CONVEX DUALITY AND MATHEMATICAL FINANCE

by

KRISTINA ROGNLIEN DAHL

THESIS

for the degree of

MASTER OF SCIENCE IN MATHEMATICS

(Master i Matematikk)

Faculty of Mathematics and Natural Sciences
University of Oslo

May 2012

Det matematisk- naturvitenskaplige fakultet
University of Oslo

Abstract

The theme of this thesis is duality methods in mathematical finance. This is a hot topic in the field of mathematical finance, and there is currently a lot of research activity regarding this subject. However, since it is a fairly new field of study, a lot of the material available is technical and difficult to read. This thesis aims to connect the duality methods used in mathematical finance to the general theory of duality methods in optimization and convexity, and hence clarify the subject. This requires the use of stochastic, real and functional analysis, as well as measure and integration theory.

The thesis begins with a presentation of convexity and conjugate duality theory. Then, this theory is applied to convex risk measures. The financial market is introduced, and various duality methods, including linear programming duality, Lagrange duality and conjugate duality, are applied to solve utility maximization, pricing and arbitrage problems. This leads to both alternative proofs of known results, as well as some (to my knowledge) new results.

Contents

List of Figures

Preface

This is a thesis for the degree of Master of Science, written under the Mathematics program, specialization stochastic analysis, at the Department of Mathematics, University of Oslo. The work resulting in this thesis corresponds to 60 credits, and has been done at the Mathematics Department, Blindern, from spring 2011 to spring 2012. The work has been done independently, but with useful comments from my supervising professor, Professor Bernt Øksendal (Center of Mathematics for Applications, University of Oslo).

Acknowledgements: I would like to thank my supervisor, Bernt Øksendal for helpful comments, leading the way, and for having a clear direction and overview at times where I was confused by the tiniest details. I am very grateful to the University of Oslo, in particular the Department of Mathematics, for excellent professors and quiet study halls. I would also like to thank all the other master students in mathematics and applied mathematics at the University of Oslo for interesting mathematical discussions, lunches filled with laughter and lasting memories. I would like to thank my father, Geir Dahl, for interesting mathematical discussions. Finally, I would like to thank my fantastic boyfriend, Lars, and my wonderful family: mom, dad, my awesome brother Eirik, my grandmother and my grandfather, for their endless love and support. Tusen takk.

Chapter 1

Introduction

1.1 The theme of the thesis and some related work

The topic of this thesis is duality methods and their connection to mathematical finance. The term *duality method* is used in many different meanings in mathematical finance (see e.g. Kramkov and Schachermayer [22], Pennanen [29] and Rogers [37]). This thesis illustrates some of these meanings, and attempts to place the duality methods of mathematical finance into a more general convexity framework. This provides a better understanding of the methods, and connects mathematical finance to convexity theory. In order to connect duality methods and mathematical finance, theory from stochastic analysis, real analysis, and functional analysis is applied. Different perspectives of duality theory are considered, and applied to financial models. For example:

- Convex duality, also called conjugate duality (see for example Sections 2.5, 3.2 and 5.5, as well as Chapters 6 and 7). This is a field pioneered by Rockafellar [34], who is also a prominent researcher within the field of convex analysis (see [33]).
- Linear programming duality (see e.g. Sections 4.3, 6.1 and 6.3). This is a well-developed field with a strong duality theorem and efficient computational methods, such as the simplex algorithm and interior point methods. Pliska [32] provides a good introduction to this theory.
- Lagrange duality (see for example Sections 5.4, 5.5 and 5.6 as well as Chapter 6). This field is somewhat less known than linear programming (which is a special case), but also provides duality theorems (under certain assumptions). An introduction can be found in Bertsekas et al. [2].

That duality methods can be used to solve problems in mathematical finance has been known for some time. For example, Pliska [32] applies linear programming duality to show that there exits a linear pricing measure if and only if there are no dominant trading strategies. Variations of duality methods are used by Kartzas and Shreve [17] to derive solutions to utility maximization and pricing problems. However, it seems like convexity theory, such as conjugate duality theory, has not been utilized in this context until recently. Over the last ten years, several researchers have applied duality methods to problems in mathematical finance. Examples of such researchers, and papers, are Henclova [15], King [18], Kramkov and Schachermayer [22], [21], Pennanen [29], Pennanen and Perkkiö [30], and Schachermayer [40].

Since active use of duality methods in mathematical finance is a fairly new field, there are few books on the topic. This has made this thesis extra challenging, as I have worked mainly with research papers. On the other hand, it has been very interesting to be guided into such an active field, where there is a lot happening, and a lot to be done.

The theoretical background for the thesis is stochastic analysis, measure- and integration theory, convexity theory, real analysis, and functional analysis. Throughout the thesis, these mathematical fields are combined to study duality theory and its applications to mathematical finance.

1.2 My work and challenges

The work resulting in this thesis has been of the following form:

1. My supervisor, Professor Bernt Øksendal, suggested a topic to guide my work in a natural direction. Sometimes he would give me a relevant article to read for inspiration, otherwise I would find articles regarding the topic myself.

2. After reading the article(s) on the topic, I would fill in the missing details or look up definitions I did not understand. This often took a lot of time, as I am not used to reading research papers, which are at a much more advanced level than the text books I have read previously in my studies.

3. Finally, I would write my own presentation of the topic, based on the article(s) I had read. Sometimes (especially during the final semester of my work), I derived some new results inspired by what I had read.

I have enjoyed this kind of work. It has been challenging to work so independently, but I have learned a lot from it. Reading research material has also been difficult (as I had never done this prior to the work on my master thesis), but after a while it became easier (though still difficult!), and I now feel like I can find and read relevant research material on my own. Another obstacle was that the term "duality method" is used in many ways in mathematical finance (see e.g. Pennanen [29] and Rogers [37]), and therefore it has been difficult to see the underlying ideas of the methods applied. In order to understand the duality methods used in mathematical finance, I had to learn a lot of background material, such as convex analysis, conjugate duality, Lagrange duality and optimization theory. Most of this theory is at an advanced level; it is done in general infinite-dimensional vector spaces, and is not covered in any classes. Some infinite-dimensional convexity has been covered in our two courses in functional analysis, but not a lot. For example, conjugate duality and Lagrange duality was completely new to me. Also, there is not a lot of literature regarding infinite-dimensional convexity (exceptions are the Hahn-Banach theorems, see Pedersen [28], and the work of Rockafellar [34] and Ekeland [10]). Another challenge was that a lot of the required background background theory from stochastic analysis, measure- and integration theory and functional analysis is fairly advanced. Finally, I had to take the class in mathematical finance in the middle of writing my thesis. Hence, getting to know the field of financial mathematics also presented a lot of work.

1.3 My contributions

A main contribution of this thesis is a presentation of convex duality theory and its connections to several areas in mathematical finance. Moreover, I have clarified some new connections, proved a number of theorems and "missing results" and solved several examples illustrating the results. Also, towards the end of the thesis, I have proved some (to my knowledge) new results regrading pricing of claims under a general level of inside information, and I have also proved some known results in a new way using duality theory.

The following summarizes my main contributions:

- Chapter 2:
  - The proof of Theorem 2.41 (on $(L^{\perp})^{\perp} = \bar{L}$ for L a subspace) is by me.
- Chapter 3:
  - Theorem 3.3 (on how to make new convex risk measures) and its proof is by me.
- Chapter 4:
  - Lemma 4.4 (finding the dual problem of the arbitrage problem) and its proof is by me.
- Chapter 5:
  - In Section 5.5, the connection between Lagrange duality and Kramkov-Schachermayer's work has been done by me. Also, the use of the Slater condition is introduced by me (Schachermayer [40] does a direct argument using optimization in $\mathbb{R}^n$). Hence, the reasoning has been simplified and connected to a general theory.

  - In Section 5.6, the connection to Lagrange duality has been made by me, and hence the reasoning is simplified.
  - Sections 5.8 (a special case of Section 5.7) and 5.9 (on utility maximization under risk constraints) consist of my own ideas.
  - In Section 5.10, I have summarized some connections between duality methods and utility maximization (from my point of view).

- Chapter 6: This entire chapter consists of ideas by me. In particular:
  - Theorem 6.3 (on the price of a claim offered by a seller with general information: finite scenario space) is new (to my knowledge) and done by me. The proof uses Lagrange duality.
  - Lemma 6.4 (used to prove Theorem 6.5) is by me.
  - Theorem 6.5 (on the relationship between the prices offered by sellers with different information levels) is by me.
  - The example of Section 6.1 is by me.
  - Section 6.2 generalizes the results mentioned above to the case where the scenario space is arbitrary, using the conjugate duality theory of Rockafellar. In particular, Lemma 6.6, Lemma 6.11 Theorem 6.10, Lemma 6.12 and Theorem 6.13 are by me.
  - Section 6.3 concerns the pricing problem of a seller facing a short selling constraint on one of the assets. Theorem 6.14 summarizes the results derived in this section, and is by me.
  - Section 6.4 generalizes the results of Section 6.3 to arbitrary scenario space Ω. The results are based on my ideas and they are summarized in Theorem 6.15.
  - Also, Section 6.5, which considers the pricing problem of a seller facing a constraint on how much she may sell short or buy of a certain asset, is by me.

- Chapter 7: This chapter also consists only of my ideas. In particular:
  - Section 7.1 proves for an arbitrary scenario space and discrete time that there is no free lunch with vanishing risk if and only if there exists an equivalent martingale measure. This is shown via a generalized version of Lagrange duality (see Section 5.4). The proof is by me.
  - Section 7.2 displays a close connection between conjugate duality and the fundamental theorem of mathematical finance. The ideas of this section are by me.

There are also some minor contributions, in the sense that I have filled in missing details in proofs, written proofs briefly etc.

- Chapter 2:

  - The proof of Theorem 2.8 (on properties of convex sets) is written in brief form by me.
  - The proof of Theorem 2.14 (on equivalence of definitions of convex functions) is by me.
  - The proof of Theorem 2.23 (on properties of the indicator function of a set) is by me.
  - The proof of Theorem 2.24 (on properties of convex functions) is written in brief form by me.
  - Example 2.40 is by me.
  - In most examples, for instance in Example 2.45, some details have been filled in.

- Chapter 3:
  - Example 3.6 is by me.
  - The proof of Theorem 3.7 (on the connection between a convex risk measure and its acceptance set) is by me.
  - The idea of the proof of Theorem 3.8 (a dual representation of convex risk measures) is from Frittelli and Gianin [14], but the proof is written out in detail by me.
  - Some details of the proof of Theorem 3.9 (on a special dual representation of convex risk measures in $\mathbb{R}^n$) have been filled out.
  - Some details of the proof of Theorem 3.10 (an explicit form of the dual representation of convex risk measures in $\mathbb{R}^n$) have been filled out.
  - Some details of the proof of Theorem 3.11 (an explicit representation of convex risk measures in infinite dimension) have been filled out.

- Chapter 4:
  - Some details of the proof of Theorem 4.3 (theorem connecting no arbitrage and the existence of equivalent martingale measures, in finite dimension) have been filled out.

- Chapter 5:
  - In Section 5.2 the ideas are from Pham [31], but the market model has been altered slightly. The proofs in this section are based on the ideas by [31], but written out by me.
  - The example in Section 5.3 (illustrates a direct method for solving utility maximization problems, finite dimension) is by me.
  - The example in Section 5.5, which is continued from Section 5.3 (illustrates the duality method applied to the previous example) is by me.

- The proof of Lemma 5.10 (simplifies the set of equivalent martingale measures for the utility maximization problem) is by me.
- The proof of Lemma 5.11 (shows that the set of absolutely continuous martingale measures is a polytope in $\mathbb{R}^n$) is by me.
- The proof of Lemma 5.12 (reduces the set of absolutely continuous martingale measures to its extreme points, for the utility maximization problem) is by me.
- Some computations and arguments in the proof of Theorem 5.15 (a strong duality result for the utility maximization problem) has been filled in.

The theorems, proofs and sections that are mentioned in the previous lists are marked with $\diamondsuit$.

1.4 The structure of the thesis

This thesis consists of 8 chapters. Chapter 1 is this introduction. Chapter 2 begins with a summary of convexity theory, and then presents the conjugate duality theory of Rockafellar [34] with some examples. This chapter is background theory, which will be applied throughout the entire thesis, but in particular in Chapters 3, 5, 6, and 7. Chapter 3 introduces convex risk measures, and applies the convexity theory, as well as the conjugate duality theory of Chapter 2, to derive results regarding such measures. The main purpose of this chapter is to derive a dual representation theorem for convex risk measures, Theorem 3.8. This theorem is an example of how duality theory can be used in mathematical finance.

Chapter 4 introduces a model for the financial market, and shows that such a market can be modeled by a scenario tree whenever the scenario space is finite and the time is discrete. This model will be used in Chapters 5, 6, and 7. Chapter 5 introduces utility functions and the utility maximization problem for an agent in the financial market. The remainder of the chapter presents some work by Kramkov and Schachermayer, see [21], [22], and [40]. These papers are connected to general duality theory by using Lagrange duality and the Slater condition. Towards the end of the chapter, a special case of the results of [21] and [22] is shown using conjugate duality, and finally, a utility maximization problem with risk constraints is solved using Lagrange duality. Hence, duality theory can be used to solve utility maximization problems in mathematical finance.

Chapter 6 illustrates a another application of duality in mathematical finance, namely to pricing problems. This chapter considers various versions of the pricing problem of a seller of a contingent claim. In most of the chapter, the seller can have some general level of inside information, not just the information given by the prices of the assets in the market. Lagrange duality, linear programming duality, the Slater condition as well as conjugate duality is applied to derive dual problems to the pricing problem of the seller, and also to prove the absence of a duality gap.

