Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Duality and Utility Maximization

Bachelor Thesis

Niklas A. Pfister

July 11, 2013

Advisor: Prof. Dr. Halil Mete Soner
Department of Mathematics, ETH Zürich

Abstract

This thesis explores the problem of maximizing expected utility of terminal wealth in several different settings. We approach this problem through the theory of Legendre Transformation, hence arriving at a solution that has strong dual properties. These allow for an efficient method to calculate the maximizing strategy and also give a deeper understanding into the structural properties of the optimizers.

Throughout the thesis, we try to stay as self-contained as possible, building up the prerequisite material from both financial mathematics and convex analysis. We then begin in the setting of a finite underlying probability space and analyse the cases of complete and incomplete financial markets. We end the thesis by giving an outlook of how to extend our results in the case of a general underlying probability space.

Contents

Chapter 1

Introduction

Assume we know the stochastic process governing the prices of all assets in a financial market. We then consider the situation of an economic agent with known utility, who is investing in this market. This economic agent is allowed to invest during a finite time period in any way he pleases. The problem this thesis will be investigating, is finding the strategy that maximizes the expected utility of his wealth at the end of the time period. In other words, we are trying to maximize the utility of the terminal wealth of an economic agent investing in a financial market.

In this thesis we will see that for common utilities and with few assumptions on the stochastic price process, this problem can be mathematically solved. We will follow Schachermayer's paper [7] very closely.

Let us begin by introducing the mathematical model of the financial market and the investment strategies. This will be the building block in the formalization of the optimization problem described above, which is established in chapter 3.

1.1 Mathematical Model

1.1.1 Security Market

We consider a security market consisting of $d+1$ assets whose price process will be modelled by the random vector $S = [S^0, S^1, \dots, S^d] = ((S^i_t)_{t\in[0,T]})_{0\leqslant i\leqslant d}$ adapted to the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbf{P})$. We assume the standard condition that the measure $\mathbf{P}$ is saturated and the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is right continuous. The price of the 0^{th} asset S^0 will be called "bond" or "cash account" and will be constant, i.e. $S^0_t \equiv 1$. It corresponds to a no-risk asset, for example money on a bank account. Notice that by fixing the value of the bond we do not restrict our model because the other price processes can be interpreted with respect to the units of the "bond". In other literature this is referred to as *discounted price process*. Furthermore we assume that S is an $\mathbb{R}^{d+1}$-valued *bounded* semi-martingale, which describes the price process of the d risky assets.

At this point let us remark that all results in this paper can also be shown for S a *locally* bounded semi-martingale. In this situation one has to at times deal with local martingales instead of martingales. However using appropriate localizing sequences of stopping times, one can reduce back to the martingale setting. Surprisingly one can even take this one step further and make no restriction at all on S apart from it being a semi-martingale. This is done in Schachermayer's paper [7]. It requires an even more general notion called sigma-martingale. In order to evade these technicalities, we leave it to the reader to verify that this is possible.

The main part of this thesis will investigate the case where the probability space Ω is finite, in which case the price process S becomes constant except for a finite number of jumps. This allows us to write S as $(S_t)_{t=0}^T$ where $t, T \in \mathbb{N}$. A proof of this can be found in the appendix (see Prop A.1).

1.1.2 Investment Strategy

We now add investment strategies to our model. It is important that we only permit strategies that make sense in reality, which means that at each time step all investments add up to the investors wealth and no insider information is allowed. This is enforced by using the notion of a self-financing portfolio.

Definition 1.1 (self-financing portfolio) *A portfolio Π is defined as a pair (x, H), where the constant x is the initial value of the portfolio and $H = ((H_t^i)_{t\in[0,T]})_{0\leqslant i\leqslant d}$ is a predictable S-integrable process specifying the amount of each asset held in the portfolio at time t. The corresponding value process $X = (X_t)_{t\in[0,T]}$ which gives the value of the portfolio at every time t is defined as*

$$X_t = \sum_{i=0}^{d} H_t^i S_t^i \qquad \forall\ t \in [0, T]$$

The portfolio is said to be self-financing if

$$dX_t = \sum_{i=0}^{d} H_t^i dS_t^i \tag{1.1}$$

The predictable, S-integrable process H should be understood as an investment strategy by which an economic agent can invest in a financial market. The adaptedness and predictability guarantee that the investor only uses information given up to time t (i.e no insider information is used) while the

S-integrability ensures that the stochastic integral and therefore (1.1) makes sense. Moreover the condition (1.1) ensures that there is no magical withdrawal or infusion of money in the investment process.

The benefit of having a constant bond is that we can express a self-financing portfolio by an initial wealth x and an $\mathbb{R}^d$-valued predictable S-integrable process $H = ((H_t^i)_{t\in[0,T]})_{1\leqslant i\leqslant d}$ specifying only the amount of each *risky* asset held in the portfolio at time t, i.e. we are able to avoid the condition (1.1). We adopt this custom for the remainder of the paper.

Indeed, in this situation we can see the investor as only investing in the d risky assets according to the portfolio H, while either taking credit or depositing excess money on his bank account H^0. In mathematical terms this means, that we can choose $(H_t^0)_{t\in[0,T]}$ such that the condition (1.1) is satisfied. In other words, for $\tilde{H} = ((H_t^i)_{t\in[0,T]})_{0\leqslant i\leqslant d}$ we see that $(x, \tilde{H})$ is a self-financing portfolio according to Definition 1.1.

For notational convenience we will additionally assume that

$$(H.S)_t = (H.\tilde{S})_t$$

where $\tilde{S} = ((S_t^i)_{t\in[0,T]})_{1\leqslant i\leqslant d}$.

Using these conventions we can write the value process of a self-financing portfolio as a $\mathbb{R}^d$-dimensional stochastic integral

$$X_t = x + (H.S)_t \qquad \forall\ t \in [0, T]$$

It is common in order to rule out arbitrage-profits to limit the strategies in such a way that the value process does not go beyond a set negative value. This is done in the following definition.

Definition 1.2 (admissible) *A predictable, S-integrable process H is called admissible, if there exists a constant* $C \in \mathbb{R}_+$ *such that almost surely we have:*

$$(H.S)_t \geqslant -C$$

In the following we will only consider admissible strategies. This leads us to the next definition.

Definition 1.3

$$\mathcal{H} = \left\{H : H \textit{ is an } \mathbb{R}^d\textit{-valued predictable, S-integrable and admissible process}\right\}$$

is called the set of all trading strategies for a financial market S.

This completes our mathematical model. In the following chapter we will discuss some useful properties of this model that will be required for our analysis on the optimization problem.

Chapter 2

Prerequisites

In this chapter we discuss the necessary background material that is required to fully understand the crucial part of this thesis. It is divided into two main parts. The first part deals with basic and important results on asset pricing such as the fundamental theorem of asset pricing. While the second part introduces the Legendre Transform and various basic concepts of duality theory and convex analysis that are used in the following chapters. Each part is roughly divided into two sections, one summarizing definitions and the other collecting the important results. However, the material builds up chronological, so in order to understand everything, it is advisable to read top to bottom. Readers who are familiar with these topics can skip this chapter entirely and continue to the optimization problem.

2.1 Asset Pricing

We state and prove all results in this section for the general case when Ω is not necessarily finite.

2.1.1 Some Definitions

Definition 2.1 *We call the subspace* K *of* $L^0(\Omega, \mathcal{F}, \mathbf{P})$ *defined by*

$$K = \{(H.S)_T : H \in \mathcal{H}\}$$

the set of contingent claims attainable at price 0*. Where* $L^0(\Omega, \mathcal{F}, \mathbf{P})$ *is the space of measurable functions equipped with the topology induced by convergence in probability.*

We can simply shift the set K by the constant function $x\mathbf{1}$, to get the set of contingent claims attainable at price x. Hence we define $K_x := K + x\mathbf{1}$.

Definition 2.2 *We call the convex cone* C *in* $L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ *defined by*

$$C = \{g \in L^\infty(\Omega, \mathcal{F}, \mathbf{P}) : \exists f \in K \text{ with } f \geqslant g\}$$

the set of contingent claims super-replicable at price 0.

In a similar manner as before we can translate the set C by $x\mathbb{1}$, to get the set of contingent claims super-replicable at price x. Hence we define $C_x := C + x\mathbb{1}$.

Definition 2.3 *A financial market S satisfies the no-arbitrage condition (NA) if*

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

or, equivalently,

$$C \cap L^\infty_+(\Omega, \mathcal{F}, \mathbf{P}) = \{0\}$$

where 0 denotes the function identically equal to zero.

Definition 2.4 *A financial market S satisfies the no-free-lunch-with-vanishing-risk condition (NFLVR) if*

$$\overline{C} \cap L^\infty_+(\Omega, \mathcal{F}, \mathbf{P}) = \{0\}$$

where 0 denotes the function identically equal to zero and $\overline{C}$ *is the norm closure of* C *in* $L^\infty(\Omega, \mathcal{F}, \mathbf{P})$.

Definition 2.5 *A financial market S is called complete if every bounded contingent claim is attainable. In other words every* $f \in L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ *can be represented as*

$$f = x + (H.S)_T \quad \text{for } x \in \mathbb{R},\ H \in \mathcal{H}$$

Notice that completeness of the market means that $L^\infty(\Omega, \mathcal{F}, \mathbf{P}) = \bigcup_{x \in \mathbb{R}} C_x$. Later we see that it is possible to connect completeness to the martingale theory, which leads to some very useful simplifications when we start talking about utility maximization.

Definition 2.6 *A probability measure* $\mathbf{Q}$ *on* $(\Omega, \mathcal{F})$ *is called and equivalent martingale measure for S, if* $\mathbf{Q} \sim \mathbf{P}$ *and S is a martingale under* $\mathbf{Q}$. *The set of all equivalent martingale measures is denoted by* $\mathcal{M}^e(S)$.

Similarly a probability measure $\mathbf{Q}$ *on* $(\Omega, \mathcal{F})$ *is called an absolutely continuous martingale measure for S, if* $\mathbf{Q} << \mathbf{P}$ *and S is a martingale under* $\mathbf{Q}$. *The set of all absolutely continuous martingale measures is denoted by* $\mathcal{M}^a(S)$.

2.1.2 Important Results

Lemma 2.7 *For a probability measure* $\mathbf{Q}$ *on* $(\Omega, \mathcal{F})$ *the following are equivalent:*

(i) $\mathbf{Q} \in \mathcal{M}^a(S)$

(ii) $\mathbb{E}_{\mathbf{Q}}(f) = 0 \qquad \forall f \in K$

(iii) $\mathbb{E}_{\mathbf{Q}}(g) \leqslant 0 \qquad \forall g \in C$

Proof (i) $\Leftrightarrow$ (ii)

This is just an application of the definition of a martingale. We start by assuming (i). Then $f \in K$ implies that there exists $H \in \mathcal{H}$ s.t. $f = (H.S)_T$. Now as S is a martingale under Q it follows that $(H.S)_t$ is also a martingale under Q which implies:

$$\mathbb{E}_{\mathbf{Q}}(f) = \mathbb{E}_{\mathbf{Q}}((H.S)_T)) = \mathbb{E}_{\mathbf{Q}}((H.S)_0)) = 0$$

For the reverse we need to show S^i is a martingale with respect to $\mathbf{Q}$. Consider for $s < t$ and $B \in \mathcal{F}_s$ the following admissible trading strategy:

$$H_u^i = \begin{cases} 0 & \text{if } 0 \leqslant u < s \\ \mathbb{1}_B & \text{if } s \leqslant u < t \\ 0 & \text{if } t \leqslant u \leqslant T \end{cases}$$

and for $j \neq i$

$$H_u^j = 0 \qquad \forall u \in [0, T]$$

Now set $f = (H.S)_T = (S_t^i - S_s^i)\mathbb{1}_B \in K$. By assumption we get:

$$\mathbb{E}_{\mathbf{Q}}((S_t^i - S_s^i)\mathbb{1}_B) = 0$$

which implies that S^i is a martingale and hence S is a martingale with respect to $\mathbf{Q}$. To show that $\mathbf{Q} << \mathbf{P}$ notice that for every $A \in \mathcal{F}$ with $\mathbf{P}(A) = 0$ we have $\mathbb{1}_A = 0$ $\mathbf{P}$–a.s. and thus we have $f = \mathbb{1}_A \in K$ which leads to $\mathbf{Q}(A) = \mathbb{E}_{\mathbf{Q}}(\mathbb{1}_A) = 0$.

(ii) $\Leftrightarrow$ (iii)

Assume (ii) holds. $g \in C$ implies that there exists an $f \in K$ s.t. $g \leqslant f$. Which leads to:

$$\mathbb{E}_{\mathbf{Q}}(g) \leqslant \mathbb{E}_{\mathbf{Q}}(f) = 0$$

On the other hand if (iii) holds let $f \in K$ then we know in particular that $f \in C$ and thus $\mathbb{E}_{\mathbf{Q}}(f) \leqslant 0$. Using that K is a linear space we get $\mathbb{E}_{\mathbf{Q}}(-f) \leqslant 0$ and hence (ii). □

Theorem 2.8 (Fundamental Theorem of Asset Pricing) *Let S be a bounded real valued semi-martingale. The following are equivalent:*

(i) S satisfies (NFLVR)

(ii) $\mathcal{M}^e(S) \neq \varnothing$

Proof The proof of (ii) $\Rightarrow$ (i) is quite simple while the other direction (i) $\Rightarrow$ (ii) is fairly technical and uses some non-trivial results from functional analysis which we will only state.

(ii) $\Rightarrow$ (i):

Let $\mathbf{Q} \in \mathcal{M}^e(S)$. Assume $\overline{C} \cap L^\infty_+(\Omega, \mathcal{F}, \mathbf{P}) \neq \{0\}$. This implies there exists a $g \in \overline{C} \cap L^\infty_+(\Omega, \mathcal{F}, \mathbf{P})$ which is not constant zero.

