0
\end{aligned}
$$

as $n \to \infty$ for $0 \leq t \leq b$ uniformly on compact intervals. Therefore we obtain convergence everywhere. □

Now we can prove a version of the central limit theorem straight forward:

5.5. Theorem. *Let $(X_n)_{n \geq 1}$ be a sequence of independent, identically distributed random variables on $(\Omega, \mathcal{F}, P)$ and and assume that $E(X_n) = a$ and $var(X_n) = \sigma^2 > 0$, then the sum*

$$
S_n = \frac{1}{\sqrt{n\sigma^2}} \sum_{i=1}^{n} (X_n - a)
$$

converges in distribution to $N(0,1)$. We shall prove that $\phi_{S_n}(u) \to e^{-\frac{u^2}{2}}$ uniformly on compact intervals.

Proof. By the properties of independent random variables we can define

$$
\begin{aligned}
\phi(u) &:= E(\exp(iu(X_n - a))) \\
\phi_{S_n}(u) &= \phi(\frac{u}{\sqrt{n\sigma^2}})^n
\end{aligned}
$$

for $n \geq 1$. ϕ is twice differentiable with derivative

$$\phi'(u) = iE((X-a)\exp(iuX))$$
$$\phi''(u) = -E((X-a)^2\exp(iuX))$$

by dominated convergence. We define next for $t \geq 0$

$$c(t) := \phi(\frac{\sqrt{t}}{\sqrt{\sigma^2}})$$

and obtain that $c(0) = 1$ and $c'(0) = -\frac{1}{2}$. Consequently by Chernoff's theorem

$$c(\frac{t}{n})^n \to e^{-\frac{t}{2}}$$

on compact intervals. Therefore with $t = u^2$ for $u \geq 0$

$$\phi(\frac{u}{\sqrt{n\sigma^2}})^n \to e^{-\frac{u^2}{2}}$$

for $u \geq 0$. For $u \leq 0$ we proceed in the same way, which yields the desired result. □

5.6. Corollary. *Let $(X_i^N)_{i=1,\dots,N}$ for $N \geq 1$ be independent sequences of i.i.d. random variables, such that X_i^N takes values $\frac{\sigma}{\sqrt{N}}, -\frac{\sigma}{\sqrt{N}}$ and has expectation μ_N for each $N \geq 1$ and*

$$\lim_{N\to\infty} N\mu_N = \mu,$$

then $\sum_{i=1}^N X_i^N$ converges in law to $N(\mu, \sigma^2)$.

Proof. Fix $N \geq 1$ and take an i.i.d. sequence X_n with the distribution of $\sqrt{N}X_i^N$, so it takes values σ and $-\sigma$. The expectation is given by $\sqrt{N}\mu_N$ and the variance $\sigma^2 - N\mu_N$, therefore

$$\frac{1}{\sqrt{M}}\sum_{i=1}^{M}(X_n - \sqrt{N}\mu_N) \to N(0, \sigma^2)$$

in law. By uniformity of the convergence with respect to N and M (the rate depends on derivative near 0 of the function c), we can conclude

$$\sum_{i=1}^{N} X_i^N \to N(\mu, \sigma^2).$$

□

CHAPTER 4

Questions for the oral exam

The following Theorems are relevant for the oral exam:

- No-Arbitrage in discrete time
- CRR
- Main theorem for complete markets
- Main theorem for incomplete markets
- Change of numeraire
- Solvable utility optimization problems lead to arbitrage-free markets
- dual method of utility optimization
- Bachelier and Black-Scholes formula
- Hahn-Banach
- Krein-Milman
- Bipolar Theorem
- Lagrangian multipliers
- Doob's optional sampling
- Change of measure, Radon-Nikodym-derivative