Chapter 7 considers the connection between equivalent martingale measures, arbitrage in the financial market and the concept of "no free lunch with vanishing risk". Here, a slightly weaker version of the fundamental theorem of mathematical finance is proved using conjugate duality and a theorem by Pennanen and Perkkiö [30] to close the duality gap (this theorem requires an extra assumption, which is why the resulting theorem is "slightly weaker" than the fundamental theorem). This chapter shows that duality theory is useful in the study of arbitrage problems.

Finally, Chapter 8 gives some concluding remarks summarizing the methods and results of the thesis.

Chapter 2

Convexity, optimization, and convex duality

The purpose of this chapter is to cover some background theory in convexity, optimization, and convex duality needed to understand how duality methods are used in mathematical finance. Convexity is very important for this, see for instance Pennanen [29] and Karatzas and Shreve [17]. Therefore, this chapter will be applied throughout the rest of the thesis. The structure of this chapter is as follows: Section 2.1 recalls some basic notions of convexity theory, such as convex sets, convex functions and properties of these. Section 2.2 covers some of the most central theorems and ideas of optimization theory. Section 2.3 introduces the convex (conjugate) duality framework of Rockafellar [34]. Some examples of optimization using convex duality is given in Section 2.4. Section 2.5 introduces conjugate functions and proves the important Theorem 2.39. Finally, Section 2.6 introduces the Lagrange function of convex duality theory, and contains another important result, namely Theorem 2.44.

2.1 Basic convexity

This section summarizes some of the most important definitions and properties of convexity theory. The material of this section is mainly based on the presentation of convexity in Rockafellar's book [34], the book by Hiriart-Urruty and Lemarèchal [16], and the report by Dahl [4]. The last two consider $X = \mathbb{R}^n$, but the extension to a general inner product space is straightforward. Therefore, in the following, let X be a real inner product space, i.e. a vector space X equipped with an inner product $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$ (so the function $\langle \cdot, \cdot \rangle$ is symmetric, linear in the first component and positive definite in the sense that $\langle x, x \rangle \geq 0$ for all $x \in X$, with equality if and only if $x = 0$).

We begin with some core definitions.

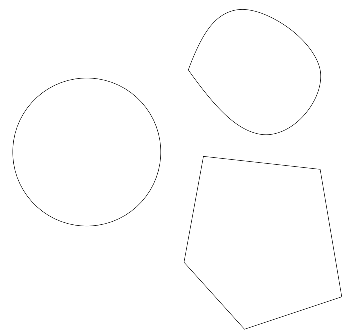

Figure 2.1: Some convex sets in the plane.

Definition 2.1 *(i) (Convex set) A set $C \subseteq X$ is called convex if $\lambda x_1 + (1-\lambda)x_2 \in C$ for all $x_1, x_2 \in C$ and $0 \leq \lambda \leq 1$.*

(ii) (Convex combination) A convex combination of elements $x_1, x_2, \ldots, x_k$ in X is an element of the form $\sum_{i=1}^k \lambda_i x_i$ where $\sum_{i=1}^k \lambda_i = 1$ and $\lambda_i \geq 0$ for all $i = 1, \ldots, k$.

(iii) (Convex hull, $\mathrm{conv}(\cdot)$) Let $A \subseteq X$ be a set. The convex hull of A, denoted $\mathrm{conv}(A)$ is the set of all convex combinations of elements of A.

(iv) (Extreme points) Let $C \subseteq X$ be a convex set. An extreme point of C is a point that cannot be written as a convex combination of any other points than itself. That is: $e \in C$ is an extreme point for C if $\lambda x + (1-\lambda)y = e$ for some $x, y \in C$ implies $x = y = e$.

(v) (Hyperplane) $H \subset X$ is called a hyperplane if it is of the form $H = \{x \in X : \langle a, x \rangle = \alpha\}$ for some nonzero vector $a \in X$ and some real number α.

(vi) (Halfspace) A hyperplane H divides X into two sets $H^+ = \{x \in X : \langle a, x \rangle \geq \alpha\}$ and $H^- = \{x \in X : \langle a, x \rangle \leq \alpha\}$, these sets intersect in H. These sets are called halfspaces.

The following definitions are from Rockafellar's book [33].

Definition 2.2 *A set $K \subseteq \mathbb{R}^n$ is called a polyhedron if it can be described as the intersection of finitely many closed half-spaces.*

Hence, a polyhedron can be described as the solution set of a system of finitely many (non-strict) linear inequalities. It is straightforward to show that a polyhedron is a convex set.

A *(convex) polytope* is a set of the following form:

Definition 2.3 *(Polytope) A set $K \subseteq \mathbb{R}^n$ is called a (convex) polytope if it is the convex hull of finitely many points.*

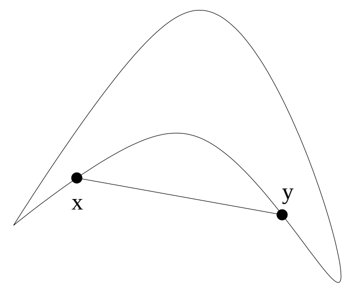

Figure 2.2: A non-convex set.

Clearly, all polytopes are convex since a convex hull is always convex. Examples of (convex) polytopes in $\mathbb{R}^2$ are triangles, squares and hexagons.

Actually, all polytopes in $\mathbb{R}^n$ are compact sets.

Lemma 2.4 *Let $K \subseteq \mathbb{R}^n$ be a polytope. Then K is a compact set.*

Proof: Since K is a polytope, it is the convex hull of finitely many points, say $K = \text{conv}(\{k_1\}, ..., \{k_m\})$, so

$$K = \{\sum_{i=1}^{m} \lambda_i k_i : \sum_{i=1}^{m} \lambda_i = 1, \lambda_i \geq 0 \text{ for all } i = 1, ..., m\}.$$

Consider the continuous function $f : \mathbb{R}^m \to \mathbb{R}^n$, $f(x_1, ..., x_m) = \sum_{i=1}^{m} x_i k_i$, and the compact set

$$S = \{(\lambda_1, ..., \lambda_m) : \sum_{i=1}^{m} \lambda_i = 1, \lambda_i \geq 0 \text{ for all } i = 1, ..., m\} \subseteq \mathbb{R}^m$$

(S is closed and bounded, hence compact in $\mathbb{R}^m$)

Then, since f is continuous and S is compact, $f(S) := \{x : x = f(s) \text{ for some } s \in S\} \subseteq \mathbb{R}^n$ is a compact set (see for example Munkres [26]). But $f(S) = K$ from the definitions, and hence K is compact. $\square$

From Lemma 2.4, any polytope is a closed and bounded set, since compactness is equivalent to being closed and bounded in $\mathbb{R}^n$.

The following theorem connects the notion of polytope and polyhedron.

Theorem 2.5 *A set $K \subseteq \mathbb{R}^n$ is a polytope if and only if it is a bounded polyhedron.*

For a proof of this, see Ziegler [43].

Sometimes, one needs to consider what is called the relative interior of a set.

Definition 2.6 *(Relative interior,* $\mathrm{rint}(\cdot)$*) Let $S \subseteq X$. $x \in S$ is a relative interior point of S if it is contained in some open set whose intersection with* $\mathrm{aff}(S)$ *is contained in S.* $\mathrm{rint}(S)$ *is the set of all relative interior points of S. Here,* $\mathrm{aff}(S)$ *is the smallest affine set that contains S (where a set is affine if it contains any affine combination of its points; an affine combination is like a convex combination except the coefficients are allowed to be negative).*

Another useful notion is that of a convex cone.

Definition 2.7 *(Convex cone) $C \subseteq X$ is called a convex cone if for all $x, y \in C$ and all $\alpha, \beta \geq 0$:*

$$\alpha x + \beta y \in C$$

From these definitions, one can derive some properties of convex sets.

Theorem 2.8 *(Properties of convex sets)*

(i) If $\{C_j\}_{j \in J} \subseteq X$ is an arbitrary family of convex sets, then the intersection $\cap_{j \in J} C_j$ is also a convex set.

(ii) $\mathrm{conv}(A)$ *is a convex set, and it is the smallest (set inclusion-wise) convex set containing A.*

(iii) If $C_1, C_2, \ldots, C_m \subseteq X$ are convex sets, then the Cartesian product $C_1 \times C_2 \times \ldots \times C_m$ is also a convex set.

(iv) If $C \subseteq X$ is a convex set, then the interior of C, $\mathrm{int}(C)$*, the relative interior* $\mathrm{rint}(C)$ *and the closure of C,* $\mathrm{cl}(C)$*, are convex sets as well.*

Proof: $\diamondsuit$
Follows from the definitions of convex set, $\mathrm{conv}(\cdot)$, intersection, Cartesian product, interior, relative interior and closure. Statement (i) also uses the fact that any convex set must contain all convex combinations of its elements. This can be proved by induction, using that C is convex and that a convex combination of convex combinations is a convex combination. $\square$

Sometimes, one considers not just $\mathbb{R}$, but $\bar{\mathbb{R}}$, the extended real numbers.

Definition 2.9 *(The extended real numbers, $\bar{\mathbb{R}}$) Let $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ denote the extended real numbers.*

When working with the extended real numbers the following computational rules apply: $a - \infty = -\infty$, $a + \infty = \infty$, $\infty + \infty = \infty$, $-\infty - \infty = -\infty$ and $\infty - \infty$ is not defined.

The following function is often useful, in particular in optimization.

Definition 2.10 *(The indicator function for a set M, δ_M) Let $M \subseteq X$ be a set. The indicator function for the set M, $\delta_M : X \to \bar{\mathbb{R}}$ is defined as*

$$\delta_M(x) = \begin{cases} 0 & \text{if } x \in M \\ +\infty & \text{if } x \notin M. \end{cases}$$

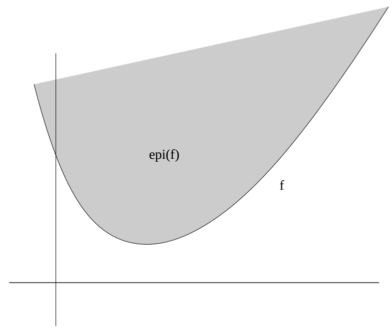

Figure 2.3: The epigraph of a function f.

The following example shows why this function is useful in optimization. Consider the constrained minimization problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in M \end{array}$$

for some function $f : X \to \bar{\mathbb{R}}$ and some set $M \subseteq X$. This can be transformed into an unconstrained minimization problem by altering the objective function as follows

$$\min f(x) + \delta_M(x).$$

This is the same problem as before because the minimum above cannot be achieved for $x \notin M$, because then $\delta_M = +\infty$, so the objective function is infinitely large as well.

The next definition is very important.

Definition 2.11 *(Convex function) Let $C \subseteq X$ be a convex set. A function $f : C \to \mathbb{R}$ is called convex if the inequality*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \tag{2.1}$$

holds for all $x, y \in C$ and every $0 \leq \lambda \leq 1$.

There is an alternative way of defining convex functions, which is based on the notion of epigraph.

Definition 2.12 *(Epigraph, $\mathrm{epi}(\cdot)$) Let $f : X \to \bar{\mathbb{R}}$ be a function. Then the epigraph of f is defined as $\mathrm{epi}(f) = \{(x, \alpha) : x \in X,\, \alpha \in \mathbb{R},\, \alpha \geq f(x)\}$.*

Definition 2.13 *(Convex function) Let $A \subseteq X$. A function $f : A \to \bar{\mathbb{R}}$ is called convex if the epigraph of f is convex (as a subset of the vector space $X \times \mathbb{R}$).*

Of course, these definitions are actually equivalent.

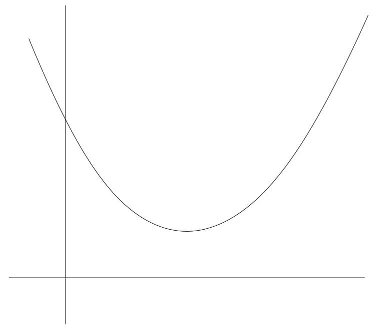

Figure 2.4: A convex function.

Theorem 2.14 *Definitions 2.11 and 2.13 are equivalent if the set A in Definition 2.13 is convex (A must be convex in order for Definition 2.11 to make sense).*

Proof: $\diamondsuit$
2.11 $\Rightarrow$ 2.13: Assume that f is a convex function according to Definition 2.11. Let $(x, a), (y, b) \in \mathrm{epi}(f)$ and let $\lambda \in [0, 1]$. Then

$$\lambda(x, a) + (1 - \lambda)(y, b) \quad = \quad (\lambda x + (1 - \lambda)y, \lambda a + (1 - \lambda)b).$$

But $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ from Definition 2.11, so

$$\begin{aligned} f(\lambda x + (1 - \lambda)y) &\leq \lambda f(x) + (1 - \lambda)f(y) \\ &\leq \lambda a + (1 - \lambda)b. \end{aligned}$$

So $(\lambda x + (1 - \lambda)y, \lambda a + (1 - \lambda)b) \in \mathrm{epi}(f)$.
2.13 $\Rightarrow$ 2.11 uses the same type of arguments, thus it is omitted. $\square$

Definition 2.15 *(Concave function) A function g is concave if the function $f := -g$ is convex.*

When minimizing a function, the points where it is infinitely large are uninteresting, this motivates the following definitions.