Now g is not constant zero and takes values in $\mathbb{R}_{\geqslant 0}$ $\mathbf{P}$–$a.s.$ and $\mathbf{Q} \sim \mathbf{P}$. Hence

$$\mathbb{E}_{\mathbf{Q}}(g) > 0 \tag{2.1}$$

We also know that there exists $(g_n)_{n\in\mathbb{N}} \subseteq C$ with $\|g_n - g\|_{L^\infty} \xrightarrow{n\to\infty} 0$

As a result Lemma 2.7 above gives us $\mathbb{E}_{\mathbf{Q}}(g_n) \leqslant 0$ for all $n \in \mathbb{N}$. In particular using the dominated convergence theorem we have:

$$\lim_{n\to\infty} \mathbb{E}_{\mathbf{Q}}(g_n) = \mathbb{E}_{\mathbf{Q}}(g) \leqslant 0 \tag{2.2}$$

(2.1) and (2.2) are a contradiction which means: $\overline{C} \cap L^\infty_+(\Omega, \mathcal{F}, \mathbf{P}) = \{0\}$

(i) $\Rightarrow$ (ii):

As mentioned above this is the hard part of the proof and it requires two technical results, which we state without proofs in the appendix. The first is the Kreps-Yan Separation Theorem (Thm A.4) and the second is a theorem that guarantees us that C is weak*-closed (Thm A.5).

Taking these results for granted we can now proof this implication. Assume S satisfies (NFLVR). Trivially this implies (NA) (i.e. $C \cap L^\infty_+ = \{0\}$) and by Thm A.5 it follows that C is weak* closed. Therefore we are able to apply the Kreps-Yan Seperation (Thm A.4) to ensure the existence of a random variable $L \in L^1$ such that L is $\mathbf{P}$–$a.s.$ strictly positive, and

$$\mathbb{E}_{\mathbf{P}}(LX) \leqslant 0 \qquad \forall X \in C$$

We now scale L if necessary such that $\mathbb{E}_{\mathbf{P}}(L) = 1$ and can then use L as a Radon-Nikodym derivative to define a new measure $\mathbf{Q}$ by $d\mathbf{Q} := Ld\mathbf{P}$ which satisfies $\mathbf{Q} << \mathbf{P}$. Furthermore we have

$$\mathbb{E}_{\mathbf{P}}(LX) = \mathbb{E}_{\mathbf{Q}}(X) \leqslant 0 \qquad \forall X \in C \tag{2.3}$$

(2.3) implies that we can apply Lemma 2.7 to get that $\mathbf{Q}$ is an absolutely continuous martingale measure. Then since L is strictly positive we also have $\mathbf{Q} \sim \mathbf{P}$. So in particular $\mathbf{Q} \in \mathcal{M}^e(S)$. □

Lemma 2.9 *Suppose that S satisfies (NFLVR). Then for an element* $g \in L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ *the following are equivalent:*

(i) $g \in C$

(ii) $\mathbb{E}_{\mathbf{Q}}(g) \leqslant 0 \qquad \forall \mathbf{Q} \in \mathcal{M}^a(S)$

Proof The proof consists of an application of the bipolar theorem which is stated in the appendix (Thm A.6) and a density argument using Lemma A.3. First, we remind ourselves of the definition of the polar cone D° of a set $D \subseteq X$:

$$D^\circ := \{g \in X^* : \langle g, f \rangle_{X^* \times X} \leqslant 0 \quad \forall f \in D\}$$

where X^* is the dual of X

Next let us recall the dual space of $L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ which should be known from functional analysis.

$$(L^\infty(\Omega, \mathcal{F}, \mathbf{P}))^* = ba(\Omega, \mathcal{F}, \mathbf{P})$$

where $ba(\Omega, \mathcal{F}, \mathbf{P})$ is the set of all bounded finitely additive measures which are absolutely continuous with respect to $\mathbf{P}$ with the total variation norm. This duality is given by the following isometric isomorphism

$$\begin{aligned} R : ba(\Omega, \mathcal{F}, \mathbf{P}) &\longrightarrow (L^\infty(\Omega, \mathcal{F}, \mathbf{P}))^* \\ \mathbf{Z} &\longmapsto \Phi_{\mathbf{Z}} \end{aligned}$$

where $\Phi_{\mathbf{Z}}(f) := \int_\Omega f(\omega)\mathbf{Z}(d\omega)$. Note that this integral refers to an integral with respect to a finitely additive measure. The corresponding inverse is simply

$$\begin{aligned} R^{-1} : (L^\infty(\Omega, \mathcal{F}, \mathbf{P}))^* &\longrightarrow ba(\Omega, \mathcal{F}, \mathbf{P}) \\ \Phi &\longmapsto \mathbf{Z}_\Phi \end{aligned}$$

where $\mathbf{Z}_\Phi(A) := \Phi(\mathbb{1}_A)$ for all $A \in \mathcal{F}$.

A proof of this correspondence can be found in any functional analysis book (e.g. [3]).

This is all the background we need from functional analysis, and we can now begin the proof. As a first step we define the following set

$$\mathcal{M}_{ba}(S) = \{\mathbf{Q} \in ba(\Omega, \mathcal{F}, \mathbf{P}) : \mathbf{Q}(\Omega) = 1 \text{ and } R(\mathbf{Q}) \in C^\circ\}$$

Lemma A.3 shows that this set is the weak* closure of $\mathcal{M}^a(S)$ in $ba(\Omega, \mathcal{F}, \mathbf{P})$ and therefore turns out to be a good way to deal with the subtleties about finitely additive measures versus countably additive measures.

The next step is to prove a slightly weaker statement, that we will then extend to the desired result by using the density given in Lemma A.3.

We now prove the following equivalence

$$g \in C \Leftrightarrow \mathbb{E}_{\mathbf{Q}}(g) \leqslant 0 \quad \forall \mathbf{Q} \in \mathcal{M}_{ba}(S) \tag{2.4}$$

To do this we use the cone of $\mathcal{M}_{ba}(S)$ which is defined as

$$\text{cone}\,\mathcal{M}_{ba}(S) = \{\mathbf{Z} \in ba : \exists \mathbf{Q} \in \mathcal{M}_{ba}(S) \text{ and } \lambda \geqslant 0 \text{ s.t. } \mathbf{Z} = \lambda \mathbf{Q}\}$$

We now wish to show the equality $C^\circ = R(\text{cone}\,\mathcal{M}_{ba}(S))$.

$\underline{R(\text{cone}\,\mathcal{M}_{ba}(S)) \subseteq C^\circ}$:

Let $\mathbf{Z} = \lambda \mathbf{Q}$ with $\mathbf{Q} \in \mathcal{M}_{ba}(S)$ and $\lambda \geqslant 0$. Then for $f \in C$ we have

$$R(\mathbf{Z})(f) = \Phi_{\mathbf{Z}}(f) = \lambda \mathbb{E}_{\mathbf{Q}}(f) \leqslant 0$$

This however implies that $R(\mathbf{Z}) \in C^\circ$.

$\underline{C^\circ \subseteq R(\text{cone}\,\mathcal{M}_{ba}(S))}$:

Let $\Phi \in C^\circ$ be non-zero. Then consider the bounded finitely additive measure $\mathbf{Z} = R^{-1}(\Phi)$. Notice that $\mathbf{Z}$ is in particular positive. Indeed, let $A \in \mathcal{F}$ then $-\mathbb{1}_A \in C \subseteq L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ which by the definition of C° implies that $\mathbf{Z}(A) = \Phi(\mathbb{1}_A) \geqslant 0$. So we have shown that $\mathbf{Z}$ is positive. Now since $\mathbf{Z}$ is a non-zero positive bounded and finitely additive measure, there exists $\mathbf{Q} \in \mathcal{M}_{ba}(S)$ and $\lambda \geqslant 0$ such that $\mathbf{Z} = \lambda \mathbf{Q}$. So in particular $\Phi \in R(\text{cone}\,\mathcal{M}_{ba}(S))$.

Therefore we have shown that

$$C^\circ = R(\text{cone}\,\mathcal{M}_{ba}(S))$$

Now since S satisfies the NFLVR condition we can apply Theorem A.5 to get that C is weak* closed and use this to apply the Bipolar Theorem A.6 to get

$$C = [R(\text{cone}\,\mathcal{M}_{ba}(S))]^\circ \tag{2.5}$$

To complete the proof of 2.4 notice that the direction $\Rightarrow$ is trivial. For the reverse direction let $g \in L^\infty$ such that

$$\mathbb{E}_{\mathbf{Q}}(g) \leqslant 0 \quad \forall \mathbf{Q} \in \mathcal{M}_{ba}(S)$$

Now let $\lambda \geqslant 0$ and $\mathbf{Z} = \lambda \mathbf{Q}$ then

$$i(g)(R(\mathbf{Z})) = R(\mathbf{Z})(g) = \lambda \mathbb{E}_{\mathbf{Q}}(g) \leqslant 0$$

where i is the canonical embedding into the bidual. This implies $g \in [R(\text{cone}\,\mathcal{M}_{ba}(S))]^\circ = C$ by (2.5) and we are done.

Next we prove the full lemma. For this we will be using the density given in Lemma A.3. First notice that the direction $[(i) \Rightarrow (ii)]$ is a trivial consequence of the above result. For the reverse direction let $g \in L^\infty$ such that

$$\mathbb{E}_{\mathbf{Q}}(g) \leqslant 0 \quad \forall \mathbf{Q} \in \mathcal{M}^a(S) \tag{2.6}$$

Furthermore let $\mathbf{Q}^* \in \mathcal{M}_{ba}(S)$. Now by the density given in Lemma A.3 there exists a sequence $(\mathbf{Q}_n)_{n \geqslant 1} \subseteq \mathcal{M}^a(S)$ such that

$$\lim_{n\to\infty} \mathbb{E}_{\mathbf{Q}_n}(g) = \mathbb{E}_{\mathbf{Q}^*}(g)$$

Using (2.6) this leads to $\mathbb{E}_{\mathbf{Q}^*}(g) \leqslant 0$. So (2.4) gives us $g \in C$. This completes the proof. □

Theorem 2.10 (Super-Replication Theorem) *Suppose that S satisfies (NFLVR),* $g \in L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ *and* $x \in \mathbb{R}$. *Then the following are equivalent:*

(i) $\mathbb{E}_{\mathbf{Q}}(g) \leqslant x \qquad \forall \mathbf{Q} \in \mathcal{M}^a(S)$

(ii) $\exists H \in \mathcal{H}$ *s.t.* $g \leqslant x + (H.S)_T$

Proof In fact this is just Lemma 2.9. Simply replace the g in the lemma by $g - x$ and then notice that (ii) is the same as $g - x \in C$. □

Often the Super-Replication theorem is stated in a slightly different fashion. For completion we state it here as well.

Theorem 2.11 (Alternate Version of the Super-Replication Theorem) *Suppose that S satisfies (NFLVR) and* $g \in L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ *then the following equality holds*

$$\sup_{\mathbf{Q} \in \mathcal{M}^a(S)} \mathbb{E}_{\mathbf{Q}}(g) = \inf\{x : \exists h \in C \text{ s.t. } x + h \geqslant g\}$$

Theorem 2.12 (Complete Markets) *Let S be a financial market satisfying the No-Arbitrage (NA) condition. Then we have the following equivalence:*

(i) S is complete

(ii) $\mathcal{M}^e(S)$ *consists of a single element* $\mathbf{Q}$

Moreover if one of these hold and $f \in L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ *we have that*

$$\mathbb{E}_{\mathbf{Q}}(f \mid \mathcal{F}_t) = \mathbb{E}_{\mathbf{Q}}(f) + (H.S)_t \quad t = 0, \dots, T$$

and the stochastic integral (H.S) is unique.

Proof see [9] □

2.2 Conjugate Duality Theory

In this section, we introduce the Legendre-Transform and derive some of its properties that are important for the maximization of the utility function. Some of this may seem quite tedious, but it is necessary in order to limit the assumptions we need to make regarding the utility function later on.

2.2.1 Some Definitions

Convex Functions

There are many ways to define convex functions. We closely follow Rockafella's approach from his book *Convex Analysis* [5] and define convex functions in such a way that we are able to connect them with the theory about convex sets. Therefore we shortly recall the definition of a convex set.

Definition 2.13 (convex set) *A subset* C *of a vector space* X *is said to be convex if for all* $x, y \in C$ *and* $\lambda \in [0, 1]$ *we have that* $\lambda x + (1 - \lambda)y \in C$.

Now let us define the notion of epigraphs that allow us to connect functions to sets.

Definition 2.14 (epigraph) *Let* $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ *be a function. We define the epigraph of f as*

$$\mathrm{epi}(f) := \{(x, \mu) : x, \mu \in \mathbb{R} \text{ and } \mu \geqslant f(x)\}$$

We now have everything we need to define convex functions.

Definition 2.15 (convex/concave function) *Let* $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ *be a function.* f *is said to be convex if epi(f) is a convex set. The function* f *is said to be concave if* $-f$ *is convex.*

Notice that by only considering functions with domain $\mathbb{R}$ we do not actually restrict the theory, because we can always extend a convex (or concave) function f defined on a subset $D \subset \mathbb{R}$ by defining $f(x) = +\infty$ (or $f(x) = -\infty$) for all $x \in \mathbb{R} \setminus D$. In order to still be able to consider the finite parts of the functions, we introduce the effective domain of a function.

Definition 2.16 (effective domain) *Let* $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ *be a function. We define the (positive) effective domain of f as*

$$\mathrm{dom}^{+}(f) := \{x \in \mathbb{R} : f(x) < +\infty\}$$

and the (negative) effective domain of f as

$$\mathrm{dom}^{-}(f) := \{x \in \mathbb{R} : f(x) > -\infty\}$$

When we use the positive or the negative effective domain depends on the properties of f that we are interested in. In this paper we use dom^+ for convex functions and dom^- for concave functions, because this ensures that extending the function as mentioned above has no effect on the effective domains. The usefulness of this can be seen in the following definition of strict convexity (and strict concavity).