Definition 2.16 *(Effective domain, $\mathrm{dom}(\cdot)$) Let $A \subseteq X$ and let $f : A \to \bar{\mathbb{R}}$ be a function. The effective domain of f is defined as $\mathrm{dom}(f) = \{x \in A : f(x) < +\infty\}$.*

Definition 2.17 *(Proper function) Let $A \subseteq X$ and let $f : A \to \bar{\mathbb{R}}$ be a function. f is called proper if $\mathrm{dom}(f) \neq \emptyset$ and $f(x) > -\infty$ for all $x \in A$.*

For definitions of general topological terms, such as convergence, continuity and neighborhood, see any basic topology book, for instance *Topology* by James Munkres [26].

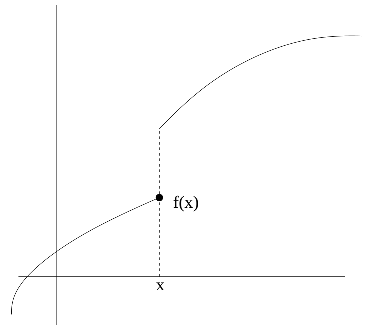

Figure 2.5: A lower semi-continuous function f.

Definition 2.18 *(Lower semi-continuity,* lsc*) Let $A \subseteq X$ be a set, and let $f : A \to \bar{\mathbb{R}}$ be a function. f is called lower semi-continuous,* lsc*, at a point $x_0 \in A$ if for each $k \in \mathbb{R}$ such that $k < f(x_0)$ there exists a neighborhood U of x_0 such that $f(u) > k$ for all $u \in U$. Equivalently: f is lower semi-continuous at x_0 if and only if $\liminf_{x \to x_0} f(x) \geq f(x_0)$.*

Definition 2.19 *(α-sublevel set of a function, $S_\alpha(f)$) Let $f : X \to \bar{\mathbb{R}}$ be a function and let $\alpha \in \mathbb{R}$. The α-sublevel set of f, $S_\alpha(f)$, is defined as*

$$S_\alpha(f) = \{x \in X : f(x) \leq \alpha\}.$$

Theorem 2.20 *Let $f : X \to \bar{\mathbb{R}}$ be a function. Then, f is lower semi-continuous if and only if the sublevel sets $S_\alpha(f)$ are closed for all $\alpha \in \bar{\mathbb{R}}$.*

Proof: The sublevel sets $S_\alpha(f) := \{x \in X : f(x) \leq \alpha\}$ are closed for all $\alpha \in \mathbb{R}$ iff. the complement sets $Y = X - S_\alpha(f) = \{x \in X : f(x) > \alpha\}$ are open for all α. But this happens iff. all $y \in Y$ are interior points, which is equivalent with that for each $y \in Y$ there is a neighborhood U such that $U \subseteq Y$, i.e. $f(U) > \alpha$. But this is the definition of f being lower semi-continuous at the point y. Since this argument holds for all $y \in X$ (by choosing different α), f is lower semi-continuous. $\square$

Definition 2.21 *(Convex hull of a function,* co(f)*) Let $A \subseteq X$ be a set, and let $f : A \to \bar{\mathbb{R}}$ be a function. Then the convex hull of f is the (pointwise) largest convex function h such that $h(x) \leq f(x)$ for all $x \in A$.*

Clearly, if f is a convex function $\mathrm{co}(f) = f$. One can define *the lower semi-continuous hull*, $\mathrm{lsc}(f)$ of a function f in a similar way.

Definition 2.22 *(Closure of a function,* cl$f$*) Let $A \subseteq X$ be a set, and let $f : A \to \bar{\mathbb{R}}$ be a function. We define: $\mathrm{cl}(f)(x) = \mathrm{lsc}(f(x))$ for all $x \in A$ if $\mathrm{lsc}(f(x)) > -\infty \ \forall\, x \in X$ and $\mathrm{cl}(f)(x) = -\infty$ for all $x \in A$ if $\mathrm{lsc}(f(x')) = -\infty$ for some $x' \in A$.*

We say that a function f is *closed* if $\mathrm{cl}(f) = f$. Hence, f is closed if it is lower semi-continuous and $f(x) > -\infty$ for all x or if $f(x) = -\infty$ for all x.

Theorem 2.23 *Let $M \subseteq X$, and consider the indicator function for the set M, δ_M, as defined in Definition 2.10. Then, the following properties hold:*

- *If $N \subseteq X$, then $M \subseteq N \iff \delta_N \leq \delta_M$.*
- *M is a convex set $\iff$ δ_M is a convex function.*
- *δ_M is lower semi-continuous $\iff$ M is a closed set.*

Proof: $\diamondsuit$

- From Definition 2.10: $\delta_N \leq \delta_M$ iff. (If $\delta_M(x) < +\infty$ then $\delta_N(x) < +\infty$) iff. ($x \in M \Rightarrow x \in N$) iff. $M \subseteq N$.
- δ_M is convex if and only if $\delta_M(\lambda x + (1-\lambda)y) \leq \lambda \delta_M(x) + (1-\lambda)\delta_M(y)$ holds for all $0 \leq \lambda \leq 1$ and all $x, y \in X$ such that $\delta_M(x), \delta_M(y) < +\infty$, that is, for all $x, y \in M$. But this means that $\lambda x + (1-\lambda)y \in M$, equivalently, M is convex.
- Assume δ_M is lower semi-continuous. Then it follows from Theorem 2.20 that $S_\alpha(\delta_M)$ is closed for all $\alpha \in \mathbb{R}$. But, for any $\alpha \in \mathbb{R}$, $S_\alpha(\delta_M) = \{x \in X : \delta_M(x) \leq \alpha\} = M$ (from the definition of δ_M), so M is closed. Conversely, assume that M is closed. Then, for any $\alpha \in \mathbb{R}$, $S_\alpha(\delta_M) = M$, hence δ_M is lower semi-continuous from Theorem 2.20.

$\square$

A *global minimum* for a function $f : A \to \bar{\mathbb{R}}$, where $A \subset X$, is an $x' \in A$ such that $f(x') \leq f(x)$ for all $x \in A$. A *local minimum* for f is an $x' \in A$ such that there exists a neighborhood U of x' such that $x \in U \Rightarrow f(x') \leq f(x)$.

Based on all these definitions, one can derive the following properties of convex functions.

Theorem 2.24 *(Properties of convex functions) Let $C \subseteq X$ be a convex set, $f : C \to \mathbb{R}$ be a convex function. Then the following properties hold:*

1. *If f has a local minimum x', then x' is also a global minimum for f.*
2. *If $C = \mathbb{R}$, so that $f : \mathbb{R} \to \mathbb{R}$ and f is differentiable, then f' in monotonically increasing.*
3. *If a function $g : \mathbb{R} \to \mathbb{R}$ is twice differentiable and $g''(x) > 0$, then g is convex.*
4. Jensen's inequality*: For $x_1, \dots, x_n \in X, \lambda_1, \dots, \lambda_n \in \mathbb{R}, \lambda_k \geq 0$, for $k = 1, \dots, n$, $\sum_{k=1}^n \lambda_k = 1$, the following inequality holds*

$$f(\sum_{k=1}^n \lambda_k x_k) \leq \sum_{k=1}^n \lambda_k f(x_k).$$

5. *The sum of convex functions is convex.*

6. αf *is convex if* $\alpha \in \mathbb{R}, \alpha \geq 0$.

7. *If* $(f_n)_{n \in \mathbb{N}}$ *is a sequence of functions,* $f_n : X \to \mathbb{R}$, *and* $f_n \to f$ *pointwise as* $n \to \infty$, *then* f *is convex.*

8. $\mathrm{dom}(f)$ *is a convex set*

9. *If* $\alpha \in \bar{\mathbb{R}}$, *then the sublevel set for* f, $S_\alpha(f)$ *is a convex set. Similarly,* $\{x \in C : f(x) < \alpha\}$ *is a convex set.*

10. Maximization*: Let* $\{f_\lambda\}$ *be an arbitrary family of convex functions, then* $g(x) = \sup_\lambda f_\lambda(x)$ *is convex. Also,* $g(x) = \sup_y f(x, y)$ *is convex if* $f(x, y)$ *is convex in* x *for all* y.

11. Minimization*: Let* $f : X \times X \to \bar{\mathbb{R}}$ *be a convex function. Then* $g(x) = \inf_y f(x, y)$ *is convex.*

Proof: $\diamondsuit$

1. Suppose x' is a local minimum for f, that is: There exists a neighborhood $U \subseteq C$ of x' such that $f(x') \leq f(x)$ for all $x \in U$. We want to show that $f(x') \leq f(x)$ for all $x \in C$. Let $x \in C$. Consider the convex combination $\lambda x + (1 - \lambda)x'$. This convex combination converges towards x' as $\lambda \to 0$. Therefore, for a sufficiently small λ^*, $\lambda^* x + (1 - \lambda^*)x' \in U$, so since f is convex

$$\begin{aligned} f(x') &\leq f(\lambda^* x + (1 - \lambda^*)x') \\ &\leq \lambda^* f(x) + (1 - \lambda^*)f(x') \end{aligned}$$

which, by rearranging the terms, shows that $f(x') \leq f(x)$. Therefore, x' is a global minimum as well.

2. Follows from Definition 2.11 and the definition of the derivative.

3. Use Definition 2.11 and the mean value inequality, see for example *Kalkulus* by Lindstrøm [23], or any other basic calculus book.

4. Follows from Definition 2.11 by induction, and the fact that a convex combination of convex combinations is a convex combination.

5. Use Definition 2.11 and induction.

6. Follows from Definition 2.11.

7. Use Definition 2.11 and the homogeneity and additivity of limits.

8. Follows from the definitions.

9. Follows from the definitions, but is included here as an example of a typical basic proof. Let $x, y \in S_\alpha(f)$. Then $f(x), f(y) \leq \alpha$. Then $\lambda x + (1-\lambda)y \in S_\alpha(f)$ because

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \leq \lambda\alpha + (1-\lambda)\alpha = \alpha$$

where the first inequality follows from the convexity of f, and the second inequality follows from that $x, y \in S_\alpha(f)$.

10. sup is a limit, so the result is a consequence of property 7.

11. Same as property 10.

$\square$

2.2 Optimization

Optimization is the mathematical theory of maximization and minimization problems. It is useful in many applications, for example in logistic problems, finding the best spot to drill for oil, and in mathematical finance. In finance, one often considers an investor who wishes to maximize her utility, given various constraints (for instance her salary). The question is how one can solve such problems. This section gives a short summary of some of the background theory of optimization that will be used in this thesis.

Let X be a vector space, $f : X \to \bar{\mathbb{R}}$, $g : X \to \mathbb{R}^n$ and $S \subseteq X$. Consider an optimization problem of the form

$$\begin{array}{ll} \text{min} & f(x) \\ \text{subject to} & \\ & g(x) \leq 0 \text{ (componentwise)} \\ & x \in S. \end{array} \tag{2.2}$$

In problem (2.2), f is called the *objective function*, while $g(x) \leq 0$, $x \in S$ are called the *constraints* of the problem.

A useful technique when dealing with optimization problems is transforming the problem. For example, a constraint of the form $h(x) \geq y$ (for $h : X \to \mathbb{R}^n$, $y \in \mathbb{R}^n$) is equivalent to $y - h(x) \leq 0$, which is of the form $g(x) \leq 0$ with $g(x) = y - h(x)$. Similarly, any maximization problem can be turned into a minimization problem (and visa versa) by using that $\inf f(x) = -\sup(-f(x))$. Also, any equality constraint can be transformed into two inequality constraints: $h(x) = 0$ is equivalent to $h(x) \leq 0$ *and* $h(x) \geq 0$.

One of the most important theorems of optimization is the extreme value theorem (see Munkres [26]).

Theorem 2.25 *(The extreme value theorem) If $f : X \to \mathbb{R}$ is a continuous function from a compact set into the real numbers, then there exist points $a, b \in X$ such that $f(a) \geq f(x) \geq f(b)$ for all $x \in X$. That is, f* attains *a maximum and a minimum.*

The importance of the extreme value theorem is that it gives the existence of a maximum and a minimum in a fairly general situation. However, these may not be unique. But, for convex (or concave) functions, Theorem 2.24 implies that any local minimum (maximum) is a global minimum (maximum). This makes convex functions useful in optimization.

For a function $f : \mathbb{R}^n \to \mathbb{R}$, the maximum and minimum are attained in *critical points*. Critical points are points x such that

- $f'(x) = 0$, where f is differentiable at x,
- the function f is not differentiable at x or
- x is on the boundary of the set one is optimizing over.

Hence, for a differentiable function which is optimized without extra constraints, one can find maximum and minimum points by solving $f'(x) = 0$ and comparing the objective value in these points to those of the points on the boundary.

Constrained optimization can be tricky to handle. An example of constrained optimization is linear programming (LP); maximization of linear functions under linear constraints. In this situation, strong theorems regarding the solution has been derived. It turns out that corresponding to each LP problem, there is a "dual" problem, and these two problems have the same optimal value. This dual problem is introduced in order to get a second chance at solving an otherwise difficult problem. There is also an effective numerical method for solving LP problems, called the *simplex algorithm*. See Vanderbei [42] for more about linear programming.

The concept of deriving a "dual" problem to handle constraints is the idea of Lagrange duality (see Section 5.4) as well. Lagrange duality begins with a problem of the form (2.2) (or the corresponding maximization problem), and derives a dual problem which gives lower (upper) bounds on the optimal value of the problem. Actually, linear programming duality is a special case of Lagrange duality, but since Lagrange duality is more general, one cannot get the strong theorems of linear programming. The duality concept is generalized even more in convex duality theory, which is the topic of Section 2.3.

2.3 Convex duality and optimization

This section is based on *Conjugate Duality and Optimization* by Rockafellar [34]. As mentioned, convex functions are very handy in optimization problems because of property 1 of Theorem 2.24: For any convex function, a local minimum is also a global minimum.

Another advantage with convex functions in optimization is that one can exploit *duality properties* in order to solve problems. In the following, let X be a linear space, and let $f : X \to \mathbb{R}$ be a function. The main idea of convex duality is to view a given minimization problem $\min_{x \in X} f(x)$ (note that it is common to write min instead of inf when introducing a minimization problem even though one does not know that the minimum is attained) as one half of a minimax problem where a saddle value exists. Very roughly, one does this by looking at an abstract optimization problem

$$\min_{x \in X} F(x, u) \tag{2.3}$$

where $F : X \times U \to \mathbb{R}$ is a function such that $F(x, 0) = f(x)$, U is a linear space and $u \in U$ is a parameter one chooses depending on the particular problem at hand. For example, u can represent time or some stochastic vector expressing uncertainty in the problem data. Corresponding to this problem, one defines an optimal value function

$$\varphi(u) = \inf_{x \in X} F(x, u), \quad u \in U. \tag{2.4}$$

We then have the following theorem:

Theorem 2.26 *Let X, U be real vector spaces, and let $F : X \times U \to \overline{\mathbb{R}}$ be a convex function. Then φ is convex as well.*

Proof: This follows from property 10 of Theorem 2.24. □

The following is a more detailed illustration of the dual optimization method: Let X and Y be general linear spaces, and let $K : X \times Y \to \overline{\mathbb{R}}$ be a function. Define

$$f(x) = \sup_{y \in Y} K(x, y) \tag{2.5}$$

and

$$g(y) = \inf_{x \in X} K(x, y). \tag{2.6}$$

Then, consider two optimization problems

$$(P) \qquad \min_{x \in X} f(x)$$

and

$$(D) \qquad \max_{y \in Y} g(y).$$

From the definitions

$$g(y) \le K(x, y) \le f(x), \quad \forall\, x \in X, \forall\, y \in Y. \tag{2.7}$$

By taking the infimum over x and then the supremum over y in equation (2.7)

$$\inf_{x \in X} \sup_{y \in Y} K(x, y) = \inf_{x \in X} f(x) \ge \sup_{y \in Y} g(y) = \sup_{y \in Y} \inf_{x \in X} K(x, y). \tag{2.8}$$

If there is equality in equation (2.8), then the common value is called the *saddle value* of K.

The saddle value exists if K has a *saddle point*, i.e. there exists a point (x', y') such that

$$K(x', y) \leq K(x', y') \leq K(x, y') \tag{2.9}$$

for all $x \in X$ and for all $y \in Y$. If such a point exists, the saddle value of K is $K(x', y')$.

From these definitions, one can prove the following theorem.

Theorem 2.27 *A point (x', y') is a saddle point for K if and only if x' solves (P), y' solves (D), and the saddle value of K exists, i.e.*

$$\inf_{x \in X} f(x) = \sup_{y \in Y} g(y)$$

Proof: One can rewrite the saddle point condition (2.9) as

$$f(x') = K(x', y') = g(y').$$

The theorem then follows from equation (2.8). □

Because of this theorem, (P) and (D) are called *dual problems*, since they can be considered as half of the problem of finding a saddle point for K.

Hence, in order to prove that (P) and (D) have a solution, and actually find this solution one can instead attempt to find a saddle point for the function K.

In convex optimization, one is interested in what has been done above in the opposite order: If one starts with (P), where $f : X \to \mathbb{R}$, how can one choose a space Y and a function K on $X \times Y$ such that $f(x) = \sup_{y \in Y} K(x, y)$ holds? This approach gives freedom to choose Y and K in different ways, so that one can (hopefully) achieve the properties one would like Y and K to have. This idea is called *the duality approach*.

2.4 Examples of convex optimization via duality

Example 2.28 *(Nonlinear programming) Let $f_0, f_1, \dots, f_m$ be real valued, convex functions on a nonempty, convex set C in the vector space X. The duality approach consists of the following steps:*

1. *The given problem:* $\min f_0(x)$ *over* $\{x \in C : f_i(x) \leq 0 \ \forall \ i = 1, \dots, m\}$.

2. *Abstract representation:* $\min f$ *over X, where*

$$f(x) = \begin{cases} f_0(x) & x \in C, \ f_i(x) \leq 0 \text{ for } i = 1, \dots, m \\ +\infty & \text{for all other } x \in X. \end{cases}$$

3. *Parametrization: Define (for example) $F(x, u)$ for $u = (u_1, \dots, u_m) \in \mathbb{R}^m$ by $F(x, u) = f_0(x)$ if $x \in C, f_i(x) \leq u_i$ for $i = 1, \dots, m$, and $F(x, u) = +\infty$ for all other x. Then, $F : X \times \mathbb{R}^m \to [-\infty, +\infty]$ is convex and $F(x, 0) = f(x)$*

Example 2.29 *(Nonlinear programming with infinitely many constraints) Let $f_0 : C \to \mathbb{R}$ where $C \subset X$ is convex, and let $h : X \times S \to \bar{\mathbb{R}}$ be convex in the x-argument, where S is an arbitrary set*

1. *The problem:* $\min f_0(x)$ *over* $K = \{x \in C : h(x, s) \leq 0 \ \forall \ s \in S\}$.
2. *Abstract representation:* $\min f(x)$ *over X, where $f(x) = f_0(x)$ if $x \in K$, and $f(x) = +\infty$ for all other x.*
3. *Parametrization: Choose u analogously with Example 2.28: Let U be the linear space of functions $u : S \to \mathbb{R}$ and let $F(x, u) = f_0(x)$ if $x \in C$, $h(x, s) \leq u(s) \ \forall \ s \in S$ and $F(x, u) = +\infty$ for all other x. As in the previous example, this makes F convex and satisfies $F(x, 0) = f(x)$.*

Example 2.30 *(Stochastic optimization) Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $h : X \times \Omega \to \bar{\mathbb{R}}$ be convex in the x-argument, where X is a linear, topological space. Let C be a closed, convex subset of X.*

1. *The general problem:* $\min h(x, \omega)$ *over all $x \in C$, where ω is a stochastic element with a known distribution. The difficulty here is that x must be chosen before ω has been observed.*
2. *We therefore solve the following problem: Minimize the expectation $f(x) = \int_\Omega h(x, \omega) dP(\omega)$ over all $x \in X$. Here, it is assumed that h is measurable, so that f is well defined. Rockafellar then shows in [34], Theorem 3, that f actually is convex.*
3. *Parametrization: Let $F(x, u) = \int_\Omega h(x - u(\omega), \omega) dP(\omega) + \delta_C(u)$ for $u \in U$, where U is a linear space of measurable functions and δ_C is the indicator function of C, as defined in Definition 2.10. Then F is (by the same argument as for f) well defined and convex, with $F(x, 0) = f(x)$.*

2.5 Conjugate functions in paired spaces

The material in this section is based on Rockafellar [34] and Rockafellar and Wets [36].

Definition 2.31 *(Pairing of spaces) A pairing of two linear spaces X and V is a real valued bilinear form $\langle \cdot, \cdot \rangle$ on $X \times V$.*

The pairing associates for each $v \in V$ a linear function $\langle \cdot, v \rangle : x \mapsto \langle x, v \rangle$ on X, and similarly for X.

Definition 2.32 *(Compatible topology) Assume there is a pairing between the spaces X and V. A topology on X is compatible with the pairing if it is a locally convex topology such that the linear function $\langle \cdot, v \rangle$ is continuous, and any continuous linear function on X can be written in this form for some $v \in V$. A compatible topology on V is defined similarly.*

Definition 2.33 *(Paired spaces) X and V are paired spaces if one has chosen a pairing between X and V, and the two spaces have compatible topologies with respect to the pairing.*

Example 2.34 *Let $X = \mathbb{R}^n$ and $V = \mathbb{R}^n$. Then, the standard Euclidean inner product is a bilinear form, so X and V become paired spaces.*

Example 2.35 *Let $X = L^1(\Omega, \mathcal{F}, P)$ and $V = L^\infty(\Omega, \mathcal{F}, P)$. Then X and V are paired via the bilinear form $\langle x, v\rangle = \int_\Omega x(s)v(s)dP(s)$. Similarly, the spaces $X = L^p(\Omega, F, P)$ and $V = L^q(\Omega, F, P)$, where $\frac{1}{p} + \frac{1}{q} = 1$, are paired.*

We now come to a central notion of convex duality, the conjugate of a function.

Definition 2.36 *(Convex conjugate of a function, f^*) Let X and V be paired spaces. For a function $f : X \to \bar{\mathbb{R}}$, define the conjugate of f, denoted by $f^* : V \to \bar{\mathbb{R}}$, by*

$$f^*(v) = \sup\{\langle x, v\rangle - f(x) : x \in X\}. \tag{2.10}$$

Finding f^* is called taking the conjugate of f in the convex sense. One may also define the conjugate g^* of a function $g : V \to \bar{\mathbb{R}}$ similarly.

Similarly, define

Definition 2.37 *(Biconjugate of a function, f^{**}) Let X and V be paired spaces. For a function $f : X \to \bar{\mathbb{R}}$, define* the biconjugate *of f, f^{**}, to be the conjugate of f^*, so $f^{**}(x) = \sup\{\langle x, v\rangle - f^*(v) : v \in V\}$.*

Definition 2.38 *(The Fenchel transform) The operation $f \mapsto f^*$ is called* the Fenchel transform.

If $f : \mathbb{R}^n \to \bar{\mathbb{R}}$, then the operation $f \mapsto f^*$ is sometimes called the *Legendre-Fenchel transform*.

To understand why the conjugate function f^* is important, it is useful to consider it via the epigraph. This is most easily done in $\mathbb{R}^n$, so let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ and consider $X = \mathbb{R}^n = V$. From equation (2.10), it is not difficult to show that

$$(v, b) \in \text{epi}(f^*) \iff b \geq \langle v, x\rangle - a \text{ for all } (x, a) \in \text{epi}(f). \tag{2.11}$$

This can also be expressed as

$$(v, b) \in \text{epi}(f^*) \iff l_{v,b} \leq f \tag{2.12}$$

where $l_{v,b}(x) := \langle v, x\rangle - b$. So, since specifying a function on $\mathbb{R}^n$ is equivalent to specifying its epigraph, equation (2.12) shows that f^* describes the family of all affine functions that are majorized by f (since all affine functions on $\mathbb{R}^n$ are of the form $\langle v, x\rangle - b$ for fixed v, b).

But from equation (2.11)

$$b \geq f^*(v) \iff b \geq l_{x,a}(v) \text{ for all } (x, a) \in \text{epi}(f).$$

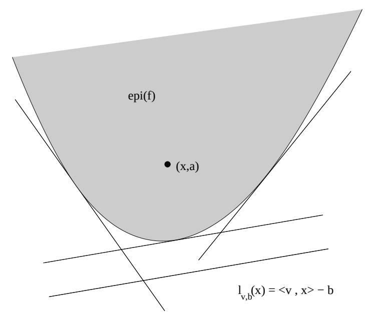

Figure 2.6: Affine functions majorized by f.

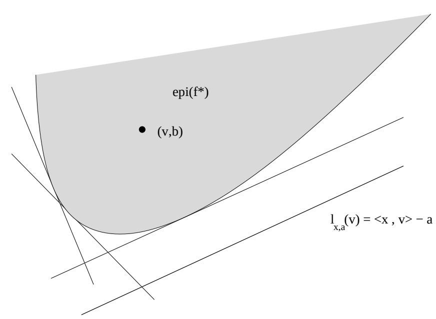

Figure 2.7: Affine functions majorized by f^*.

This means that f^* is the pointwise supremum of all affine functions $l_{x,a}$ for $(x, a) \in \text{epi}(f)$.

This is illustrated in Figures 2.6 and 2.7.

We then have the following very central theorem on duality, which is Theorem 5 in Rockafellar [34]:

Theorem 2.39 *Let $f : X \to \bar{\mathbb{R}}$ be arbitrary. Then the conjugate f^* is a closed (as defined in Section 2.1), convex function on V and $f^{**} = \text{cl}(\text{co}(f))$. Similarly if one starts with a function in V. In particular, the Fenchel transform induces a one-to-one correspondence $f \mapsto h, h = f^*$ between the closed, convex functions on X and the closed, convex functions on V.*

Proof: By definition f^* is the pointwise supremum of the continuous, affine functions $V \mapsto \langle x, v\rangle - \alpha$, where $(x, \alpha) \in \text{epi}(f)$. Therefore, f^* is convex and lsc, hence it is closed. $(v, \beta) \in \text{epi}(f^*)$ if and only if the continuous affine function $x \mapsto \langle x, v\rangle - \beta$ satisfies $f(x) \geq \langle x, v\rangle - \beta$ for all $x \in X$, that is if the epigraph of this affine function contains the epigraph of f. Thus, $\text{epi}(f^{**})$ is the intersection of all the nonvertical, closed halfspaces in $X \times \mathbb{R}$ containing $\text{epi}(f)$. This implies, using what a closed, convex set is, that $f^{**} = \text{cl}(\text{co}(f))$. $\square$

Theorem 2.39 implies that if f is convex and closed, then $f = f^{**}$. This gives a one-to-one correspondence between the closed convex functions on X, and the same type of functions on V. Hence, all properties and operations on such functions must have conjugate counterparts (see [36]).