Definition 2.17 (strictly convex/concave) *Let* $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ *be a function.* f *is said to be strictly convex if for every convex set* $C \subseteq \mathrm{dom}^+(f)$ *we have*

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y) \qquad \forall \lambda \in (0,1) \text{ and } x, y \in C$$

f *is said to be strictly concave if for every convex set* $C \subseteq \mathrm{dom}^-(f)$ *we have*

$$f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)f(y) \qquad \forall \lambda \in (0,1) \text{ and } x, y \in C$$

As of right now we are still allowing some trivial functions to be part of the theory (e.g. $f \equiv +\infty$). Because these lead to some cumbersome exceptions, we make the following useful definition.

Definition 2.18 (proper) *Let* $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ *be a function. Then f is called proper if the following holds:*

- $f(x) > -\infty \qquad \forall x \in \mathbb{R}$
- $f(x) < +\infty \qquad$ *for some* $x \in \mathbb{R}$

Legendre-Transform

The Legendre-Transform is the tool we use to solve maximization and minimization problems. Depending on the assumptions made on the differentiability of f there are different ways of defining the Legendre-Transform we use the most general version that requires no assumptions on f.

Definition 2.19 (Legendre Transformation) *Let* $f : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ *be a function. Then we define the Legendre Transformation* f^* *of* f *as follows:*

$$f^*(y) := \sup_{x \in \mathbb{R}} \{yx - f(x)\} \qquad \forall y \in \mathbb{R}$$

We will call f^* *the conjugate of* f.

We now start imposing some restrictions on f to derive some nice properties of the Transform. As usual the more restrictions we impose, the nicer the results we achieve.

2.2.2 Important Results

First, we see that even without imposing any restrictions on f the conjugate f^* still has the property that it is always closed and convex.

Proposition 2.20 f^* *is a closed convex function.*

Proof Define the affine functions $h^x(y) := xy - f(x)$. Notice that affine functions are in particular closed and convex which means $\mathrm{epi}(h^x)$ is a closed convex set for all $x \in \mathbb{R}$. We write f^* as follows

$$f^*(y) = \sup_{x\in\mathbb{R}}\{yx - f(x)\} = \sup_{x\in\mathbb{R}}\{h^x(y)\}$$

This implies

$$\mathrm{epi}(f^*) = \bigcap_{x\in\mathbb{R}} \mathrm{epi}(h^x)$$

The intersection of closed convex sets is again closed convex which means f^* is closed and convex. □

The most fundamental question one can ask about the Legendre Transform is when the duality relation

$$(f^*)^* = f \tag{2.7}$$

holds.

The theory of Legendre-Transforms becomes especially useful when this duality is satisfied. It is therefore quite nice that we have the following simple theorem that gives us two assumptions on f such that (2.7) holds.

Theorem 2.21 *Let f be a closed convex function then*

$$f^{**} = f$$

Proof To prove this theorem we use the well known fact that a closed convex function f can be written as the pointwise supremum of the collection of all affine functions h satisfying $h \leqslant f$. A proof of this can be found in [5].

So

$$f(x) = \sup\{h(x) = yx - \mu : h \leqslant f\} \qquad \forall x \in \mathbb{R}$$

This leads to the following calculation:

$$\begin{aligned} & yx - \mu \leqslant f(x) \quad \forall x \in \mathbb{R} \\ \Leftrightarrow\ & yx - f(x) \leqslant \mu \quad \forall x \in \mathbb{R} \\ \Leftrightarrow\ & f^*(y) \leqslant \mu \\ \Leftrightarrow\ & (y, \mu) \in \mathrm{epi}(f^*) \end{aligned}$$

Hence we have shown:

$$f(x) = \sup_{(y,\mu)\in epi(f^*)} \{yx - \mu\} \qquad \forall x \in \mathbb{R}$$

Using the simple fact that for $(y, \mu) \in epi(f^*)$ we have $yx - \mu \leqslant yx - f^*(x)$ we get the final result:

$$f(x) = \sup_{y\in\mathbb{R}}\{yx - f^*(x)\} = f^{**}(x) \qquad \forall x \in \mathbb{R} \qquad \square$$

As the biconjugate f^{**} of f is just the conjugate of f^* Proposition 2.20 implies that f^{**} is convex and closed which shows that these are in fact the minimal conditions on f such that (2.7) holds. For the rest of this chapter we restrict ourselves to closed convex functions.

We now turn to the differentiability properties of conjugate functions. In order to have minimal restrictions on f this turns out to be a little bit of work. First, off we need to define a less restrictive version of differentiability that takes into account that we allow our functions to be infinite valued at some points. The right notion for this is called essentially differentiable and is defined as follows:

Definition 2.22 (essentially differentiable) *A proper convex function is essentially differentiable if the following holds:*

- $int(dom^+(f)) \neq \varnothing$
- f *is differentiable on* $int(dom^+(f))$
- $\lim_{n\to\infty} |f'(x_n)| = +\infty \qquad \forall x_n \in int(dom^+(f)) \xrightarrow{n\to\infty} x \in \partial(dom^+(f))$

Next we introduce the subdifferential, which can be defined for any convex function and is a generalization of the derivative. This turns out to be the right tool to analyse the differentiability properties of convex functions.

Definition 2.23 (subgradient and subdifferential) *Let f be a convex function and* $x \in \mathbb{R}$ *then the subdifferential of* f *at the point* x *is the set defined as*

$$\partial f(x) := \{y \in \mathbb{R} : f(z) \geqslant f(x) + y(z - x) \qquad \forall z \in \mathbb{R}\}$$

Every element $y \in \partial f(x)$ *is called subgradient of f at the point x.*

The inequality in the definition is referred to as the subgradient inequality. To make sense of it notice that $h(z) := f(x) + y(z - x)$ is an affine function through the point $f(x)$ that lies underneath the graph of f if and only if $y \in \partial f(x)$. So $\partial f(x)$ consists of all possible slopes such that the line with that slope through the point $f(x)$ is a tangent line. Hence it is intuitively clear that the subdifferentiable is an extension of the standard derivative.

It can be shown that the set $\partial f(x)$ is closed and convex. A useful way of looking at the subdifferential is to consider it as a multi-valued map that sends x to the set $\partial f(x)$. It turns out that if this is a single-valued mapping (i.e. $\partial f(x)$ is a singleton $\forall x \in \mathbb{R}$) it will under certain assumptions coincide with the ordinary differential. We now state the theorem that describes this connection.

Theorem 2.24 *Let f be a proper closed convex function. Then* ∂f *being a single-valued mapping in the sense that:*

- $\partial f(x) = f'(x) \quad \forall x \in \operatorname{int}(\operatorname{dom}^+(f))$
- $\partial f(x) = \varnothing \quad \forall x \notin \operatorname{int}(\operatorname{dom}^+(f))$

is equivalent to f being essentially differentiable.

The proof can be found in [5].

Using this characterization of the subgradient we can prove the following simple but useful Proposition.

Proposition 2.25 *Let f be a strictly convex essentially differentiable function then f is continuously differentiable on* $\operatorname{int}(\operatorname{dom}^+(f))$*.*

Proof Assume $x_0 \in \operatorname{int}(\operatorname{dom}^+(f))$ is a discontinuity point of f'. Since f is strictly convex we in particular know that f' is increasing so we have:

$$\lim_{\varepsilon \to 0} f'(x_0 + \varepsilon) > f'(x_0) \tag{2.8}$$

Using the subgradient inequality we get the following two inequalities:

(a) $f(z) \geqslant f(x_0) + f'(x_0)(z - x_0) \qquad \forall z \in \mathbb{R}$

(b) $f(z+\varepsilon) \geqslant f(x_0+\varepsilon) + f'(x_0+\varepsilon)(z - x_0) \qquad \forall z \in \mathbb{R},\ x_0 + \varepsilon \in \operatorname{int}(\operatorname{dom}^+(f))$

Now we can simply subtract (a) from (b) to get for all $z \in \mathbb{R}$ and $x_0 + \varepsilon \in \operatorname{int}(\operatorname{dom}^+(f))$

$$f(z+\varepsilon) - f(z) \geqslant [f(x_0+\varepsilon) - f(x_0)] + \left[f'(x_0+\varepsilon) - f(x_0)\right](z - x_0)$$

Next let ε tend to zero and use the continuity of f to arrive at

$$0 \geqslant \left[\lim_{\varepsilon \to 0} f'(x_0+\varepsilon) - f'(x_0)\right](z - x_0) \qquad \forall z \in \mathbb{R}$$

Together with (2.8) and $z > x_0$ we have a contradiction. Therefore f' has to be continuous. □

The importance of the subdifferential in the theory of Legendre transforms stems from the following theorem.

Theorem 2.26 *Let f be a proper convex function. Then for* $x \in \mathbb{R}$ *the following are equivalent:*

(i) $y \in \partial f(x)$

(ii) $f^*(y) = yx - f(x)$

Proof First let $y \in \partial f(x)$. Then y satisfies the subgradient inequality:

$$\begin{aligned} f(z) &\geqslant f(x) + y(z-x) & \forall z \in \mathbb{R} \\ xy - f(x) &\geqslant yz - f(z) & \forall z \in \mathbb{R} \end{aligned}$$

This implies $f^*(y) = \sup_{z \in \mathbb{R}} \{zy - f(z)\} = xy - f(x)$.

For the opposite direction let $f^*(y) = xy - f(x)$. Using the definition of f^* this leads to the inequality $xy - f(x) \geqslant zy - f(z)$ for all $z \in \mathbb{R}$ which is exactly the subgradient inequality for y. This means $y \in \partial f(x)$. □

Theorem 2.27 *Let f be a proper closed convex function. Then the following holds:*

$$\partial f^* \equiv (\partial f)^{-1}$$

i.e. $x \in \partial f^*(y) \Leftrightarrow y \in \partial f(x)$

Proof Recall that Prop 2.20 tells us that f^* is a proper closed convex function. Let $x \in \partial f^*(y)$ then this is equivalent using Thm 2.26 and Thm 2.21 to $f(x) = xy - f^*(y)$. If we rearrange this to $f^*(y) = xy - f(x)$ and apply Thm 2.26 again we get that this is equivalent to $y \in \partial f(x)$. □

This theorem tells us in particular that if f and f^* are differentiable we have the very useful relationship $(f^*)'(y) = (f')^{-1}(y) \quad \forall y \in \operatorname{int}(\operatorname{dom}^+(f))$.

After having become slightly more familiar with the notion of the subdifferential, we are now ready to prove the main result of this section that gives us sufficient conditions on f such that the conjugate f^* is essentially differentiable.

Theorem 2.28 *Let* f *be a closed proper strictly convex function then its conjugate* f^* *is essentially differentiable.*

Proof To begin notice Thm 2.24 implies we only need to prove

$$f \text{ closed proper strictly convex } \Rightarrow \partial f^* \text{ is a single-valued map}$$

which using $\partial f^* \equiv (\partial f)^{-1}$ (Thm 2.27) is equivalent to

$$f \text{ closed proper strictly convex } \Rightarrow \partial f(x_1) \cap \partial f(x_2) = \varnothing \quad \forall x_1 \neq x_2$$

We do this by proving the negation. So start by assuming $y \in \partial f(x_1) \cap \partial f(x_2)$ for some $x_1 \neq x_2$. Applying Thm 2.26 this gives us

$$f^*(y) = x_1 y - f(x_1) \text{ and } f^*(y) = x_2 y - f(x_2) \tag{2.9}$$

Define $h(z) := yz - f^*(y)$ then by the definition of f^* we have

$$\begin{aligned} f^*(y) &\geqslant yz - f(z) && \forall z \in \mathbb{R} \\ f(z) &\geqslant h(z) && \forall z \in \mathbb{R} \end{aligned}$$

Now using (2.9) we get $h(x_1) = f(x_1)$ and $h(x_2) = f(x_2)$. We have therefore shown that the image of h is a line that lies below $\mathrm{epi}(f)$ and touches it at the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$. This implies that f is not strictly convex. □

It is interesting to note that if we tweak our notion of strict convexity to make it slightly weaker, this theorem can be extended to a necessary and sufficient condition.

Chapter 3

Problem description

After having introduced some of the formal notation and basic results of asset pricing and conjugate duality theory, we can now formalize what it means to maximize the utility of the wealth of an economic agent investing in a financial market given a certain fixed initial endowment.

Firstly we want to restrict our market in such a way that it does not allow arbitrage profits. We choose to do this by assuming the no-free-lunch-with-vanishing-risk (NFLVR) condition on S which in the case when Ω is finite coincides with the no-arbitrage (NA) condition. Notice however that even in the general case the differences between (NA) and (NFLVR) are mainly of mathematical, rather than economical, nature. Applying the Fundamental Theorem of asset pricing (Thm 2.8) we see that the assumption that S satisfies (NFLVR) is equivalent with

Assumption 1 *The set $\mathcal{M}^e(S)$ is not empty.*

Next we introduce the function $U(x)$ that models the utility of an economic agents wealth x at the final time T. We will have to distinguish between the cases when negative wealth x is allowed and when it is not. This is done by assuming $U : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ with $\overline{\mathrm{dom}}(U) = (0, \infty)$ if negative wealth is disallowed and $\overline{\mathrm{dom}}(U) = \mathbb{R}$ if negative wealth is allowed. Additionally, we make the classical assumptions that U is increasing on $\mathbb{R}$, strictly concave, continuous on $\overline{\mathrm{dom}}(U)$ and differentiable on $\mathrm{int}(\overline{\mathrm{dom}}(U))$. Furthermore we need to make some restrictions on the marginal utility U'. We enforce that U' tends to zero when the wealth tends to infinity. Moreover U' has to tend to infinity if the wealth tends to its smallest allowed value. These assumptions all arise naturally in economics and thus do not restrict our theory. We now summarize all assumptions we have made on the utility function.