Example 2.40 $\diamondsuit$

Let X and V be paired spaces, and let $f = \delta_L$ where $L \subseteq X$ is a subspace (so in particular, L is convex) and δ_L is the indicator function of L, as in Definition 2.10. It follows from Example 2.23 that $f = \delta_L$ is convex. Then

$$\begin{aligned} \delta_L^*(v) &= \sup\{\langle x, v\rangle - \delta_L(x) : x \in X\} \\ &= \sup\{\langle x, v\rangle ; x \in L\} \end{aligned}$$

since $\langle x, v\rangle - \delta_L(x) = -\infty$ if $x \notin L$. This function δ_L^* is called the support function *of L (and is often denoted by ψ_L). Note also that*

$$\delta_L^*(v) = \delta_{L^\perp}(v)$$

because if $v \in L^\perp$, then $\langle x, v\rangle = 0$ for all $x \in L$, but if $v \notin L^\perp$ then $\langle x', v\rangle \neq 0$ for some $x' \in L$. Hence, since L is a subspace, $\langle x', v\rangle$ can be made arbitrarily large by multiplying x' by either $+t$ or $-t$ (in order to make $\langle x', v\rangle$ positive), and letting $t \to +\infty$.

By a similar argument

$$\delta_L^{**} = \delta_{(L^\perp)^\perp}. \tag{2.13}$$

We will now use conjugate duality to prove a central result in functional analysis, namely that for any subspace $L \subseteq X$, $(L^\perp)^\perp = \bar{L}$ (see for instance *Linear Functional Analysis* by Rynne and Youngston [39]).

Theorem 2.41 *Let $L \subseteq X$ be a subspace. Then $(L^{\perp})^{\perp} = \bar{L}$.*

Proof: $\diamondsuit$
From Example 2.40

$$\delta_L^{**} = \delta_{(L^{\perp})^{\perp}} \tag{2.14}$$

But then, Theorem 2.39 implies that $\delta_{(L^{\perp})^{\perp}} = \mathrm{cl}(\mathrm{co}(\delta_L))$. δ_L is convex, so $\mathrm{co}(\delta_L) = \delta_L$. To proceed, we make the following claim:

Claim: $\mathrm{cl}(\delta_L) = \delta_{\bar{L}}$.

Proof of Claim: From Definition 2.22, $\mathrm{cl}(\delta_L) = \mathrm{lsc}(\delta_L) =$ the largest lower semi-continuous function that is less than or equal to δ_L. $\delta_{\bar{L}}$ is lower semi-continuous since $\bar{L}$ is closed (from Theorem 2.20). Also, from Example 2.40, $\delta_{\bar{L}} \leq \delta_L$ since $L \subseteq \bar{L}$. All that remains to be proved is that if f is lower semi-continuous and $f \leq \delta_L$, then $f \leq \delta_{\bar{L}}$.

So assume that f is lower semi-continuous and $f \leq \delta_L$. We know that $\delta_L(L) = \delta_{\bar{L}}(L)$, so $f(L) \leq \delta_L(L) \leq \delta_{\bar{L}}(L)$, from the assumption that $f \leq \delta_L$.

If $x \in (\bar{L})^{\perp}$, then $\delta_{\bar{L}}(x) = +\infty$, so $f(x) \leq \delta_{\bar{L}}(x)$.

Finally, if $x \in \bar{L} \setminus L$, then $\delta_L(x) = +\infty$, but $\delta_{\bar{L}}(x) = 0$. Hence, we must show that $f(x) \leq 0$. Since f is lower semi-continuous, Theorem 2.20 implies that the sublevel set $S_0(f) = \{x \in X : f(x) \leq 0\}$ is closed. Because $f \leq \delta_L$, $L \subseteq S_0(f)$, hence (since $S_0(f)$ is closed) $\bar{L} \subseteq S_0(f)$, so $f(x) \leq 0$ for all $x \in \bar{L}$.

So the claim is proved.

The arguments above imply that

$$\delta_{(L^{\perp})^{\perp}} = \delta_L^{**} = \mathrm{cl}(\mathrm{co}(\delta_L)) = \mathrm{cl}(\delta_L) = \delta_{\bar{L}}$$

where the final equality uses the claim. But this again implies that $(L^{\perp})^{\perp} = \bar{L}$.
$\square$

For a concave function $g : X \to \bar{\mathbb{R}}$ one can define the conjugate as:

$$g^*(v) = \inf\{\langle x, v\rangle - g(x) : x \in X\} \tag{2.15}$$

This is called taking the *conjugate of g in the concave sense.*

2.6 Dual problems and Lagrangians

This is our situation as of now. We have an abstract minimization problem:

$$(P) \qquad \min_{x \in X} f(x)$$

which is assumed to have the representation:

$$f(x) = F(x, 0), \quad F : X \times U \to \bar{\mathbb{R}} \tag{2.16}$$

(where U is some linear space). Everything now depends on the choice of U and F. We want to exploit duality, so let X be paired with V, and U paired with Y, where U and Y are linear spaces (the choice of pairings may also be important in applications). Preferably, we want to choose (F, U) such that F is a closed, jointly convex function of x and u.

Definition 2.42 *(The Lagrange function, $K(x, y)$) Define the Lagrange function $K : X \times Y \to \bar{\mathbb{R}}$ to be*

$$K(x, y) = \inf\{F(x, u) + \langle u, y \rangle : u \in U\}. \tag{2.17}$$

The following theorem is Theorem 6 in Rockafellar [34]. It says that K is a closed convex function which satisfies a certain inequality, and that all functions of this form actually are the Lagrange function associated with some function f.

Theorem 2.43 *The Lagrange function K is closed, concave in $y \in Y$ for each $x \in X$, and if $F(x, u)$ is closed and convex in u*

$$f(x) = \sup_{y \in Y} K(x, y). \tag{2.18}$$

Conversely, if K is an arbitrary extended-real valued function on $X \times Y$ such that (2.18) holds, and if K is closed and concave in y, then K is the Lagrange function associated with a unique representation $f(x) = F(x, 0), F : X \times U \to \bar{\mathbb{R}}$ where F is closed and convex in u. This means that

$$F(x, u) = \sup\{K(x, y) - \langle u, y \rangle : y \in Y\}.$$

Further, if F is closed and convex in u, K is convex in x if and only if $F(x, u)$ is jointly convex in x and u.

Proof: Everything in the theorem, apart from the last statement, follows from Theorem 2.39. For the last statement, assume that F and K respectively are convex, use the definitions of F and K and that the supremum and infimum of convex functions are convex (see Theorem 2.24). $\square$

We now define, motivated by equation (2.18), the dual problem of (P),

$$(D) \qquad \max_{y \in Y} g(y)$$

where $g(y) = \inf_{x \in X} K(x, y)$.

Note that this dual problem gives a lower bound on the primal problem, from (2.18) since

$$K(x, y) \geq \inf_{x \in X} K(x, y) = g(y).$$

But then

$$\sup_{y \in Y} K(x,y) \geq \sup_{y \in Y} g(y).$$

So from equation (2.18), $f(x) \geq \sup_{y \in Y} g(y)$. Therefore, taking the infimum with respect to $x \in X$ on the left hand side implies $(D) \leq (P)$. This is called *weak duality*. Sometimes, one can prove that the dual and primal problems have the same optimal value. If this is the case, there is *no duality gap* and *strong duality holds*.

The next theorem (Theorem 7 in Rockafellar [34]) is important:

Theorem 2.44 *The function g in (D) is closed and concave. By taking the conjugate in concave sense, $g = -\varphi^*$, hence $-g^* = \mathrm{cl}(\mathrm{co}(\varphi))$, so*

$$\sup_{y \in Y} g(y) = \mathrm{cl}(\mathrm{co}(\varphi))(0)$$

while

$$\inf_{x \in X} f(x) = \varphi(0)$$

In particular, if $F(x,u)$ is convex in (x,u), then $-g^* = \mathrm{cl}(\varphi)$ and $\sup_{y \in Y} g(y) = \liminf_{u \to 0} \varphi(u)$ (except if $0 \notin \mathrm{cl}(\mathrm{dom}(\varphi)) \neq \emptyset$, and $\mathrm{lsc}(\varphi)$ is nowhere finite valued).

For the proof, see Rockafellar [34].

What makes this theorem important is that it converts the question of whether $\inf_{x \in X} f(x) = \sup_{y \in Y} g(y)$ and the question of whether the saddle value of the Lagrange function K exists, to a question of whether the optimal value function φ satisfies $\varphi(0) = (\mathrm{cl}(\mathrm{co}(\varphi)))(0)$. Hence, *if the value function φ is convex, the lower semi-continuity of φ is a sufficient condition for the absence of a duality gap.*

By combining the results of the previous sections, we get the following rough summary of the duality method, based on conjugate duality:

- To begin, there is a minimization problem $\min_{x \in X} f(x)$ which cannot be solved directly.
- Find a function $F : X \times U \to \bar{\mathbb{R}}$, where U is a vector space, such that $f(x) = F(x,0)$.
- Introduce the linear space Y, paired with U, and define the Lagrange function $K : X \times Y \to \bar{\mathbb{R}}$ by $K(x,y) = \inf_{u \in U}\{F(x,u) + \langle u, y \rangle\}$.
- Try to find a saddle point for K. If this succeeds, Theorem 2.27 tells us that this gives the solution of (P) and (D).
- Theorem 2.44 tells us that K has a saddle point if and only if $\varphi(0) = (\mathrm{cl}(\mathrm{co}(\varphi)))(0)$. Hence, if the value function φ is convex, the lower semi-continuity of φ is a sufficient condition for the absence of a duality gap.

We can look at an example illustrating these definitions, based on Example 2.28.

Example 2.45 $\diamondsuit$
(Nonlinear programming) The Lagrange function takes the form

$$\begin{aligned}
K(x,y) &= \inf\{F(x,u) + \langle u, y\rangle : u \in U\} \\
&= \inf\{\begin{cases} f_0(x) + \langle u, y\rangle; x \in C, f_i(x) \le u_i \\ +\infty + \langle u, y\rangle; \forall \text{ other } x \end{cases} : u \in U\} \\
&= \begin{cases} f_0(x) + \inf\{\langle u, y\rangle : u \in U, f_i(x) \le u_i\}, x \in C \\ +\infty, \text{ otherwise.} \end{cases} \\
&= \begin{cases} \inf\{f_0(x) + f_1(x)y_1 + \ldots + f_m(x)y_m\}, u \in U, x \in C, y \in \mathbb{R}^m_+ \\ -\infty, x \in C, y \notin \mathbb{R}^m_+ \\ +\infty, \text{ otherwise.} \end{cases}
\end{aligned}$$

where the last equality follows because if there is at least one negative y_j, one can choose u_j arbitrarily large and make the above expression arbitrarily small. Therefore, the dual function is

$$\begin{aligned}
g(y) &= \inf_{x \in X} K(x,y) \\
&= \inf_{x \in X} \begin{cases} f_0(x) + f_1(x)y_1 + \ldots + f_m(x)y_m \text{ if } x \in C, y \in \mathbb{R}^m_+ \\ -\infty, x \in C, y \notin \mathbb{R}^m_+ \\ +\infty, \text{ otherwise.} \end{cases} \\
&= \begin{cases} \inf_{x \in C}\{f_0(x) + f_1(x)y_1 + \ldots + f_m(x)y_m\} \text{ if } y \in \mathbb{R}^m_+ \\ -\infty, y \notin \mathbb{R}^m_+. \end{cases}
\end{aligned}$$

By making some small alterations to the approach above, Rockafellar [34] shows that by beginning with the standard primal linear programming problem (abbreviated LP-problem)

$$\max\{\langle c, x\rangle : Ax \le b, \; x \ge 0\}$$

where c and b are given vectors and A is a given matrix, and finding its dual problem (in the above sense), one gets the standard dual LP-problem back. That is

$$\min\{\langle b, y\rangle : A^T y \ge c, \; y \ge 0\}$$

(see Vanderbei [42]).

The purpose of this chapter has been to introduce the convexity theory necessary to understand duality methods in mathematical finance. The definitions and theorems from convex analysis will be used throughout the thesis, and the conjugate duality theory by Rockafellar [34] will be applied in Sections 3.2, 6.2, 6.4, and 7.2. The next chapter introduces convex risk measures. This is a topic where convexity theory is essential, and it is also useful in mathematical finance.

Chapter 3

Convex risk measures

The purpose of this chapter is to define the notion of convex risk measure, derive some properties of such risk measures and prove a dual representation theorem (Theorem 3.8) via the conjugate duality theory of Chapter 2. Convex risk measures will be applied in a utility maximization problem in Section 5.9. Section 3.1 is devoted to the definition of convex risk measures (Definition 3.1), and some corresponding notions, such as coherent risk measures (Definition 3.2) and acceptance sets (se Definition 3.4). This section also considers some properties of convex risk measures; see Theorem 3.3 and Theorem 3.7.

Section 3.2 is derives dual representation theorems for convex risk measures. The most important of these theorems is Theorem 3.8, which is proved using the convex duality theorem, Theorem 2.39, of Chapter 2.

3.1 The basics of convex risk measures

This section is based on the papers Artzner et al. [1], Föllmer and Schied [13], [12], Frittelli and Gianin [14] and Rockafellar [35]. The notion of coherent risk measure was introduced by Artzner et al. in the seminal paper [1].