Assumption 2 *The utility function $U : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ satisfies*

(i) ***case 1:*** $\mathrm{dom}^{-}(U) = (0, \infty)$
case 2: $\mathrm{dom}^{-}(U) = \mathbb{R}$

(ii) *U is increasing on* $\mathrm{dom}^{-}(U)$

(iii) *U is continuous on* $\mathrm{dom}^{-}(U)$

(iv) *U is strictly concave*

(v) *U is differentiable on* $\mathrm{dom}^{-}(U)$

(vi) $U'(\infty) := \lim_{x \to \infty} U'(x) = 0$

(vii) ***case 1:*** $U'(0) := \lim_{x \downarrow 0} U'(x) = \infty$
case 2: $U'(-\infty) := \lim_{x \downarrow -\infty} U'(x) = \infty$

In the notation of chapter 2.2, (v) and (vii) together are equivalent to U being essentially differentiable.

Notice that as a consequence of these assumptions, we get for case 2 that

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

This does not hold true in case 1. Consider for example the utility function $U(x) = \frac{x^\alpha}{\alpha}$ for $0 < \alpha < 1$ and $x > 0$, which satisfies Assumption 2 (case 1), but $\lim_{x \downarrow 0} U(x) = 0$.

Given such a utility function we are ready to formalize the problem of maximizing utility. In abstract terms, with the notation introduced in the previous chapters, we can express the problem corresponding to the initial endowment $x \in \mathrm{dom}^{-}(U)$ as the maximization of

$$\mathbb{E}\left(U\left(x + (H.S)_T\right)\right)$$

over all 'admissible' trading strategies H (i.e. $H \in \mathcal{H}$ as defined in Chapter 1). In other words, we are trying to calculate the value function

$$u(x) := \sup_{H \in \mathcal{H}} \mathbb{E}\left(U\left(x + (H.S)_T\right)\right) \qquad \forall x \in \mathrm{dom}^{-}(U) \tag{3.1}$$

From this definition it is clear that u is increasing simply because U is increasing. Additionally we have $u(x) > -\infty$ by noticing that $u(x) \geqslant U(x)$ for $x \in \mathrm{dom}^{-}(U)$ (because $H \equiv 0 \in \mathcal{H}$).

Furthermore we would like to ensure that

$$u(x) < U(\infty) := \lim_{x \to \infty} U(x) \quad \text{for all } x \in \mathrm{dom}^{-}(U) \tag{3.2}$$

This can be achieved by making the following formally weaker assumption.

Assumption 3

$$u(x) < U(\infty) \quad \textit{for some } x \in \mathrm{dom}^{-}(U)$$

Remark 3.1
- *One can show that Assumption 3 implies* (3.2) *using basic properties of* u *and* U*. A full proof of this can be found in the appendix (Prop A.2).*

- *In the case where* $\mathrm{dom}^{-}(U) = (0, \infty)$ *the formally weaker assumption that* $u(x) < \infty$ *for some* $x \in \mathrm{dom}^{-}(U)$ *can be shown to be equivalent using Assumption 1 (see [8]).*

Chapter 4

The Finite Case

In this chapter we start our investigation into the maximization of the expected utility when Ω is finite. As previously mentioned, this assumption simplifies the setting tremendously. So first of all let us recall the important aspects of our model. We have an $\mathbb{R}^{d+1}$-valued discrete time price process $S = (S_t)_{t\in[0,1,\dots,T]}$ with $T \in \mathbb{N}$ and $S_t^0 \equiv 1$, which is adapted to the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, \mathbf{P})$. The proof that the price process degenerates to a discrete time process can be found in the appendix Prop A.1. Furthermore we will let $\Omega = \{\omega_1, \dots, \omega_N\}$ and can assume without loss of generality that $\mathcal{F}_0$ is trivial, that $\mathcal{F} = \mathcal{F}_T$ is the power set of Ω and that $\mathbf{P}[\omega_n] =: p_n > 0$ for all $1 \leqslant n \leqslant N$.

Assumption 1 from chapter 3 guarantees the existence of a measure $\mathbf{Q} \sim \mathbf{P}$ such that S is a $\mathbf{Q}$-martingale. Clearly we have that $\mathbf{Q}(\omega_n) =: q_n > 0$ holds for every such $\mathbf{Q}$.

4.1 Complete Market

We begin by analysing the case when the financial market is complete. As shown in Thm 2.12 in Chapter 2 this is equivalent to the set $\mathcal{M}^e(S)$, of equivalent probability measures under which S is a martingale, being reduced to a singleton. Let $\mathbf{Q}$ be this unique measure.

As it turns out, trying to solve the maximization problem (3.1) directly is hard because the maximizing variable H is contained in a stochastic integral. The trick here is to apply the Super-Replication Theorem (Thm 2.10) to transform the maximization problem to a problem we are more capable of solving. We do this in the following theorem.

Theorem 4.1 *For initial endowment* $x \in \operatorname{dom}^{-}(U)$ *the following maximization problems are equivalent*

(a) *maximization of*

$$\mathbb{E}_{\mathbf{P}}(U(x + (H.S)_T))$$

over all $H \in \mathcal{H}$

(b) *maximization of*

$$\mathbb{E}_{\mathbf{P}}(U(X_T)) = \sum_{n=1}^{N} p_n U(\xi_n) \tag{4.1}$$

over all $X_T \in C(x) := \{X_T \in L^0(\Omega, \mathcal{F}_T, \mathbf{P}) : \mathbb{E}_{\mathbf{Q}}(X_T) = \sum_{n=1}^{N} q_n \xi_n \leqslant x\}$

where we defined $\xi_n := X_T(\omega_n)$.

Proof Notice that since $(H.S)_T \in L^0(\Omega, \mathcal{F}_T, \mathbf{P})$ we know immediately that the maximization of (b) is greater than that of (a).

For the other direction Assumption 1 allows us to apply the Super-Replication Theorem (Thm 2.10) which gives for $X_T \in L^0(\Omega, \mathcal{F}_T, \mathbf{P}) = L^\infty(\Omega, \mathcal{F}_T, \mathbf{P})$ (since Ω is finite):

$$\mathbb{E}_{\mathbf{Q}}(X_T) \leqslant x \quad \forall \mathbf{Q} \in \mathcal{M}^a(S) \quad \Leftrightarrow \quad \exists H \in \mathcal{H} : \; X_T \leqslant x + (H.S)_T$$

This implies that the maximization of (a) is greater than that of (b) and therefore completes the proof. □

The problem written in the form (b) is now just a concave maximization problem in $\mathbb{R}^N$ with one linear constraint, which can be solved using the classical approach of Lagrange multipliers. The Lagrangian is given by

$$L(\xi_1, \dots, \xi_N, y) = \sum_{n=1}^{N} p_n U(\xi_n) - y \left(\sum_{n=1}^{N} q_n \xi_n - x \right) \tag{4.2}$$

$$= \sum_{n=1}^{N} p_n \left(U(\xi_n) - y \frac{q_n}{p_n} \xi_n \right) + yx \tag{4.3}$$

for $y \geqslant 0$. Next we define the following two functions:

$$\Phi(\xi_1, \dots, \xi_N) := \inf_{y>0} L(\xi_1, \dots, \xi_N, y) \quad \xi_n \in \operatorname{dom}^-(U) \tag{4.4}$$

$$\Psi(y) := \sup_{\xi_1, \dots, \xi_N} L(\xi_1, \dots, \xi_N, y) \quad y \geqslant 0 \tag{4.5}$$

In the following we will use the conjugate duality theory to determine Ψ and use this to find an explicit expression for $\inf_{y>0} \Psi$. Comparing this to $\sup_{\xi_1, \dots, \xi_N} \Phi$ will then lead us to an explicit expression for u.

Begin by making the observation that because of (4.3) the optimization problem in (4.5) splits up into N independent optimization problems over $\mathbb{R}$. In other words, for all $1 \leqslant n \leqslant N$ we are trying to find the $\xi_n \in \mathbb{R}$ that maximizes

$$U(\xi_n) - y\frac{q_n}{p_n}\xi_n$$

At this point we can apply our knowledge on the Legendre transform to see that we can express this optimization using the conjugate function

$$V(\eta) := (-U(-\,\cdot\,))^*(\eta) = \sup_{\xi\in\mathbb{R}}\{U(\xi) - \eta\xi\} \tag{4.6}$$

Using the theory we derived in chapter 2.2 we can derive some useful properties of V which we collect in the following proposition.

Proposition 4.2 *If* U *satisfies Assumption 2, then* $V : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ *has the following properties:*

(i) $\operatorname{dom}^+(V) = (0, +\infty)$

(ii) for all $\xi \in \mathbb{R}$*:*

$$V^*(\xi) = -U(-\xi)$$
$$U(\xi) = \inf_{\eta\in\mathbb{R}}\{V(\eta) + \eta\xi\}$$

(iii) V is continuously differentiable

(iv) $I(y) := -V'(y) = (U')^{-1}(y)$

(v) V is strictly convex

(vi) $V(0) := \lim_{x\downarrow 0} V(x) = U(\infty)$

case 1: $V(\infty) := \lim_{x\to\infty} V(x) = U(0)$

case 2: $V(\infty) := \lim_{x\to\infty} V(x) = \infty$

(vii) $V'(0) := \lim_{y\downarrow 0} V'(y) = -\infty$

case 1: $V'(\infty) := \lim_{y\to\infty} V'(y) = 0$

case 2: $V'(\infty) := \lim_{y\to\infty} V'(y) = \infty$

Proof (i) This follows directly from (4.6).

(ii) Since $-U(-\,\cdot\,)$ is closed and convex we can apply Thm 2.21 and get the result.

(iii) V is a conjugate function and hence by Prop 2.20 it is convex. Furthermore $-U(-\,\cdot\,)$ is strictly convex hence applying Thm 2.28 we know that V is essentially differentiable. Together with Prop 2.25 this implies that V is continuously differentiable.

(iv) This relationship follows from the differentiability of U and V and an application of Thm 2.27.

(v) Begin by noting that U is strictly concave which implies that U' is strictly decreasing. Now the inverse of a strictly decreasing function is again strictly decreasing, so using $-V'(y) = (U')^{-1}(y)$ we have that V' is strictly increasing which implies that V is strictly convex.

(vi) These limits are a consequence of (4.6). To determine them, consider V as the solution of the maximization of a strictly convex function and then determine the optimizers. Together with the properties of U this gives the result.

(vii) The limit as y goes to zero is just a side result of the essential differentiability of V while the limits as y goes to infinity can be calculated using the relation $-V'(y) = (U')^{-1}(y)$ proved above. □

We can use these properties to determine Ψ. So expressing (4.5) with V we get

$$\begin{aligned}\Psi(y) &= \sum_{n=1}^{N} p_n V\left(y\frac{q_n}{p_n}\right) + yx \\ &= \mathbb{E}_{\mathbf{P}}\left(V\left(y\frac{d\mathbf{Q}}{d\mathbf{P}}\right)\right) + yx\end{aligned}$$

Next define the

$$v(y) := \mathbb{E}_{\mathbf{P}}\left(V\left(y\frac{d\mathbf{Q}}{d\mathbf{P}}\right)\right) = \sum_{n=1}^{N} p_n V\left(y\frac{q_n}{p_n}\right)$$

which, as a linear combination of V, clearly inherits the same qualitative properties as V listed in Prop 4.2. Later on we see that v is in fact the dual function of the value function u.

Prop 4.2 infers in particular that v' is continuous, that $\lim_{x\searrow 0} v'(x) = -\infty$ and that for case 1 $\lim_{x\to\infty} v'(x) = 0$ and for case 2 $\lim_{x\to\infty} v'(x) = \infty$. Therefore applying the intermediate value theorem it follows that in case 1 for every $x \in (0,\infty)$ and in case 2 for every $x \in \mathbb{R}$ there exists a $\hat{y}(x) > 0$ such that $v'(\hat{y}(x)) = -x$.