Most things in life are uncertain, therefore, risk is very important. How to quantify risk is an interesting question, and this is especially important in finance. Over the years researchers have tried many different methods for this quantification, but most of these seem to be flawed. One approach to the quantification of risk could be computing the variance. The problem in this case is that the variance does not separate between negative and positive deviations, and is therefore not suitable in finance where positive deviations are good (earning more money), but negative deviations are bad (earning less money). In order to resolve this, Artzner et al. [1] set up some economically reasonable axioms that a measure of risk should satisfy and thereby introduced the notion of coherent risk measure. This notion has later been modified to a convex risk measure, which is a more general term.

This chapter requires some measure theory, see for example Shilling [41] for more.

On a given scenario space Ω, one can define a *σ-algebra* $\mathcal{F}$, that is a family of subsets of Ω that contains the empty set $\emptyset$ and is closed under complements and countable unions. The elements in the σ-algebra $\mathcal{F}$ are called *measurable sets*. $(\Omega, \mathcal{F})$ is then called a *measurable space*. A *measurable function* is a function $f : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ (where $(\Omega', \mathcal{F}')$ is another measurable space) such that the inverse image of any measurable set is a measurable set. A *random variable* is a real-valued measurable function. On a measurable space $(\Omega, \mathcal{F})$ one can define a *measure*, i.e. a non-negative countably additive function $\mu : \Omega \to \mathbb{R}$ such that $\mu(\emptyset) = 0$. Then, $(\Omega, \mathcal{F}, \mu)$ is called a measure space. A *signed measure* is the same as a measure, but without the non-negativity requirement. A *probability measure* is a measure P such that $P(\Omega) = 1$. Let $\mathcal{P}$ denote the set of all probability measures on $(\Omega, \mathcal{F})$, and $\mathbb{V}$ the set of all measures on $(\Omega, \mathcal{F})$. Then $\mathbb{V}$ is a vector space, and $\mathcal{P} \subseteq \mathbb{V}$ is a convex set.

In the following, let Ω be a fixed set of scenarios, or possible states of the world. Consider the measure space $(\Omega, \mathcal{F}, P)$, where $\mathcal{F}$ is a given σ-algebra on Ω and P is a given probability measure on $(\Omega, \mathcal{F})$. A financial position, for example a portfolio of stocks, can be described by a mapping $X : \Omega \to \mathbb{R}$, where $X(\omega)$ is the value of the position at the end of the trading period if the state ω occurs. More formally, X is a random variable. Hence, the dependency of X on ω describes the uncertainty of the value of the portfolio. Let $\mathbb{X}$ be a given vector space of such random variables $X : \Omega \to \mathbb{R}$, which contains the constant functions (that is, the functions of the form $(c\mathbf{1})(\omega) = c$ for all $\omega \in \Omega$, where $c \in \mathbb{R}$ is some constant). An example of such a space is $L^p(\Omega, \mathcal{F}, P)$. A convex risk measure is defined as follows:

Definition 3.1 *(Convex risk measure) A convex risk measure is a function $\rho : \mathbb{X} \to \mathbb{R}$ which satisfies the following for each $X, Y \in \mathbb{X}$:*

(*i*) *(Convexity) $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$ for $0 \leq \lambda \leq 1$.*

(*ii*) *(Monotonicity) If $X \leq Y$, then $\rho(X) \geq \rho(Y)$.*

(*iii*) *(Translation invariance) If $m \in \mathbb{R}$, then $\rho(X + m\mathbf{1}) = \rho(X) - m$.*

If $\rho(X) \leq 0$, X is *acceptable* (to an investor), since the portfolio does not have a positive risk. On the other hand, if $\rho(X) > 0$, X is *unacceptable*.

If a convex risk measure also satisfies *positive homogeneity*, that is if

$$\lambda \geq 0 \Rightarrow \rho(\lambda X) = \lambda\rho(X)$$

then ρ is called a *coherent risk measure*. The original definition of a coherent risk measure, did not involve convexity directly, but instead required subadditivity:

Definition 3.2 *(Coherent risk measure) A coherent risk measure is a function $\pi : \mathbb{X} \to \mathbb{R}$ which satisfies the following for each $X, Y \in \mathbb{X}$:*

(*i*) *(Positive homogeneity) $\pi(\lambda X) = \lambda\pi(X)$ for $\lambda \geq 0$.*

(*ii*) *(Subadditivity)* $\pi(X+Y) \leq \pi(X) + \pi(Y)$.

(*iii*) *(Monotonicity) If* $X \leq Y$*, then* $\pi(X) \geq \pi(Y)$.

(*iv*) *(Translation invariance) If* $m \in \mathbb{R}$*, then* $\pi(X + m\mathbf{1}) = \pi(X) - m$.

One can interpret ρ as a *capital requirement*, that is: $\rho(X)$ is the extra amount of money which should be added to (or withdrawn from) the portfolio (in a risk free way) to make the position acceptable for an agent.

The conditions for being a convex risk measure are quite natural. The convexity reflects that diversification reduces risk. The total risk of loss in two portfolios should be reduced when the two are weighed into a mixed portfolio. Roughly speaking, spreading your eggs in several baskets should reduce the risk of broken eggs.

Monotonicity says that the downside risk, the risk of loss, is reduced by choosing a portfolio that has a higher value in every possible state of the world.

Finally, translation invariance can be interpreted in the following way: ρ is the amount of money one needs to add to the portfolio in order to make it acceptable for an agent. Hence, if one adds a risk free amount m to the portfolio, the capital requirement should be reduced by the same amount.

As mentioned, Artzner et al. [1] originally defined coherent risk measures, that is, they required positive homogeneity. The reason for skipping this requirement in the definition of a convex risk measure, is that positive homogeneity means that risk grows linearly with X, and this may not always be the case. Therefore, only convex risk measures will be considered in the following.

Starting with n convex risk measures, one can derive more convex risk measures, as in the following theorem. This was proven by Rockafellar in [35], Theorem 3, for coherent risk measures.

Theorem 3.3 $\diamondsuit$

Let $\rho_1, \rho_2, \ldots, \rho_n$ *be convex risk measures.*

1. *If* $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0$ *and* $\sum_{i=1}^n \lambda_i = 1$*, then* $\rho = \sum_{i=1}^n \lambda_i \rho_i$ *is a convex risk measure as well.*

2. $\rho = \max\{\rho_1, \rho_2, \ldots, \rho_n\}$ *is a convex risk measure.*

Proof: $\diamondsuit$

1. Let's check the requirements in Definition 3.1. Obviously, $\rho : \mathbb{X} \to \mathbb{R}$, so we check for any $X, Y \in \mathbb{X}$, $0 \leq \lambda \leq 1$:

   (*i*) : This follows from that a sum of convex functions is a convex function, and a positive constant times a convex function is also a convex function (see Theorem 2.24 parts 5 and 6, plus induction on part 5).

   (*ii*) : If $X \leq Y$, then $\rho(X) = \sum_{i=1}^n \lambda_i \rho_i(X) \geq \sum_{i=1}^n \lambda_i \rho_i(Y) = \rho(Y)$.

(iii) : If $m \in \mathbb{R}$,

$$\begin{aligned}\rho(X+m) &= \sum_{i=1}^{n} \lambda_i \rho_i(X+m) \\ &= \sum_{i=1}^{n} \lambda_i(\rho_i(X)-m) \\ &= \sum_{i=1}^{n} \lambda_i \rho_i(X) - m \sum_{i=1}^{n} \lambda_i \\ &= \rho(X)-m.\end{aligned}$$

2. Again, let's simply check Definition 3.1 for any $X, Y \in \mathbb{X}$:

(i) : If $0 \leq \lambda \leq 1$,

$$\begin{aligned}\rho(\lambda X+(1-\lambda)Y) &= \max\{\rho_1(\lambda X+(1-\lambda)Y), \\ &\quad \ldots, \rho_n(\lambda X+(1-\lambda)Y)\} \\ &\leq \max\{\lambda\rho_1(X)+(1-\lambda)\rho_1(Y), \\ &\quad \ldots, \lambda\rho_n(X)+(1-\lambda)\rho_n(Y)\} \\ &\leq \lambda \max\{\rho_1(X),\ldots,\rho_n(X)\} \\ &\quad +(1-\lambda)\max\{\rho_1(Y),\ldots,\rho_n(Y)\} \\ &= \lambda\rho(X)+(1-\lambda)\rho(Y).\end{aligned}$$

(ii) : If $X \leq Y$,

$$\begin{aligned}\rho(X) &= \max\{\rho_1(X),\ldots,\rho_n(X)\} \\ &\geq \max\{\rho_1(Y),\ldots,\rho_n(Y)\} \\ &= \rho(Y).\end{aligned}$$

(iii) : For $m \in \mathbb{R}$,

$$\begin{aligned}\rho(X+m) &= \max\{\rho_1(X+m),\ldots,\rho_n(X+m)\} \\ &= \max\{\rho_1(X)-m,\ldots,\rho_n(X)-m\} \\ &= \max\{\rho_1(X),\ldots,\rho_n(X)\}-m \\ &= \rho(X)-m.\end{aligned}$$

□

Associated with every convex risk measure ρ, there is a natural set of all acceptable portfolios, called the *acceptance set*, $\mathcal{A}_\rho$, of ρ.

Definition 3.4 *(The acceptance set of a convex risk measure, $\mathcal{A}_\rho$) A convex risk measure ρ induces a set*

$$\mathcal{A}_\rho = \{X \in \mathbb{X} : \rho(X) \leq 0\}$$

The set $\mathcal{A}_\rho$ is called the acceptance set of ρ.

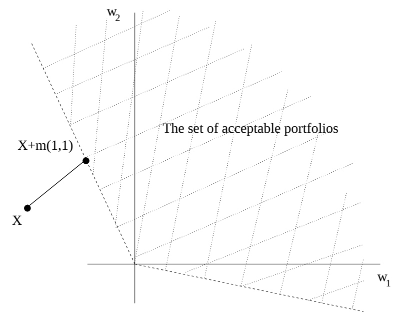

Figure 3.1: Illustration of the risk measure $\rho_{\mathcal{A}}$ associated with a set $\mathcal{A}$ of acceptable portfolios.

Conversely, given a class $\mathcal{A} \subseteq \mathbb{X}$ of acceptable portfolios, one can associate a quantitative risk measure $\rho_{\mathcal{A}}$ to it.

Definition 3.5 *Let $\mathcal{A} \subseteq \mathbb{X}$ be a set of "acceptable" random variables. This set has an associated measure of risk $\rho_{\mathcal{A}}$ defined as follows: For $X \in \mathbb{X}$, let*

$$\rho_{\mathcal{A}}(X) = \inf\{m \in \mathbb{R} : X + m \in \mathcal{A}\}. \tag{3.1}$$

This means that $\rho_{\mathcal{A}}(X)$ measures how much one must add to the portfolio X, in a risk free way, to get the portfolio into the set $\mathcal{A}$ of acceptable portfolios. This is the same interpretation as for a convex risk measure.

The previous definitions show that one can either start with a risk measure, and derive an acceptance set, or one can start with a set of acceptable portfolios, and derive a risk measure.

Example 3.6 $\diamondsuit$
(Illustration of the risk measure $\rho_{\mathcal{A}}$ associated with a set $\mathcal{A}$ of acceptable portfolios) Let $\Omega = \{\omega_1, \omega_2\}$, and let $X : \Omega \to \mathbb{R}$ be a portfolio. Let $x = (X(\omega_1), X(\omega_2))$. If the set of acceptable portfolios is as in Figure 3.1 , the risk measure $\rho_{\mathcal{A}}$ associated with the set $\mathcal{A}$ can be illustrated as in the figure.

Based on this theory, a theorem on the relationship between risk measures and acceptable sets can be derived. The following theorem is a version of Proposition 2.2 in Föllmer and Schied [12].

Theorem 3.7 *Let ρ be a convex risk measure with acceptance set $\mathcal{A}_\rho$. Then:*

(i) $\rho_{\mathcal{A}_\rho} = \rho$

(ii) *$\mathcal{A}_\rho$ is a nonempty, convex set.*

(iii) *If $X \in \mathcal{A}_\rho$ and $Y \in \mathbb{X}$ are such that $X \leq Y$, then $Y \in \mathcal{A}_\rho$*

(iv) *If ρ is a coherent risk measure, then $\mathcal{A}_\rho$ is a convex cone.*

Conversely, let $\mathcal{A}$ be a nonempty, convex subset of $\mathbb{X}$. Let $\mathcal{A}$ be such that if $X \in \mathcal{A}$ and $Y \in \mathbb{X}$ satisfy $X \leq Y$, then $Y \in \mathcal{A}$. Then, the following holds:

(*v*) *$\rho_{\mathcal{A}}$ is a convex risk measure.*

(*vi*) *If $\mathcal{A}$ is a convex cone, then $\rho_{\mathcal{A}}$ is a coherent risk measure.*

(*vii*) *$\mathcal{A} \subseteq \mathcal{A}_{\rho_{\mathcal{A}}}$.*

Proof: $\diamondsuit$

(*i*) For any $X \in \mathcal{A}_\rho$

$$\begin{aligned}
\rho_{\mathcal{A}_\rho}(X) &= \inf\{m \in \mathbb{R} : m + X \in \mathcal{A}_\rho\} \\
&= \inf\{m \in \mathbb{R} : m + X \in \{Y \in \mathbb{X} : \rho(Y) \leq 0\}\} \\
&= \inf\{m \in \mathbb{R} : \rho(m + X) \leq 0\} \\
&= \inf\{m \in \mathbb{R} : \rho(X) - m \leq 0\} \\
&= \inf\{m \in \mathbb{R} : \rho(X) \leq m\} \\
&= \rho(X)
\end{aligned}$$

where we have used the definition of a convex risk measure (Definition 3.1) and an acceptance set (Definition 3.4).