Hence for every $x \in \overline{\mathrm{dom}}(U)$ we have found a $\hat{y}(x)$ such that $\Psi'(\hat{y}(x)) = v'(\hat{y}(x)) + x = -x + x = 0$. Furthermore since Ψ is a linear combination of strictly convex functions it is in particular also strictly convex and therefore $\hat{y}(x)$ is the unique minimizer of Ψ.

$$\Psi(\hat{y}(x)) = \inf_{y>0} \Psi(y)$$

Next we fix $\hat{y}(x)$ and consider the function

$$\tilde{L}(\xi_1, \dots, \xi_N) := L(\xi_1, \dots, \xi_N, \hat{y}(x))$$

It is strictly concave so we can find the maximum by setting the derivative equal to zero.

$$\nabla\tilde{L}(\xi_1,\dots,\xi_N) = \begin{pmatrix} p_1\left(U'(\xi_1)-\hat{y}(x)\dfrac{q_1}{p_1}\right) \\ \vdots \\ p_N\left(U'(\xi_N)-\hat{y}(x)\dfrac{q_N}{p_N}\right)\end{pmatrix} \overset{!}{=} 0$$

We therefore know that the maximum is taken at the point $(\hat{\xi}_1,\dots,\hat{\xi}_N)$ satisfying for all $1 \leqslant n \leqslant N$

$$U'(\hat{\xi}_n) = \hat{y}(x)\frac{q_n}{p_n} \text{ or, equivalently, } \hat{\xi}_n = I\left(\hat{y}(x)\frac{q_n}{p_n}\right) \tag{4.7}$$

Thus we have shown

$$\begin{aligned}\inf_{y>0}\Psi(y) &= \inf_{y>0}(v(y)+xy) \\ &= v(\hat{y}(x)) + x\hat{y}(x) \\ &= L(\hat{\xi}_1,\dots\hat{\xi}_N,\hat{y}(x))\end{aligned} \tag{4.8}$$

Furthermore differentiating v explicitly at $\hat{y}(x)$ gives us

$$\begin{aligned} v'(\hat{y}(x)) = \sum_{n=1}^{N} q_n V'\left(\hat{y}(x)\frac{q_n}{p_n}\right) &\overset{!}{=} -x \\ \sum_{n=1}^{N} q_n I\left(\hat{y}(x)\frac{q_n}{p_n}\right) &= x \\ \sum_{n=1}^{N} q_n\hat{\xi}_n &= x \end{aligned} \tag{4.9}$$

Now combining (4.8) and (4.9) leads us to the result

$$\inf_{y>0}\Psi(y) = \sum_{n=1}^{N} p_n U(\hat{\xi}_n) \tag{4.10}$$

We can now connect this expression to u. To do this notice that using the form (4.2) of the Lagrangian we can write Φ as

$$\Phi(\xi_1,\dots,\xi_N) = \begin{cases} -\infty & \text{if } \sum_{n=1}^{N} q_n\xi_n > x \\ \sum_{n=1}^{N} p_n U(\xi_n) & \text{if } \sum_{n=1}^{N} q_n\xi_n \leqslant x \end{cases}$$

Hence it follows

$$u(x) = \sup_{\substack{\xi_1,\dots,\xi_N \\ \sum_{n=1}^{N} q_n\xi_n \leqslant x}} \sum_{n=1}^{N} p_n U(\xi_n) = \sup_{\xi_1,\dots,\xi_N} \Phi(\xi_1,\dots,\xi_N) \tag{4.11}$$

The final step is now to show that

$$\inf_{y>0} \Psi(y) = \sup_{\xi_1,\dots,\xi_N} \Phi(\xi_1,\dots,\xi_N) \tag{4.12}$$

We do this by showing both inequalities. Firstly the inequality $\inf_{y>0}\Psi(y) \geqslant \sup_{\xi_1,\dots,\xi_N} \Phi(\xi_1,\dots,\xi_N)$ follows easily using the properties of the supremum and the infimum. For the reverse inequality realize that if $(\xi_1,\dots,\xi_N)$ satisfies $\sum_{n=1}^{N} q_n\xi_n \leqslant x$ we have $\hat{y}(x)\left(\sum_{n=1}^{N} q_n\xi_n - x\right) \leqslant 0$ since $\hat{y}(x) > 0$. So together with (4.2) we get

$$\sum_{n=1}^{N} p_n U(\xi_n) \leqslant L(\xi_1,\dots,\xi_N,\hat{y}(x)) \leqslant L(\hat{\xi}_1,\dots,\hat{\xi}_N,\hat{y}(x))$$

which using (4.8) and (4.11) gives us the desired inequality. So we have proven (4.12).

Now we can put (4.10),(4.11) and (4.12) together and get

$$u(x) = \sum_{n=1}^{N} p_n U(\hat{\xi}_n) \tag{4.13}$$

Additionally (4.12) implies

$$\inf_{y>0}(v(y) + xy) = \inf_{y>0}\Psi(y) = \sup_{\xi_1,\dots,\xi_N} \Phi(\xi_1,\dots,\xi_N) = u(x) \tag{4.14}$$

So v and u satisfy the same duality as U and V. We sum up our findings in the following theorem.

Theorem 4.3 (finite Ω, complete market) *Let the financial market $S = (S_t)_{t=0}^{T}$ be defined over the finite filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^{T}, \mathbf{P})$ and satisfy $\mathcal{M}^e(S) = \{\mathbf{Q}\}$, and let the utility function U satisfy Assumption 2. Denote by $u(x)$ and $v(x)$ the value functions*

$$u(x) = \sup_{X_T \in C(x)} \mathbb{E}_{\mathbf{P}}(U(X_T)) \qquad x \in \operatorname{dom}^{-}(U) \tag{4.15}$$

$$v(y) = \mathbb{E}_{\mathbf{P}}\left(V\left(y\frac{d\mathbf{Q}}{d\mathbf{P}}\right)\right) \qquad y > 0 \tag{4.16}$$

We then have:

(i) *The value functions* $u(x)$ *and* $v(y)$ *are conjugate and* u *inherits the qualitative properties of* U *listed in Assumption 2.*

(ii) *The optimizer* $\hat{X}_T(x)$ *in* (4.15) *exists, is unique and satisfies*

$$\hat{X}_T(x) = I(y\tfrac{d\mathbf{Q}}{d\mathbf{P}}) \text{ or, equivalently, } y\tfrac{d\mathbf{Q}}{d\mathbf{P}} = U'(\hat{X}_T(x))$$

where $x \in \mathrm{dom}^{-}(U)$ *and* $y > 0$ *are related via* $u'(x) = y$ *or, equivalently,* $x = -v'(y)$.

(iii) *The following formulae for* u' *and* v' *hold true:*

$$u'(x) = \mathbb{E}_{\mathbf{P}}(U'(\hat{X}_T(x))) \qquad v'(y) = \mathbb{E}_{\mathbf{Q}}(V'(y\tfrac{d\mathbf{Q}}{d\mathbf{P}})) \quad (4.17)$$

$$xu'(x) = \mathbb{E}_{\mathbf{P}}(\hat{X}_T(x)U'(\hat{X}_T(x))) \qquad yv'(y) = \mathbb{E}_{\mathbf{P}}(y\tfrac{d\mathbf{Q}}{d\mathbf{P}}V'(y\tfrac{d\mathbf{Q}}{d\mathbf{P}})) \quad (4.18)$$

Proof We already proved part (ii) of the theorem.

(i):

Equation (4.14) implies

$$u(x) = \inf_{y>0}(v(y) + xy) = (-v(-\cdot))^*(x) \quad \forall x \in \mathrm{dom}^{-}(U)$$

Now v is by (4.16) just a linear combination of V so it inherits the qualitative properties of V listed in Prop 4.2. In particular, this means that we have the same dual relationship between u and v as between U and V. Hence it follows that u inherits the qualitative properties from U.

(iii):

Recall that

$$v(y) = \mathbb{E}_{\mathbf{P}}(V(y\tfrac{d\mathbf{Q}}{d\mathbf{P}})) = \sum_{n=1}^{N} p_n V(y\tfrac{q_n}{p_n})$$

so differentiating gives us

$$v'(y) = \sum_{n=1}^{N} q_n V'(y\tfrac{q_n}{p_n}) = \mathbb{E}_{\mathbf{Q}}(V'(y\tfrac{d\mathbf{Q}}{d\mathbf{P}}))$$

The expression for $yv'(y)$ is trivial. To get the expressions for u' we use (ii) to get the following equivalences

$$u'(x) = \mathbb{E}_{\mathbf{P}}(U'(\hat{X}_T(x))) \quad \Leftrightarrow \quad y = \mathbb{E}_{\mathbf{P}}(y\tfrac{d\mathbf{Q}}{d\mathbf{P}})$$

$$xu'(x) = \mathbb{E}_{\mathbf{P}}(\hat{X}_T(x)U'(\hat{X}_T(x))) \quad \Leftrightarrow \quad yv'(y) = \mathbb{E}_{\mathbf{P}}(y\tfrac{d\mathbf{Q}}{d\mathbf{P}}V'(y\tfrac{d\mathbf{Q}}{d\mathbf{P}}))$$

for which the first is trivially true and the second is just the identity for $yv'(y)$ and hence also true. □

Let us make some quick remarks on this theorem, to fully understand what we have actually shown. First, recall that both U and V are explicit functions and hence we can use them to calculate $v(y)$ using (4.16). Our theorem then gives easy formulae to calculate all other quantities we are interested in, e.g., $\widehat{X}_T(x)$ or $u(x)$. In particular the theorem gives us the existence of the optimizer $\widehat{X}_T(x)$ and allows us to explicitly calculate it. The super replication theorem (Thm 2.10) connects this with the existence of an admissible trading strategy H that can be used to invest in the financial market and maximizes expected utility.

4.2 Incomplete Market

We now turn to the case of an incomplete financial market. This means our previous assumption that $\mathcal{M}^e(S)$ is a singleton does not hold true anymore. However, Assumption 1 asserts that $\mathcal{M}^e(S)$ is non-empty. We adapt our strategy from the complete case and extend our argument to take into account that there are different possible equivalent martingale measures.

As in the previous section in Thm 4.1, we first translate the maximization problem 3.1 to eliminate the trading strategies H.

Theorem 4.4 *For initial endowment* $x \in \mathrm{dom}(U)$ *the following maximization problems are equivalent*

(a) maximization of

$$\mathbb{E}_{\mathbf{P}}(U(x + (H.S)_T))$$

over all $H \in \mathcal{H}$

(b) maximization of

$$\mathbb{E}_{\mathbf{P}}(U(X_T)) = \sum_{n=1}^{N} p_n U(\xi_n) \tag{4.19}$$

over all $X_T \in C(x) := \{X_T \in L^0(\Omega, \mathcal{F}_T, \mathbf{P}) : \mathbb{E}_{\mathbf{Q}}(X_T) \leqslant x \quad \forall \mathbf{Q} \in \mathcal{M}^e(S)\}$

where we defined $\xi_n := X_T(\omega_n)$.

Proof same as for Thm 4.1 □

The difference to the complete case is that it could be that $\mathcal{M}^e(S)$ contains infinitely many elements and we thus have an optimization problem with infinitely many linear constraints. Technically, this is not a problem and can still be solved using a general minimax theorem that applies to this condition. In fact, this is what has to be done in the case when Ω is not finite anymore. However, because of the nice structure of the problem we are able to reduce back to the case of finitely many linear constraints and then solve the problem as in the case of the complete financial market.

In order to do this we prove the following proposition that gives us the "shape" of $\mathcal{M}^a(S)$.

Proposition 4.5 $\mathcal{M}^a(S)$ *is a bounded, closed convex polytope in* $\mathbb{R}^N$

Proof We note that this statement is to be understood using the following embedding of $\mathcal{M}^a(S)$ into $\mathbb{R}^N$:

$$\begin{aligned} \mathcal{M}^a(S) &\hookrightarrow \mathbb{R}^N \\ \mathbf{Q} &\mapsto (\mathbf{Q}(\omega_1), \dots, \mathbf{Q}(\omega_N)) \end{aligned}$$

Let $q_i := \mathbf{Q}(\omega_i)$.

First, we define $D := \{x \in \mathbb{R}^N : x_i \geqslant 0,\ \sum_{i=1}^N x_i = 1\}$ and observe that D is a closed convex polytope and that it is just the set of all probability measures which are absolutely continuous with respect to **P**. Therefore we have

$$\mathcal{M}^a(S) \subseteq D$$

This gives us the obvious boundedness of $\mathcal{M}^a(S)$. Next, recall that P is called a closed convex polytope in $\mathbb{R}^N$ if it is the intersection of a finite number of closed half-spaces in $\mathbb{R}^N$. Note that this definition automatically implies that the intersection of two closed convex polytopes is again a closed convex polytope. Every half-space can be expressed as the solution of an inequality of the form

$$a_1 x_1 + \cdots + a_N x_N \leqslant b$$

In particular this implies that every solution to an equality of the form

$$a_1 x_1 + \cdots + a_N x_N = b$$

is just the intersection of two half-spaces and thus a polytope.

Now we determine how the restriction that $\mathcal{M}^a(S)$ only consists of martingale measures affects the "shape". To this end let $\mathbf{Q} \in \mathcal{M}^a(S)$. This implies that S is a **Q**-martingale or equivalently we have for all $0 \leqslant n \leqslant T$ and for all $A \in \mathcal{F}_n$ that

$$\begin{aligned} \mathbb{E}_{\mathbf{Q}}(S_{n+1}\mathbb{1}_A) &= \mathbb{E}_{\mathbf{Q}}(S_n\mathbb{1}_A) \\ \Leftrightarrow \sum_{i=1}^N q_i S_{n+1}(\omega_i)\mathbb{1}_A(\omega i) &= \sum_{i=1}^N q_i S_n(\omega_i)\mathbb{1}_A(\omega i) \end{aligned}$$

Defining $c_i^{n,A} := S_{n+1}(\omega_i) - S_n(\omega_i)\mathbb{1}_A(\omega i)$ and rearranging the above we get that for **Q** to be a martingale measure for S we need that for all $0 \leqslant n \leqslant T$ and for all $A \in \mathcal{F}_n$

$$\sum_{i=1}^N c_i^{n,A} q_i = 0$$

Since there are only finitely many combinations for $n \in \{0, 1, \ldots, T\}$ and $A \in \mathcal{F}_n$ we get that the set of all martingale measures for S is a closed convex polytope.

Using that

$$\mathcal{M}^a(S) = D \cap \{\text{set of martingale measures for } S\}$$

we have therefore shown that $\mathcal{M}^a(S)$ is a bounded closed convex polytope.□

Every bounded closed convex polytope can be written as the closed convex hull of its finitely many extreme points $\{\mathbf{Q}^1, \ldots, \mathbf{Q}^M\}$. So every $\mathbf{Q} \in \mathcal{M}^a(S)$ can be written as a convex combination of these extreme points and thus using the linearity of the expectation we get that the maximization problem (4.19) is equivalent to the concave optimization problem with finitely many linear constraints given by

$$\begin{cases} \mathbb{E}_{\mathbf{P}}(U(X_T)) = \sum_{n=1}^{N} p_n U(\xi_n) \to \max! \\ \mathbb{E}_{\mathbf{Q}^m}(X_T) = \sum_{n=1}^{N} q_n^m \xi_n \leqslant x, \quad \text{for } m = 1, \ldots, M \end{cases}$$

The corresponding Lagrangian is given by

$$L(\xi_1, \ldots, \xi_N, \eta_1, \ldots, \eta_M) = \sum_{n=1}^{N} p_n U(\xi_n) - \sum_{m=1}^{M} \eta_m \left(\sum_{n=1}^{N} q_n^m \xi_n - x \right) \tag{4.20}$$

$$= \sum_{n=1}^{N} p_n \left(U(\xi_n) - \sum_{m=1}^{M} \frac{\eta_m q_n^m}{p_n} \xi_n \right) + \sum_{m=1}^{M} \eta_m x \tag{4.21}$$

with $(\xi_1, \ldots, \xi_N) \in \operatorname{dom}^-(U)^N$ and $(\eta_1, \ldots, \eta_M) \in \mathbb{R}_+^M$.

We can simplify the Lagrangian by defining $y := \eta_1 + \cdots + \eta_M$ and $\mu_m := \frac{\eta_m}{y}$ for all $1 \leqslant m \leqslant M$ and then considering the convex combination

$$\mathbf{Q}^\mu = \sum_{m=1}^{M} \mu_m \mathbf{Q}^m$$

Now since $\{\mathbf{Q}^1, \ldots, \mathbf{Q}^M\}$ are the extreme points of $\mathcal{M}^a(S)$ we know that as $(\eta_1, \ldots, \eta_M)$ runs through $\mathbb{R}_+^M$ the pairs $(y, \mathbf{Q}^\mu)$ run through $\mathbb{R}_+ \times \mathcal{M}^a(S)$. Therefore the Lagrangian can be expressed as

$$L(\xi_1, \dots, \xi_N, y, \mathbf{Q}) = \mathbb{E}_{\mathbf{P}}(U(X_T)) - y\mathbb{E}_{\mathbf{Q}}(X_T - x) \tag{4.22}$$
$$= \sum_{n=1}^{N} p_n \left(U(\xi_n) - y\frac{q_n}{p_n}\xi_n \right) + yx \tag{4.23}$$

with $\xi_n \in \text{dom}^{-}(U)$, $y > 0$ and $\mathbf{Q} = (q_1, \dots, q_N) \in \mathcal{M}^a(S)$.

This is now the exact same Lagrangian as in the complete case with the exception that now $\mathbf{Q}$ additionally runs through the set $\mathcal{M}^a(S)$ instead of remaining fixed. Thus, similarly to the complete case we define

$$\Phi(\xi_1, \dots, \xi_N) := \inf_{y>0, \mathbf{Q}\in\mathcal{M}^a(S)} L(\xi_1, \dots, \xi_N, y, \mathbf{Q}) \quad \xi_n \in \text{dom}^{-}(U) \tag{4.24}$$
$$\Psi(y, \mathbf{Q}) := \sup_{\xi_1,\dots,\xi_N} L(\xi_1, \dots, \xi_N, y, \mathbf{Q}) \quad y \geqslant 0 \tag{4.25}$$

Analogous to the complete case we can apply conjugate duality theory and the dual function V to get

$$\Psi(y, \mathbf{Q}) = \sum_{n=1}^{N} p_n V\left(\frac{yq_n}{p_n}\right) + yx, \quad y > 0, \ \mathbf{Q} \in \mathcal{M}^a(S)$$

Now we minimize Ψ in two steps. We begin by fixing $y > 0$ and minimizing over $\mathcal{M}^a(S)$, which means we wish to determine

$$\inf_{\mathbf{Q}\in\mathcal{M}^a(S)} \Psi(y, \mathbf{Q}), \quad y > 0$$

To do this consider for fixed $y > 0$ the function $(q_1, \dots, q_N) \mapsto \Psi(y, (q_1, \dots, q_N))$. Since it is a linear combination of the function V it is in particular continuous and strictly convex. Furthermore Proposition 4.5 implies that $\mathcal{M}^a(S)$ is compact. Consequently $\mathbf{Q} \mapsto \Psi(y, \mathbf{Q})$ attains a unique minimum on $\mathcal{M}^a(S)$, set $\hat{\mathbf{Q}}(y) = (\hat{q}_1, \dots, \hat{q}_N)$ for the minimizer. Moreover $\hat{q}_n > 0$ for each $n = 1, \dots, N$.

Indeed assume $\hat{q}_n = 0$ for some $1 \leqslant n \leqslant N$ and fix any $\mathbf{Q} \in \mathcal{M}^e(S)$. Defining $\mathbf{Q}^\varepsilon = \varepsilon\mathbf{Q} + (1-\varepsilon)\hat{\mathbf{Q}}$ we then have that $\mathbf{Q}^\varepsilon \in \mathcal{M}^e(S)$ for every $0 < \varepsilon < 1$. Using Proposition 4.2 we know that $V'(0) = -\infty$ and therefore $\Psi(y, \mathbf{Q}^\varepsilon) < \Psi(y, \hat{\mathbf{Q}})$ for $\varepsilon > 0$ small enough. This is a contradiction.

This implies that $\hat{\mathbf{Q}}(y)$ is in fact an equivalent martingale measure for S. In analogy to the complete case we define the dual value function ν by

$$\begin{aligned} v(y) &= \inf_{\mathbf{Q}\in\mathcal{M}^a(S)} \sum_{n=1}^{N} p_n V\left(y\frac{q_n}{p_n}\right) \\ &= \sum_{n=1}^{N} p_n V\left(y\frac{\hat{q}_n(y)}{p_n}\right) \end{aligned}$$

Since the minimizer $\hat{\mathbf{Q}}$ has a dependency on y we need to be slightly more careful when determining the derivative of v. We end up with the following result.

Lemma 4.6 *v is continuously differentiable and*

$$v'(y) = \mathbb{E}_{\hat{\mathbf{Q}}(y)}\left(V\left(y\frac{d\hat{\mathbf{Q}}(y)}{d\mathbf{P}}\right)\right) = \sum_{n=1}^{N} p_n V\left(y\frac{\hat{q}_n(y)}{p_n}\right)$$

Proof The fact that $\hat{\mathbf{Q}}$ is a minimizer gives us the inequalities:

(i) $\sum_{n=1}^{N} p_n V((y+h)\frac{\hat{q}_n(y+h)}{p_n}) \leqslant \sum_{n=1}^{N} p_n V((y+h)\frac{\hat{q}_n(y)}{p_n})$

(ii) $\sum_{n=1}^{N} p_n V((y)\frac{\hat{q}_n(y)}{p_n}) \leqslant \sum_{n=1}^{N} p_n V(y\frac{\hat{q}_n(y+h)}{p_n})$

These allow us to make the following calculations:

$$\begin{aligned} \limsup_{h\to 0} \frac{v(y+h)-v(y)}{h} &= \limsup_{h\to 0} \frac{1}{h}\sum_{n=1}^{N} p_n \left[V\left((y+h)\frac{\hat{q}_n(y+h)}{p_n}\right) - V\left(y\frac{\hat{q}_n(y)}{p_n}\right)\right] \\ &\leqslant \limsup_{h\to 0} \frac{1}{h}\sum_{n=1}^{N} p_n \left[V\left((y+h)\frac{\hat{q}_n(y)}{p_n}\right) - V\left(y\frac{\hat{q}_n(y)}{p_n}\right)\right] \\ &= \sum_{n=1}^{N} p_n \limsup_{h\to 0} \frac{1}{h}\left[V\left((y+h)\frac{\hat{q}_n(y)}{p_n}\right) - V\left(y\frac{\hat{q}_n(y)}{p_n}\right)\right] \\ &= \sum_{n=1}^{N} p_n \frac{\hat{q}_n(y)}{p_n} V'\left(y\frac{\hat{q}_n(y)}{p_n}\right) \\ &= \mathbb{E}_{\hat{\mathbf{Q}}(y)}\left(V\left(y\frac{d\hat{\mathbf{Q}}(y)}{d\mathbf{P}}\right)\right) \end{aligned}$$

where in the second line we used (i) and in the forth line we used that V is differentiable.

$$\begin{aligned}
\liminf_{h\to 0}\frac{v(y+h)-v(y)}{h} &= \liminf_{h\to 0}\frac{1}{h}\sum_{n=1}^{N} p_n\left[V\left((y+h)\frac{\hat{q}_n(y+h)}{p_n}\right)-V\left(y\frac{\hat{q}_n(y)}{p_n}\right)\right]\\
&\geqslant \liminf_{h\to 0}\frac{1}{h}\sum_{n=1}^{N} p_n\left[V\left((y+h)\frac{\hat{q}_n(y+h)}{p_n}\right)-V\left(y\frac{\hat{q}_n(y+h)}{p_n}\right)\right]\\
&= \sum_{n=1}^{N} p_n \liminf_{h,t\to 0}\frac{1}{h}\left[V\left((y+h)\frac{\hat{q}_n(y+t)}{p_n}\right)-V\left(y\frac{\hat{q}_n(y+t)}{p_n}\right)\right]\\
&= \sum_{n=1}^{N} p_n \frac{\hat{q}_n(y)}{p_n} V'\left(y\frac{\hat{q}_n(y)}{p_n}\right)\\
&= \mathbb{E}_{\hat{\mathbf{Q}}(y)}\left(V\left(y\frac{d\hat{\mathbf{Q}}(y)}{d\mathbf{P}}\right)\right)
\end{aligned}$$

where now in the second line we used (ii). The step from the third to the forth line holds true because both limits for t and h exist independently of each other and can therefore be interchanged.

The lemma follows immediately. □

Using this expression for v' and Proposition 4.2 one can get

- $\lim_{y\to 0} v'(y)) = -\infty$
- **case 1:** $\lim_{y\to\infty} v'(y)) = 0$
  case 2: $\lim_{y\to\infty} v'(y)) = \infty$

Now we are in good shape, because we are in the exact same position as in the complete case. This means using the same arguments as for the complete case we can define $\hat{y}(x)$ by $v'(\hat{y}(x)) = -x$ and $\hat{\xi}_n = I(\hat{y}(x)\frac{\hat{q}_n(\hat{y}(x))}{p_n})$ which leads to

$$\inf_{y>0,\mathbf{Q}\in\mathcal{M}^a(S)} \Psi(y,\mathbf{Q}) = L(\hat{\xi}_1,\ldots,\hat{\xi}_N,\hat{y}(x),\hat{\mathbf{Q}}(\hat{y}(x))) = \sup_{\xi_1,\ldots,\xi_N} \Phi(\xi_1,\ldots,\xi_N) = u(x)$$

So as a result we get that $(\hat{\xi}_1,\ldots,\hat{\xi}_N,\hat{y}(x),\hat{\mathbf{Q}}(\hat{y}(x)))$ is the unique saddle point of the Lagrangian (4.22) and that the value functions u and v are conjugate.

We collect the result in the following theorem.

Theorem 4.7 (finite Ω, incomplete market) *Let the financial market $S = (S_t)_{t=0}^T$ be defined over the finite filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, \mathbf{P})$ and satisfy Assumption 1, and let the utility function U satisfy Assumption 2. Denote by $u(x)$*

and $v(x)$ *the value functions*

$$u(x) = \sup_{X_T \in C(x)} \mathbb{E}_{\mathbf{P}}(U(X_T)) \qquad x \in \operatorname{dom}^-(U) \tag{4.26}$$

$$v(y) = \inf_{\mathbf{Q} \in \mathcal{M}^a(S)} \mathbb{E}_{\mathbf{P}}(V(y \tfrac{d\mathbf{Q}}{d\mathbf{P}})) \qquad y > 0 \tag{4.27}$$

We then have:

(i) The value functions $u(x)$ *and* $v(y)$ *are conjugate and* u *inherits the qualitative properties of* U *listed in Assumption 2.*

(ii) The optimizers $\hat{X}_T(x)$ *in* (4.26) *and* $\hat{\mathbf{Q}}(y)$ *in* (4.27) *exist, are unique and satisfy*

$$\hat{X}_T(x) = I(y \tfrac{d\hat{\mathbf{Q}}(y)}{d\mathbf{P}}) \text{ or, equivalently, } y \tfrac{d\hat{\mathbf{Q}}(y)}{d\mathbf{P}} = U'(\hat{X}_T(x))$$

where $x \in \operatorname{dom}^-(U)$ *and* $y > 0$ *are related via* $u'(x) = y$ *or, equivalently,* $x = -v'(y)$.

(iii) The following formulae for u' *and* v' *hold true:*

$$u'(x) = \mathbb{E}_{\mathbf{P}}(U'(\hat{X}_T(x))) \qquad v'(y) = \mathbb{E}_{\hat{\mathbf{Q}}(y)}(V'(y \tfrac{d\hat{\mathbf{Q}}(y)}{d\mathbf{P}})) \tag{4.28}$$

$$x u'(x) = \mathbb{E}_{\mathbf{P}}(\hat{X}_T(x) U'(\hat{X}_T(x))) \quad y v'(y) = \mathbb{E}_{\mathbf{P}}(y \tfrac{d\hat{\mathbf{Q}}(y)}{d\mathbf{P}} V'(y \tfrac{d\hat{\mathbf{Q}}(y)}{d\mathbf{P}})) \tag{4.29}$$

4.3 Discussion

After having proved the two main theorems (Thm 4.3 and Thm 4.7) in full detail, we now want to take a step back and understand the results and recall the main ideas in the proofs.

Let us emphasize the dual character of the theorem. We began with the utility function U and its corresponding value function u. The conjugate function V of U was then used to express the maximization in Ψ explicitly and thus led us to define the value function v. Using the saddle point of the Lagrange function, it then turned out that the same duality also holds between the value functions u and v. The important message one needs to understand is that instead of calculating u we can now calculate v, which can be explicitly done, and then use the duality to switch back to u. The duality furthermore carries over to the optimizers $\hat{X}_T(x)$ of u and $\hat{\mathbf{Q}}(y)$ of v, which again gives an easy way to switch back and forth between the two.