(*ii*) $\mathcal{A}_\rho \neq \emptyset$ because $X = 0 \in \mathcal{A}_\rho$. $\mathcal{A}_\rho$ is a convex set from property 9 of Theorem 2.24 (since ρ is a convex function).

(*iii*) Since $X \in \mathcal{A}_\rho$, $\rho(X) \leq 0$, but because $Y \in \mathbb{X}$ is such that $X \leq Y$, $\rho(Y) \leq \rho(X)$ (from the definition of a convex risk measure). Hence

$$\rho(Y) \leq \rho(X) \leq 0.$$

So $Y \in \mathcal{A}_\rho$ (from the definition of an acceptance set).

(*iv*) Let ρ be a coherent risk measure, and let $X, Y \in \mathcal{A}_\rho$ and $\alpha, \beta \geq 0$. Then, from the positive homogeneity and subadditivity of coherent risk measures (see Definition 3.2), in addition to the definition of $\mathcal{A}_\rho$

$$\rho(\alpha X + \beta Y) \leq \alpha\rho(X) + \beta\rho(Y) \leq \alpha \cdot 0 + \beta \cdot 0 = 0.$$

Hence $\alpha X + \beta Y \in \mathcal{A}_\rho$, so from Definition 2.7, $\mathcal{A}_\rho$ is a convex cone.

(*v*) We check Definition 3.1: $\rho_{\mathcal{A}} : \mathbb{X} \to \mathbb{R}$. Also, for $0 \leq \lambda \leq 1$, $X, Y \in \mathbb{X}$

$$\begin{aligned}
\rho_{\mathcal{A}}(\lambda X + (1-\lambda)Y) &= \inf\{m \in \mathbb{R} : m + \lambda X + (1-\lambda)Y \in \mathcal{A}\} \\
&\leq \lambda \inf\{m \in \mathbb{R} : m + X \in \mathcal{A}\} \qquad (3.2) \\
&\quad + (1-\lambda)\inf\{m \in \mathbb{R} : m + Y \in \mathcal{A}\} \\
&= \lambda\rho_{\mathcal{A}}(X) + (1-\lambda)\rho_{\mathcal{A}}(Y)
\end{aligned}$$

where the inequality follows because $\lambda\rho_{\mathcal{A}}(X) + (1-\lambda)\rho_{\mathcal{A}}(Y) = K + L$, is a real number which will make the portfolio become acceptable because

$$\begin{aligned}(K+L) + (\lambda X + (1-\lambda)Y) &= (K+\lambda X) + (L + (1-\lambda)Y)\\ &= \lambda(\inf\{m \in \mathbb{R} : m + X \in \mathcal{A}\} + X) +\\ &\quad (1-\lambda)(\inf\{m \in \mathbb{R} : m + Y \in \mathcal{A}\} + Y) \in \mathcal{A}\end{aligned}$$

since $\mathcal{A}$ is convex (see Figure 3.2). In addition, if $X, Y \in \mathbb{X}, X \leq Y$

$$\begin{aligned}\rho_{\mathcal{A}}(X) &= \inf\{m \in \mathbb{R} : m + X \in \mathcal{A}\}\\ &\geq \inf\{m \in \mathbb{R} : m + Y \in \mathcal{A}\}\\ &= \rho_{\mathcal{A}}(Y)\end{aligned}$$

since $X \leq Y$. Finally, for $k \in \mathbb{R}$ and $X \in \mathbb{X}$

$$\begin{aligned}\rho_{\mathcal{A}}(X+k) &= \inf\{m \in \mathbb{R} : m + X + k \in \mathcal{A}\}\\ &= \inf\{s - k \in \mathbb{R} : s + X \in \mathcal{A}\}\\ &= \inf\{s \in \mathbb{R} : s + X \in \mathcal{A}\} - k\\ &= \rho_{\mathcal{A}}(X) - k.\end{aligned}$$

Hence, $\rho_{\mathcal{A}}$ is a convex risk measure.

(vi) From (v), all that remains to show is positive homogeneity. For $\alpha > 0$

$$\begin{aligned}\rho_{\mathcal{A}}(\alpha X) &= \inf\{m \in \mathbb{R} : m + \alpha X \in \mathcal{A}\}\\ &= \inf\{m \in \mathbb{R} : \alpha(\frac{m}{\alpha} + X) \in \mathcal{A}\}\\ &= \inf\{m \in \mathbb{R} : \frac{m}{\alpha} + X \in \mathcal{A}\}\\ &= \inf\{\alpha k \in \mathbb{R} : k + X \in \mathcal{A}\}\\ &= \alpha \inf\{k \in \mathbb{R} : k + X \in \mathcal{A}\}\\ &= \alpha\rho_{\mathcal{A}}(X)\end{aligned}$$

where we have used that $\mathcal{A}$ is a convex cone in equality number three. Hence, $\rho_{\mathcal{A}}$ is a coherent risk measure.

(vii) Note that $\mathcal{A}_{\rho_{\mathcal{A}}} = \{X \in \mathbb{X} : \rho_{\mathcal{A}}(X) \leq 0\} = \{X \in \mathbb{X} : \inf\{m \in \mathbb{R} : m + X \in \mathcal{A}\} \leq 0\}$.

Let $X \in \mathcal{A}$, then $\inf\{m \in \mathbb{R} : m + X \in \mathcal{A}\} \leq 0$, since $m = 0$ will suffice, since $X \in \mathcal{A}$. Hence $X \in \mathcal{A}_{\rho_{\mathcal{A}}}$.

$\square$

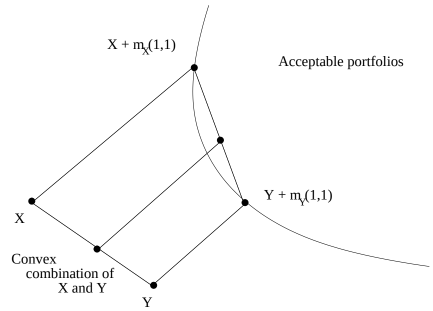

Figure 3.2: Illustration of proof of Theorem 3.7 part (v).

3.2 A dual characterization of convex risk measures

We want to use the theory presented in Chapter 2 to derive a dual characterization of a convex risk measure ρ. Therefore, let V be a vector space that is paired with the vector space $\mathbb{X}$ of financial positions (in the sense of Definition 2.33). For instance, if $\mathbb{X}$ is given a Hausdorff topology, so that it becomes a topological vector space (for definitions of these terms, see Pedersen's *Analysis Now*, [28]), V can be the set of all continuous linear functionals from $\mathbb{X}$ into $\mathbb{R}$, as in Frittelli and Gianin [14].

Using the theory presented in Chapter 2, and in particular Theorem 2.39, a dual characterization of a convex risk measure ρ can be derived. The following Theorem 3.8 is Theorem 6 in Frittelli and Gianin [14]. In Theorem 3.8, ρ^* denotes the conjugate of ρ in the sense of Definition 2.36, ρ^{**} is the biconjugate of ρ as in Definition 2.37, and $\langle \cdot, \cdot \rangle$ is a pairing as in Definition 2.31.

Theorem 3.8 *Let $\rho : \mathbb{X} \to \mathbb{R}$ be a convex risk measure. Assume in addition that ρ is lower semi-continuous. Then $\rho = \rho^{**}$. Hence for each $X \in \mathbb{X}$*

$$\begin{aligned} \rho(X) &= \sup\{\langle X, v\rangle - \rho^*(v) : v \in V\} \\ &= \sup\{\langle X, v\rangle - \rho^*(v) : v \in \mathrm{dom}(\rho^*)\} \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ is a pairing between $\mathbb{X}$ and V.

Proof: $\diamondsuit$

Since ρ is a convex risk measure, it is a convex function (see Definition 3.1 and Definition 2.11). Hence the convex hull of ρ is equal to ρ, i.e. $\mathrm{co}(\rho) = \rho$ (see Definition 2.21). In addition, since ρ is lower semi-continuous and always greater than $-\infty$, ρ is closed (see comment after Definition 2.22), so $\mathrm{cl}(\rho) = \rho$. Therefore

$$\mathrm{cl}(\mathrm{co}(\rho)) = \mathrm{cl}(\rho) = \rho.$$

But Theorem 2.39 says that $\rho^{**} = \mathrm{cl}(\mathrm{co}(\rho))$, hence $\rho = \rho^{**}$.

The second to last equation in the theorem follows directly from the definition of ρ^{**} (Definition 2.37), while the last equation follows because the supremum cannot be achieved when $\rho^* = +\infty$. $\square$

Theorem 3.8 is quite abstract, but by choosing a specific set of paired spaces, $\mathbb{X}$ and V, some nice results can be derived.

The next theorem is due to Föllmer and Schied [11]. Consider the paired spaces $\mathbb{X} = \mathbb{R}^n, V = \mathbb{R}^n$ with the standard Euclidean inner product, denoted $\cdot$, as pairing (see Example 2.34).

In the following, let $(\Omega, \mathcal{F})$ be a measurable space and let $\mathcal{P}$ denote the set of all probability measures over Ω. Also, let ρ^* be the conjugate of ρ, as in Definition 2.36.

Theorem 3.9 *Assume that Ω is finite. Then, any convex risk measure $\rho : \mathbb{X} \to \mathbb{R}$ can be represented in the form*

$$\rho(X) = \sup_{Q \in \mathcal{P}} \{E_Q[-X] - \alpha(Q)\} \tag{3.3}$$

where $E_Q[\cdot]$ denotes the expectation with respect to Q and $\alpha : \mathcal{P} \to (-\infty, \infty]$ is a "penalty function" which is convex and closed. Actually, $\alpha(Q) = \rho^*(-Q)$ for all $Q \in \mathcal{P}$.

The proof of this theorem is from *Convex Risk Measures for Portfolio Optimization and Concepts of Flexibility* by Luthi and Doege [25], some details of that proof have been filled out.

Proof: $\diamondsuit$

To show that $\rho : \mathbb{X} \to \mathbb{R}$ (as in Theorem 3.9) is a convex risk measure we check Definition 3.1: Let $\lambda \in [0,1], m \in \mathbb{R}, X, Y \in \mathbb{X}$.

(i) :

$$\begin{aligned}
\rho(\lambda X + (1-\lambda)Y) &= \sup_{Q \in \mathcal{P}} \{E_Q[-(\lambda X + (1-\lambda)Y)] - \alpha(Q)\} \\
&= \sup_{Q \in \mathcal{P}} \{\lambda E_Q[-X] + (1-\lambda)E_Q[-Y] - \alpha(Q)\} \\
&\leq \lambda \sup_{Q \in \mathcal{P}} \{E_Q[-X] - \alpha(Q)\} \\
&\quad + (1-\lambda) \sup_{Q \in \mathcal{P}} \{E_Q[-Y] - \alpha(Q)\} \\
&= \lambda \rho(X) + (1-\lambda)\rho(Y).
\end{aligned}$$

(ii) : Assume $X \leq Y$. Then $-X \geq -Y$, so

$$\begin{aligned}
\rho(X) &= \sup_{Q \in \mathcal{P}} \{E_Q[-X] - \alpha(Q)\} \\
&\geq \sup_{Q \in \mathcal{P}} \{E_Q[-Y] - \alpha(Q)\} \\
&= \rho(Y).
\end{aligned}$$

(iii) :

$$\begin{aligned}
\rho(X+m\mathbf{1}) &= \sup_{Q\in\mathcal{P}}\{E_Q[-(X+m\mathbf{1})]-\alpha(Q)\} \\
&= \sup_{Q\in\mathcal{P}}\{E_Q[-X]-mE_Q[\mathbf{1}]-\alpha(Q)\} \\
&= \sup_{Q\in\mathcal{P}}\{E_Q[-X]-m-\alpha(Q)\} \\
&= \sup_{Q\in\mathcal{P}}\{E_Q[-X]-\alpha(Q)\}-m \\
&= \rho(X)-m.
\end{aligned}$$

Hence, ρ is a convex risk measure.

So, assume that ρ is a convex risk measure. The conjugate function of ρ, denoted ρ^*, is then defined as $\rho^*(v)=\sup_{X\in\mathbb{X}}\{v\cdot X-\rho(X)\}$ (where $\cdot$ denotes Euclidean inner product) for all $v\in V(=\mathbb{R}^n)$. Fix an $X\in\mathbb{X}$ and consider $Y_m := X+m\mathbf{1}\in\mathbb{X}$ for an arbitrary $m\in\mathbb{R}$. Then

$$\rho^*(v)\geq \sup_{m\in\mathbb{R}}\{v\cdot Y_m-\rho(Y_m)\}$$

because $\{Y_m\}_{m\in\mathbb{R}}\subset\mathbb{X}$. This means that

$$\begin{aligned}
\rho^*(v) &\geq \sup_{m\in\mathbb{R}}\{v\cdot(X+m\mathbf{1})-\rho(X+m\mathbf{1})\} \\
&= \sup_{m\in\mathbb{R}}\{m(v\cdot\mathbf{1}+1)\}+v\cdot X-\rho(X)
\end{aligned}$$

where the equality follows from the translation invariance of ρ (see Definition 3.1). The first term in the last expression is only finite if $v\cdot\mathbf{1}+1=0$, (where $\mathbf{1}=(1,1,\dots,1)\in\mathbb{R}^n$) i.e. if $\sum_{i=1}^n v_i=-1$ (if not, one can make the first term go towards $+\infty$ by letting m go towards either $+\infty$ or $-\infty$). It is now proved that in order for $\rho^*(v)<+\infty$, $\sum_{i=1}^n v_i=-1$ must hold.