Next we turn to the important aspects of the proof. Recall, that after a first reduction, we were trying to determine the optimizers of the maximization problem (4.19) (resp. (4.1) in the complete case). We approached this by considering the Lagrangian and realizing that we are in the situation of minimaxing. To see this recall

$$L : C(x) \times (\mathbb{R}_+ \times \mathcal{M}^a(S)) \longrightarrow \mathbb{R}$$

Now when fixing the second coordinate $(y, \mathbf{Q})$ we wanted to maximize the concave function $L(\cdot, (y, \mathbf{Q}))$ and similar fixing the first coordinate X_T we wanted to minimize the convex function $L(X_T, (\cdot, \cdot))$. In essence, we were trying to find a saddle point of L. We then construct the saddle point explicitly and it is at this point, when the two functions Ψ and Φ came into play. The following equation illustrates the important interconnections.

saddle point condition

$$\inf_{y>0} (v(y) + yx) = \inf_{\substack{y>0 \\ \mathbf{Q} \in \mathcal{M}^a(S)}} \Psi(y, \mathbf{Q}) = L(\hat{\xi}_1, \dots, \hat{\xi}_N, \hat{y}(x), \hat{\mathbf{Q}}(\hat{y}(x))) = \sup_{\xi_1, \dots, \xi_N} \varphi(\xi_1, \dots, \xi_N) = u(x) \tag{4.30}$$

duality relation

In the next chapter we drop the condition that Ω is finite and discuss whether the theory we have developed here can be extended to the general setting.

Chapter 5

Outlook: General Setting

In the previous chapter we have solved the utility maximization problem for finite Ω. The next step is to drop this rather restrictive assumption. Unfortunately this implies that the price process S cannot be expressed as a discrete jump process anymore, and in particular, the stochastic integrals with respect to S do not simplify to sums.

The question now is under which conditions we can recover the essential features of Theorem 4.7 in the general setting?

It turns out that if we make two adjustments in the conditions we are in fact be able proof a very similar result. These two changes are:

- the sets in which $\hat{X}_T$ and $\hat{\mathbf{Q}}$ vary have to be extended
- an additional regularity condition on U is needed

In fact, there is a significant difference between the case where negative wealth is allowed and the case where it is disallowed. For the remainder of this paper we restrict ourselves to the case when negative wealth is not permitted so Assumption 2 case 1 with $\mathrm{dom}^{-}(U) = (0, \infty)$.

In this chapter we state and explain these required adjustments and then motivate the proof of a theorem similar to Theorem 4.7 for this general case.

5.1 Motivating Considerations

The proof for the theorem in the general setting can be done very similar to the case where Ω is finite. Nevertheless it becomes quite technical and lengthy, so in this section we try and emphasize the important considerations that have to be made in order to complete the proof. First, we note that similar to the finite case we can again reduce the problem of maximizing over all admissible investment strategies (3.1) to the problem of maximizing over all non-negative $\mathcal{F}_T$-measurable random variables X_T satisfying

$$\mathbb{E}_{\mathbf{Q}}(X_T) \leqslant x \quad \text{for all } \mathbf{Q} \in \mathcal{M}^a(S)$$

On the basis of the argument of the previous chapter we can fix $x > 0$, $y > 0$ and formally define the Lagrangian as

$$\begin{aligned} L^{x,y}(X_T, \mathbf{Q}) &= \mathbb{E}_{\mathbf{P}}(U(X_T)) - y[\mathbb{E}_{\mathbf{Q}}(X_T) - x] \\ &= \mathbb{E}_{\mathbf{P}}\left(U(X_T) - y\frac{d\mathbf{Q}}{d\mathbf{P}}X_T\right) + yx \end{aligned} \tag{5.1}$$

Furthermore, we define the set

$$C(x) := \{X_T \in L^0_+(\Omega, \mathcal{F}_T, \mathbf{P}) : \ \mathbb{E}_{\mathbf{Q}}(X_T) \leqslant x \text{ for all } \mathbf{Q} \in \mathcal{M}^a(S)\}$$

and take it as our candidate in which we let X_T vary when we minimax the Lagrangian in (5.1). Dually, we would like to have $\mathbf{Q}$ vary in $\mathcal{M}^a(S)$ which later turns out to be slightly too small.

As discussed at the end of the previous chapter, the main difficulty in the proof is to find the saddle point of the Lagrangian (see (4.30)). In the infinite dimensional setting this is done using a proper version of the minimax theorem. In the proof of Kramkov and Schachermayer (see [4]) the following version of the minimax theorem, for which a proof can be found in [10], is used.

Theorem 5.1 (Minimax Theorem) *Let C be a convex, compact subset of a locally convex space E and let D be a convex subset of a vector space F. Assume that $L : C \times D \mapsto \mathbb{R}$ satisfies the following conditions*

(i) $\xi \mapsto L(\xi, \mu)$ *is continuous and concave on* C *for all* $\mu \in D$

(ii) $\mu \mapsto L(\xi, \mu)$ *is convex on* D *for all* $\xi \in C$

Then:

$$\inf_{\mu \in D} \sup_{\xi \in C} L(\xi, \mu) = \sup_{\xi \in C} \inf_{\mu \in D} L(\xi, \mu)$$

The second important step in the proof is to show how to relate the Lagrangian (5.1) to the value functions u and v, which are defined similar as in the finite case. As can be seen in equation (4.30) this is needed in order to show the duality between the two value functions. The way in which this is done is proving a polar relation between the set in which X_T varies and the set in which $\mathbf{Q}$ varies. The significance of this will shortly become obvious, however we need to first make precise which C, D, E and F we want to use in the minimax theorem (Thm 5.1).

To make everything work we require a dual pairing between the spaces E and F. So what spaces should we choose E and F to be? $\mathcal{M}^a(S)$ can be naturally embedded into $L^1(\mathbf{P})$ by identifying $\mathbf{Q}$ with its Radon-Nikodym derivative $\frac{d\mathbf{Q}}{d\mathbf{P}}$. On the other hand the only L^p-space that $C(x)$ can be embedded into is $L^0(\mathbf{P})$; the space of measurable functions on Ω equipped with the topology induced by convergence in $\mathbf{P}$. This is however a problem since we have no duality between $L^1(\mathbf{P})$ and $L^0(\mathbf{P})$. The solution to this problem stems from the fact that both $C(x)$ and $\mathcal{M}^a(S)$ are in the positive orthant of $L^0(\mathbf{P})$ and $L^1(\mathbf{P})$ respectively. Both can therefore be embedded into $L^0_+(\mathbf{P})$ which we can turn into an "almost" Hilbert space using the following well-defined map:

$$\langle f, g\rangle := \mathbb{E}_{\mathbf{P}}(fg) \in [0,\infty]$$

This is possible because for non-negative functions the Lebesgue integral is well-defined if it is allowed to take the value ∞. We call this an "almost" Hilbert space because the bracket $\langle\cdot,\cdot\rangle$ has all the properties of a scalar product with the exception, that it may take values in $[0,\infty]$ and the map $(f,g) \mapsto \langle f, g\rangle$ is only lower semi-continuous and not continuous. This is however sufficient and gives us the dual pairing $\langle L^0_+(\mathbf{P}), L^0_+(\mathbf{P})\rangle$.

We now extend the set $\mathcal{M}^a(S)$ to a set $D(y)$ in a way that will give us a strong polar relation between the sets $C(x)$ and $D(y)$. This is done by defining D as the closed, convex hull of $\mathcal{M}^a(S)$ in $L^0_+(\mathbf{P})$. We can express D explicitly by the following proposition.

Proposition 5.2

$$D = \left\{ Y_T \in L^0_+(\Omega, \mathcal{F}_T, \mathbf{P}) : \exists (\mathbf{Q}_n)_{n=1}^{\infty} \subseteq \mathcal{M}^a(S) \text{ s.t. } Y_T \leqslant \lim_{n\to\infty} \frac{d\mathbf{Q}_n}{d\mathbf{P}} \right\}$$

where $\lim_{n\to\infty} \frac{d\mathbf{Q}}{d\mathbf{P}}$ *is understood in the sense of almost sure convergence.*

The proof of this proposition relies on the following lemma

Lemma 5.3 *Let A be a closed, convex bounded subset of $L^0_+(\Omega, \mathcal{F}_T, \mathbf{P})$. Then for each sequence $(h_n)_{n\geqslant 1} \subseteq A$ there exists a sequence of convex combinations $k_n \in conv(h_n, h_{n+1}, \dots)$ which converges almost surely to a function $k \in A$.*

We now define $D(y) := yD$ and are ready to state the polar relation between $C(x)$ and $D(y)$.

Proposition 5.4 *For $X_T, Y_T \in L^0_+(\Omega, \mathcal{F}_T, \mathbf{P})$ we have*

(i) $X_T \in C(x) \Leftrightarrow \mathbb{E}_{\mathbf{P}}(X_T Y_T) \leqslant xy$ *for all* $Y_T \in D(y)$

(ii) $Y_T \in D(y) \Leftrightarrow \mathbb{E}_{\mathbf{P}}(X_T Y_T) \leqslant xy$ *for all* $X_T \in C(x)$

So in particular $C^\circ = D$ and $D^\circ = C$ with respect to the dual pairing $\langle L^0_+(\mathbf{P}), L^0_+(\mathbf{P})\rangle$.

Proof Notice that we can assume w.l.o.g that $x = y = 1$ since $C(x) = xC(1)$ and $D(y) = yD(1)$. Furthermore, set $C := C(1)$.

(i)

$\underline{\Leftarrow}$: Let $X_T \in L^0_+(\mathbf{P})$ and $\mathbb{E}_{\mathbf{P}}(X_T Y_T) \leqslant 1$ for all $Y_T \in D$. Then since $\mathcal{M}^a(S) \subseteq D$ we know that

$$\mathbb{E}_{\mathbf{P}}(X_T \tfrac{d\mathbf{Q}}{d\mathbf{P}}) \leqslant 1 \text{ for all } \mathbf{Q} \in \mathcal{M}^a(S)$$

which is equivalent to

$$\mathbb{E}_{\mathbf{Q}}(X_T) \leqslant 1 \text{ for all } \mathbf{Q} \in \mathcal{M}^a(S)$$

Hence $X_T \in C$.

$\underline{\Rightarrow}$: Let $X_T \in C$. We immediately get

$$\mathbb{E}_{\mathbf{P}}(X_T \tfrac{d\mathbf{Q}}{d\mathbf{P}}) \leqslant 1 \text{ for all } \mathbf{Q} \in \mathcal{M}^a(S)$$

Now for some $Y_T \in D$ we know there exists $(\mathbf{Q}_n)_{n=1}^{\infty} \subseteq \mathcal{M}^a(S)$ such that $Y_T \leqslant \lim_{n\to\infty} \frac{d\mathbf{Q}_n}{d\mathbf{P}}$. So using Fatou's lemma we get the following:

$$\mathbb{E}_{\mathbf{P}}(X_T Y_T) \leqslant \mathbb{E}_{\mathbf{P}}\left(X_T \liminf_{n\to\infty} \tfrac{d\mathbf{Q}_n}{d\mathbf{P}}\right) \leqslant \liminf_{n\to\infty} \mathbb{E}_{\mathbf{P}}(X_T \tfrac{d\mathbf{Q}_n}{d\mathbf{P}}) \leqslant 1$$

(ii)

For this we make use of (i) and the bipolar theorem (Theorem A.6). In order to do this note the fact that for a closed convex set the polar cone and polar set are identical. Since D is closed and convex this implies that its polar cone with respect to the dual pairing $\langle L^0_+(\mathbf{P}), L^0_+(\mathbf{P})\rangle$ can be expressed as

$$D^\circ = \{X_T \in C : \mathbb{E}_{\mathbf{P}}(X_T Y_T) \leqslant 1 \text{ for all } Y_T \in D\}$$

Thus (i) implies $C = D^\circ$ and applying the bipolar theorem to D leads to

$$D^{\circ\circ} = D = C^\circ \qquad \square$$

which is simply (ii).

These polar relations can now be used to connect the Lagrangian and value functions v and u. Furthermore we are almost able to apply the minimax theorem (Theorem 5.1) to the sets $C(x), D(y) \subseteq L^0_+(\mathbf{P})$ which are in particular closed convex and bounded. The only missing ingredient is the compactness however it turns out that in general, neither $C(x)$ nor $D(y)$ are compact. Fortunately, we have Lemma 5.3 which gives us something similar to compactness that ends up being sufficient after making some additional arguments.

In order to localize the mini-maximizers and maxi-minimizer again requires some sort of a compactness argument which can again be done using lemma 5.3. However we need a further technical assumption on U to be able to prove the existence of the optimizer $\hat{X}_T$. A good such notion is called the reasonable asymptotic elasticity.

Definition 5.5 (reasonable asymptotic elasticity) *A utility function U satisfying the Assumption 2 (case 1) is said to have reasonable asymptotic elasticity if*

$$AE(U) = \limsup_{x\to\infty} \frac{xU'(x)}{U(x)} < 1$$

It was introduced by Kramkov and Schachermayer (see [4]) as a necessary and sufficient criterion on the utility function U to make the value function u inherit the qualitative properties of U and additionally ensure the existence of the optimizer $\hat{X}_T$ in this general setting.

5.2 Existence Theorem

We have now given some insight into how the ideas from the finite case can be extended to this general setting. Let us now state the general optimization theorem.

Theorem 5.6 (general Ω, incomplete market) *Let the bounded semi-martingale* $S = (S_t)_{t=0}^T$ *and the utility function* U *satisfy Assumption 1, Assumption 2 (case 1) and Assumption 3. In addition* U *has reasonable asymptotic elasticity. Denote by* $u(x)$ *and* $v(x)$ *the value functions*

$$u(x) = \sup_{X_T \in C(x)} \mathbb{E}_{\mathbf{P}}(U(X_T)) \qquad x > 0 \tag{5.2}$$

$$v(y) = \inf_{Y_T \in D(y)} \mathbb{E}_{\mathbf{P}}(V(Y_T)) \qquad y > 0 \tag{5.3}$$

Then we have:

(i) The value functions $u(x)$ *and* $v(y)$ *are conjugate; they are continuously differentiable, strictly concave (resp. convex) on* $(0, \infty)$ *and satisfy*

$$u'(0) = -v'(0) = \infty \quad u'(\infty) = v'(\infty) = 0$$

(ii) The optimizers $\hat{X}_T(x)$ *in* (5.2) *and* $\hat{Y}_T(y)$ *in* (5.3) *exist, are unique and satisfy*

$$\hat{X}_T(x) = I(\hat{Y}_T(y)) \text{ or, equivalently, } \hat{Y}_T(y) = U'(\hat{X}_T(x))$$

where $x > 0$, $y > 0$ *are related via* $u'(x) = y$ *or, equivalently,* $x = -v'(y)$.

(iii) We have the following relations between u′,v′ and $\hat{X}_T$, $\hat{Y}_T$ *respectively*

$$u'(x) = \mathbb{E}_{\mathbf{P}}\left(\frac{\hat{X}_T(x)U'(\hat{X}_T(x))}{x}\right) \qquad x > 0 \tag{5.4}$$

$$v'(y) = \mathbb{E}_{\mathbf{P}}\left(\frac{\hat{Y}_T(y)V'(\hat{Y}_T(y))}{y}\right) \qquad y > 0 \tag{5.5}$$

The full proof of this theorem can be found in [4]. The ideas are the ones from the previous section, however there are many technical difficulties that have to be dealt with.

Appendix A

Appendix

A.1 Technical Subsidiary Results

Proposition A.1 *Let* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbf{P})$ *a filtered probability space with* Ω *finite and let* $S = ((S^i_t)_{t\in[0,T]})_{0\leqslant i\leqslant N}$ *be an* $(\mathbb{R})^{d+1}$*-valued adapted martingale. Then the trajectory* $S_{.}(\omega)$ *will be constant except for finitely many jumps.*

Proof First notice: Ω finite implies $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ is also finite. This means that there are only finitely many sub-σ-algebras of $\mathcal{F}$.

For the filtration we therefore get that there exists a division of [0,T] into N intervals $(I)^N_{j=1}$ s.t.

- $I_j \cap I_i = \varnothing$
- $\bigcup\limits_{j=1}^{N} I_j = [0, T]$
- $i < j : \; x \in I_i, \; y \in I_j \;\; \Rightarrow \;\; x < y$

for which it holds that:

$$\forall 1 \leqslant j \leqslant N \text{ and } \forall t, r \in I_j : \; \mathcal{F}_t = \mathcal{F}_r$$

By the martingale property of S we know:

$$\forall \; r \leqslant t : \; \mathbb{E}(S_t \mid \mathcal{F}_r) = S_r$$

so for $r \leqslant t$ and $r, t \in I_j$ for some $1 \leqslant j \leqslant N$ then:

$$S_r = \mathbb{E}(S_t \mid \mathcal{F}_r) = \mathbb{E}(S_t \mid \mathcal{F}_t) = S_t$$

This implies $\forall 1 \leqslant j \leqslant N$ and $\forall\, r, t \in I_j$:

$$S_r = S_t$$

Which is exactly what we wanted to prove. □

Proposition A.2 *Let* U *be the utility function and* u *the corresponding value function as defined in Chapter 3. Then* u *is concave and if* $u(x) < U(\infty)$ *for some* $x \in \operatorname{dom}^-(U)$ *then we also have* $u(x) < U(\infty)$ *for all* $x \in \operatorname{dom}^-(U)$.

Proof We start by showing that u is concave. For that, notice that for every $\lambda \in [0,1]$ and $H_1, H_2 \in \mathcal{H}$, $H := \lambda H_1 + (1-\lambda)H_2$ is again an admissible trading strategy so $H \in \mathcal{H}$. Therefore using the convexity of U we get

$$\begin{aligned}
u(\lambda x + (1-\lambda)y) &= \sup_{H \in \mathcal{H}} \mathbb{E}(U(\lambda x + (1-\lambda)y + (H.S)_T)) \\
&= \sup_{H_1, H_2 \in \mathcal{H}} \mathbb{E}(U(\lambda(x + (H_1.S)_T) + (1-\lambda)(y + (H_2.S)_T))) \\
&\geqslant \sup_{H_1, H_2 \in \mathcal{H}} [\lambda \mathbb{E}(U(x + (H_1.S)_T)) + (1-\lambda)\mathbb{E}(U(y + (H_2.S)_T))] \\
&= \lambda u(x) + (1-\lambda)u(y)
\end{aligned}$$

Now assume $u(x) < U(\infty)$ for some $x \in \operatorname{dom}^-(U)$. Firstly note that $u(x) \leqslant U(\infty)$ for all $x \in \operatorname{dom}^-(U)$ holds trivially because U is increasing.

Let $y \in \operatorname{dom}^-(U)$. If $y \leqslant x$ then we get $u(y) \leqslant u(x) < U(\infty)$ using that u is increasing. Therefore it only remains to prove that if $y > x$ we still have $u(y) < U(\infty)$. Let $y > x$.

Now we distinguish two cases:

1. $U(\infty) = \infty$:

   Then by the concavity of u we have

   $$u(y) \leqslant u(x) + z(y - x) \qquad \forall z \in \partial(-u(x))$$

   Now $\partial(-u(x)) \neq \emptyset$ so we get that $u(y) < \infty$ and hence we are done.

2. $U(\infty) < \infty$:

   Assume for the sake of contradiction that

   $$u(y) = \sup_{H \in \mathcal{H}} \mathbb{E}(U(y + (H.S)_T)) = U(\infty)$$

   then there exists a sequence $(H_n)_{n \geqslant 1} \subseteq \mathcal{H}$ such that

   $$\lim_{n \to \infty} \mathbb{E}(U(y + (H_n.S)_T)) = U(\infty)$$

Using $U(\infty)$ as dominating function we can apply the dominated convergence theorem and get

$$\lim_{n\to\infty} \mathbb{E}(U(y + (H_n.S)_T)) = \mathbb{E}(\lim_{n\to\infty} U(y + (H_n.S)_T)) = U(\infty)$$

This implies $\lim_{n\to\infty} U(y + (H_n.S)_T) = U(\infty) \quad \mathbf{P}\text{–}a.s$ and using that U is continuous and increasing (by Assumption 2) we therefore have shown

$$\lim_{n\to\infty} (H_n.S)_T = \infty \quad \mathbf{P}\text{–}a.s$$

Again applying Assumption 2 we get

$$\lim_{n\to\infty} U'(x + (H_n.S)_T) = 0 \quad \mathbf{P}\text{–}a.s. \tag{A.1}$$

By the concavity of U (subgradient inequality) we have that

$$U(y + (H_n.S)_T) \leqslant U(x + (H_n.S)_T) + U'(x + (H_n.S)_T)(y - x)$$

After taking the expectation of both sides we can use (A.1) to calculate the limit as n goes to infinity.

$$U(\infty) \leqslant \lim_{n\to\infty} \mathbb{E}(U(x + (H_n.S)_T)) \leqslant u(x) \qquad \square$$

which is a contradiction to $u(x) < U(\infty)$.

Lemma A.3 $\mathcal{M}^a(S)$ *is* $\sigma(ba, L^\infty)$*- dense in* $\mathcal{M}_{ba}(S)$

Proof The proof will consist of an application of Goldstine's Theorem (Thm A.7). We begin by recalling that $ba(\Omega, \mathcal{F}, \mathbf{P})$ is the space of all bounded *finitely* additive measures which are absolutely continuous with respect to $\mathbf{P}$. Moreover the space $ca(\Omega, \mathcal{F}, \mathbf{P})$ is the space of all bounded *countably* additive measures which are absolutely continuous with respect to $\mathbf{P}$. Both spaces have the total variation norm. From functional analysis we have the following isometric isomorphic spaces

$$L^1(\Omega, \mathcal{F}, \mathbf{P}) \cong ca(\Omega, \mathcal{F}, \mathbf{P})$$
$$L^1(\Omega, \mathcal{F}, \mathbf{P})^* \cong L^\infty(\Omega, \mathcal{F}, \mathbf{P})$$
$$L^1(\Omega, \mathcal{F}, \mathbf{P})^{**} \cong ba(\Omega, \mathcal{F}, \mathbf{P})$$

If we apply Goldstine's Theorem (Thm A.7) to $X = L^1$ we get that $i(L^1)$ is $\sigma((L^1)^{**}, (L^1)^*))$- dense in $(L^1)^{**}$. Now using the isometrically isomorphic spaces stated above this implies that $ca(\Omega, \mathcal{F}, \mathbf{P})$ is $\sigma(ba, L^\infty)$- dense in $ba(\Omega, \mathcal{F}, \mathbf{P})$. This means that for every $\mathbf{Q}^* \in ba$ there exists a sequence $(\mathbf{Q}_n)_{n\geqslant 1} \subseteq ca$ such that

$$\lim_{n\to\infty} \mathbb{E}_{\mathbf{Q}_n}(g) = \mathbb{E}_{\mathbf{Q}^*}(g) \quad \forall g \in L^\infty$$

We are now ready to show the desired result. In order to do this let $\mathbf{Q}^* \in \mathcal{M}_{ba}(S)$. By the above considerations we know that there exists a sequence $(\mathbf{Q}_n)_{n\geqslant 1} \subseteq ca$ that converges to $\mathbf{Q}^*$ in the weak* topology. We now modify this sequence in such a way that it is in $\mathcal{M}^a(S)$.

Firstly, notice that since $\mathbf{Q}^*(\Omega) = 1$ we get that $\mathbf{Q}_n(\Omega) > 0$ for n large enough. So by normalizing the measures we can get that $\mathbf{Q}_n(\Omega) = 1$ for all n without losing the convergence. So in particular $\mathbf{Q}_n$ is a probability measure for every n. Furthermore since

$$\lim_{n\to\infty} \mathbb{E}_{\mathbf{Q}_n}(g) = \mathbb{E}_{\mathbf{Q}^*}(g) \leqslant 0 \quad \forall g \in C$$

we can adjust the sequence such that for every $n \geqslant 1$ we get $\mathbb{E}_{\mathbf{Q}_n}(g) \leqslant 0$ for all $g \in C$. At this point we can simply apply Lemma 2.7 to get that $\mathbf{Q}_n \in \mathcal{M}^a(S)$ for all $n \geqslant 1$. This completes the proof. □

A.2 Relevant Theorems

Theorem A.4 (Kreps-Yan Separation Theorem) *If C is weak* closed and if*

$$C \cap L^\infty_+ = \{0\}$$

then there exists a random variable $L \in L^1$ *s.t.* L *is* $\mathbf{P} - a.s.$ *strictly positive, and*

$$\mathbb{E}_{\mathbf{P}}(LX) \leqslant 0 \qquad \forall X \in C$$

Proof see [6] □

Theorem A.5 *If the asset price process S is uniformly bounded, then the condition NFLVR implies that C is weak* closed.*

Proof see [2] □

Theorem A.6 (Bipolar Theorem) *Let* $C \subseteq X$ *be a non-empty convex cone in the linear space X. Then the bipolar cone*

$$C^{\circ\circ} = \overline{C}$$

Proof In fact, this theorem can be seen as an application of our conjugate duality theory. The underlying idea is that for K convex cone we have

$$f \equiv \mathbb{1}_K \Rightarrow f^* \equiv \mathbb{1}_{K^\circ} \tag{A.2}$$

We do not prove this but it can be checked explicitly. Otherwise a proof can be found in [5].

Now since C is a non-empty convex cone it follows that $f \equiv \mathbb{1}_{\overline{C}}$ is a closed convex function and thus by Thm 2.21 we have that $f \equiv f^{**}$ which using (A.2) implies $\overline{C} = C^{\circ\circ}$. □

Theorem A.7 (Goldstine's Theorem) *Let* X *be a normed vector space. Then* $i(X)$ *is* $\sigma(X^{**}, X^*)$*- dense in* X^{**}*, where* i *is the canonical embedding into the bidual.*

Proof see [11] □

Theorem A.8 (Hahn-Banach Separation Theorem) *Let* X *be a normed vector space,* A *and* B *convex subsets of* X *such that* A *is compact and* B *is closed. Then there exists* $\Phi \in X^*$ *and* $\lambda \in \mathbb{R}$ *such that*

$$\Phi(a) > \lambda > \Phi(b) \quad \forall\, a \in A,\ b \in B$$

Bibliography

[1] Tomas Björk. *Arbitrage theory in continuous time*. Oxford University Press, 2009.

[2] Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300:463–520, 1994.

[3] Nelson Dunford and Jacob T. Schwartz. *Linear Operators: General Theory*. Interscience Publishers, Inc., New York, 1958.

[4] Dmitry Kramkov and Walter Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *The Annals of Applied Probability*, 9:904–950, 1999.

[5] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[6] Walter Schachermayer. Martingale measures for discrete-time processes with infinite horizon. *Mathematical Finance*, 4:25–55, 1994.

[7] Walter Schachermayer. Optimal investment in incomplete financial markets. In *Mathematical Finance: Bachelier Congress 2000*, pages 427–462. Springer, 2000.

[8] Walter Schachermayer. Optimal investment in incomplete markets when wealth may become negative. *The Annals of Applied Probability*, 11:694–734, 2001.

[9] Walter Schachermayer. Utility maximisation in incomplete markets. In *Stochastic Methods in Finance*, pages 255–293. Springer, 2004.

[10] Helmut Strasser. *Mathematical theory of statistics: statistical experiments and asymptotic decision theory*.

[11] Dirk Werner. *Funktionalanalysis*. Springer, 2011.