Again, consider an arbitrary, but fixed $X\in\mathbb{X}, X\geq 0$ (here, $X\geq 0$ means component-wise). Then, for all $\lambda\geq 0$, we have $\lambda X\geq 0$, and $\lambda X\in\mathbb{X}$, and hence $\rho(\lambda X)\leq\rho(0)$, from the monotonicity of ρ (again, see Definition 3.1). Therefore, by the same type of arguments as above

$$\rho^*(v)\geq\sup_{\lambda\geq 0}\{v\cdot\lambda X-\rho(\lambda X)\}\geq\sup_{\lambda\geq 0}\{v\cdot(\lambda X)\}-\rho(0).$$

Here, $\rho^*(v)$ is only finite if $v\cdot X\leq 0$ for all $X\geq 0$, hence $v\leq 0$.

We then get that the conjugate ρ^* is reduced to

$$\rho^*(v)=\begin{cases}\sup_{X\in\mathbb{X}}\{v\cdot X-\rho(X)\} \text{ where } v\cdot\mathbf{1}=-1 \text{ and } v\leq 0\\ +\infty \text{ otherwise .}\end{cases}$$

Now, define $\alpha(Q) = \rho^*(-Q)$ for all $Q \in \mathcal{P}$. From Theorem 3.8, $\rho = \rho^{**}$. But

$$\begin{aligned}
\rho^{**}(X) &= \sup_{v \in V}\{v \cdot X - \rho^*(v)\} \\
&= \sup_{Q \in \mathcal{P}}\{(-Q) \cdot X - \alpha(Q)\} \\
&= \sup_{Q \in \mathcal{P}}\{\sum_{i=1}^{n} Q_i(-X_i) - \alpha(Q)\} \\
&= \sup_{Q \in \mathcal{P}}\{E_Q[-X] - \alpha(Q)\}
\end{aligned}$$

where Q_i, X_i denote the i'th components of the vectors Q, X respectively. Hence $\rho(X) = \rho^{**}(X) = \sup_{Q \in \mathcal{P}}\{E_Q[-X] - \alpha(Q)\}$. □

Theorem 3.9 says that any convex risk measure $\rho : \mathbb{R}^n \to \mathbb{R}$ is the expected value of the negative of a contingent claim, $-X$, minus a penalty function, $\alpha(\cdot)$, under the worst case probability. Note that we consider the expectation of $-X$, not X, since losses are negative in our context.

We already know that the penalty function α in Theorem 3.9 is of the form $\alpha(Q) = \rho^*(-Q)$. Actually, Luthi and Doege [25] proved that it is possible to derive a more intuitive representation of α (see Corollary 2.5 in [25]).

Theorem 3.10 *Let $\rho : \mathbb{R}^n \to \mathbb{R}$ be a convex risk measure, and let $\mathcal{A}_\rho$ be its acceptance set (in the sense of Definition 3.4). Then, Theorem 3.9 implies that $\rho(X) = \sup_{Q \in \mathcal{P}}\{E_Q[-X] - \alpha(Q)\}$, where $\alpha : \mathcal{P} \to \mathbb{R}$ is a penalty function. Then, α is of the form*

$$\alpha(Q) = \sup_{X \in \mathcal{A}_\rho}\{E_Q[-X]\}.$$

Proof: ♢

It suffices to prove that for all $Q \in \mathcal{P}$

$$\rho^*(-Q) = \sup_{X \in \mathbb{X}}\{E_Q[-X] - \rho(X)\} = \sup_{X \in \mathcal{A}_\rho}\{E_Q[-X]\} \tag{3.4}$$

since we know that $\alpha(Q) = \rho^*(-Q)$.

For all $X \in \mathcal{A}_\rho$, $\rho(X) \leq 0$ (see Definition 3.4), so $E_Q[-X] - \rho(X) \geq E_Q[-X]$. Hence, since $\mathcal{A}_\rho \subseteq X$

$$\rho^*(-Q) \geq \sup_{X \in \mathcal{A}_\rho}\{E_Q[X] - \rho(X)\} \geq \sup_{X \in \mathcal{A}_\rho}\{E_Q[-X]\}.$$

To prove the opposite inequality, and hence to prove equation (3.4), assume for contradiction that there exists $Q \in \mathcal{P}$ such that $\rho^*(-Q) > \sup_{X \in \mathcal{A}_\rho}\{E_Q[-X]\}$. From the definition of supremum, there exists a $Y \in \mathbb{X}$ such that

$$E_Q[-Y] - \rho(Y) > E_Q[-X] \text{ for all } X \in \mathcal{A}_\rho.$$

Note that $Y + \rho(Y)\mathbf{1} \in \mathcal{A}_\rho$ since $\rho(Y + \rho(Y)\mathbf{1}) = \rho(Y) - \rho(Y) = 0$. Therefore $E_Q[-Y] - \rho(Y) > E_Q[-(Y + \rho(Y)\mathbf{1})] = E_Q(-Y) + \rho(Y)E_Q[-\mathbf{1}] = E_Q(-Y) - \rho(Y)$, which is a contradiction. Hence, the result is proved. □

Together, Theorem 3.9 and Theorem 3.10 provide a good understanding of convex risk measures in $\mathbb{R}^n$: Any convex risk measure $\rho : \mathbb{R}^n \to \mathbb{R}$ can be written in the form $\rho(X) = \sup_{Q \in \mathcal{P}}\{E_Q[-X] - \alpha(Q)\}$, where $\alpha(Q) = \sup_{X \in \mathcal{A}_\rho}\{E_Q[-X]\}$ and $\mathcal{A}_\rho$ is the acceptance set of ρ. But what about infinite-dimensional spaces? Can a similar representation of ρ be derived? This is partially answered in the following Theorem 3.11, which is Theorem 2.2 in Ruszczynski and Shapiro [38], modified slightly to our setting.

First, let's introduce the setting of Theorem 3.11. Let $(\Omega, \mathcal{F})$ be a measurable space and let $\bar{\mathbb{V}}$ be the vector space of all finite signed measures on $(\Omega, \mathcal{F})$. For each $v \in \bar{\mathbb{V}}$ we know that there is a Jordan decomposition of v, so $v = v^+ - v^-$. Let $|v| = v^+ + v^-$. Let $\mathbb{X}$ be a vector space of measurable functions $X : \Omega \to \mathbb{R}$. Also, let $\mathbb{X}_+ = \{X \in \mathbb{X} : X(\omega) \geq 0 \; \forall \; \omega \in \Omega\}$. This gives a partial order relation on X, so $X \leq Y$ where $X, Y \in \mathbb{X}$ means $Y - X \in \mathbb{X}_+$.

Let $V \subseteq \bar{\mathbb{V}}$ be the measures $v \in \mathbb{V}$ such that $\int_\Omega |X(\omega)|d|v| < +\infty$ for all $X \in \mathbb{X}$. V is a vector space because of uniqueness of the Jordan decomposition and linearity of integrals. For example: If $v, w \in V$ then $|v+w| = (v+w)^+ + (v+w)^- = (v^+ + w^+) + (v^- + w^-) = |v| + |w|$ by uniqueness of the decomposition, hence $\int_\Omega |X|d|v + w| = \int_\Omega |X|d|v| + \int_\Omega |X|d|w| < +\infty$. Define the pairing $\langle X, v\rangle = \int_\Omega X(\omega)dv(\omega)$. Let $V_- \subseteq V$ be the non-positive measures in V and let $\mathcal{P}$ be the set of probability measures in V.

Assume the following:

$$\text{(A): If } v \notin V_- = \{v \in V : v \leq 0\}, \text{ then there exists an } X' \in \mathbb{X}_+ \text{ such that } \langle X', v\rangle > 0.$$

Now, let $\mathbb{X}$ and V have topologies so that they become paired spaces under the pairing $\langle \cdot, \cdot \rangle$.

For example, let $\mathbb{X} = L^p(\Omega, \mathcal{F}, P)$ where P is a measure, and let V be as above. Each signed measure $v \in V$ can be decomposed so that $v = v_P + v'$, where v_P is absolutely continuous with respect to P (i.e. $P(A) = 0 \Rightarrow v_P(A) = 0$). Then, $dv_P = MdP$, where $M : \Omega \to \mathbb{R}$ is the Radon-Nikodym density of v w.r.t. P. Look at $V' := \{v \in V : \int_\Omega |X(\omega)|d|v_P| < +\infty\} \subseteq V$. This is a vector space for the same reasons that V is a vector space. Then, any signed measure $v \in V'$ can be identified by the Radon-Nikodym derivative of v_P w.r.t. P, that is by M. Actually, $M \in L^q(\Omega, \mathcal{F}, P)$, where $\frac{1}{p} + \frac{1}{q} = 1$, because $\int_\Omega |M|^q dP = \int_\Omega |M|^{q-1} d|v_P| < +\infty$. Hence, each signed measure $v \in V'$ is identified in L^q by its Radon-Nikodym density with respect to P.

Note that the pairing defined above actually is the usual bilinear form between L^p and L^q since for $p \in L^p, q \in L^q$

$$\begin{aligned} \langle p, q \rangle &= \int_\Omega p(\omega)q(\omega)dP(\omega) \\ &= \int_\Omega p(\omega)M(\omega)dv(\omega) \\ &= \int_\Omega p(\omega)dv(\omega) \end{aligned} \tag{3.5}$$

where the second equality follows from that any $q \in L^q$ can be viewed as a Radon-Nikodym derivative w.r.t. P for some signed measure $v \in V'$ and the third equality from the definition of a Radon-Nikodym derivative. See Example 2.35.

In the following theorem monotonicity and translation invariance mean the same as in Definition 3.1.

Theorem 3.11 *Let $\mathbb{X}$ be a vector space paired with the space V, both of the form above. Let $\rho : \mathbb{X} \to \mathbb{R}$ be a proper, lower semi-continuous convex function. From Theorem 2.39 the following holds: $\rho(X) = \sup\{\langle X, v\rangle - \rho^*(v) : v \in \mathrm{dom}(\rho^*)\}$.*

We then have the following results:

(i) ρ is monotone $\iff$ All $v \in \mathrm{dom}(\rho^*)$ are such that $v \leq 0$

(ii) ρ is translation invariant $\iff$ $v(\Omega) = -1$ for all $v \in \mathrm{dom}(\rho^*)$.

Hence, if ρ is a convex risk measure (so monotonicity and translation invariance hold), then $v \in \mathrm{dom}(\rho^*)$ implies that $Q := -v \in \mathcal{P}$ and

$$\begin{aligned} \rho(X) &= \sup_{Q \in \mathcal{P}} \{\langle X, -Q\rangle - \rho^*(-Q)\} \\ &= \sup_{Q \in \mathcal{P}} \{\langle -X, Q\rangle - \rho^*(-Q)\} \\ &= \sup_{Q \in \mathcal{P}} \{E_Q[-X] - \alpha(Q)\} \end{aligned}$$

where $\alpha(Q) := \rho^*(-Q)$ is a penalty function and the pairing, i.e. the integral, is viewed as an expectation.

Proof: $\diamondsuit$

(i) : Assume monotonicity of ρ. We want to show that $\rho^*(v) = +\infty$ for all $v \notin V_-$. From assumption (A), $v \notin V_- \Rightarrow$ there exists $X' \in X_+$ such that $\langle X', v\rangle > 0$. Take $X \in \mathrm{dom}(\rho)$, so that $\rho(X) < +\infty$ and consider $Y_m = X + mX'$. For $m \geq 0$, monotonicity implies that $\rho(X) \geq \rho(Y_m)$ (since $Y_m = X + mX' \geq X$ because $X' \geq 0$). Hence

$$\begin{aligned}
\rho^*(v) &\geq \sup_{m\in\mathbb{R}_+} \{\langle Y_m, v\rangle - \rho(Y_m)\} \\
&= \sup_{m\in\mathbb{R}_+} \{\langle X, v\rangle + m\langle X', v\rangle - \rho(X + mX')\} \\
&\geq \sup_{m\in\mathbb{R}_+} \{\langle X, v\rangle + m\langle X', v\rangle - \rho(X)\}
\end{aligned}$$

where the last inequality uses the monotonicity. But since $\langle X', v\rangle > 0$, by letting $m \to +\infty$, one gets $\rho^*(v) = +\infty$ (since $X \in \mathrm{dom}(\rho)$, so $\rho(X) < +\infty$, and $\langle X, v\rangle$ is bounded since $\langle X, \cdot\rangle$ and $\langle \cdot, v\rangle$ are bounded linear functionals).

Hence, monotonicity implies that $\rho^*(v) = +\infty$, unless $v \leq 0$, so all $v \in \mathrm{dom}(\rho^*)$ are such that $v \leq 0$.

Conversely, Assume that all $v \in \mathrm{dom}(\rho^*)$ are such that $v \leq 0$. Take $X, Y \in \mathbb{X}, Y \leq X$ (i.e. $X - Y \geq 0$). Then $\langle Y, v\rangle \geq \langle X, v\rangle$ (from the linearity of the pairing). Since $\rho(X) = \sup_{v\in\mathrm{dom}(\rho^*)}\{\langle X, v\rangle - \rho^*(v)\}$, it follows that $\rho(X) \leq \rho(Y)$. Hence (i) is proved.

(ii) : Assume translation invariance. Let $\mathbf{1} : \Omega \to \mathbb{R}$ denote the random variable constantly equal to 1, so $\mathbf{1}(\omega) = 1 \ \forall\ \omega \in \Omega$. This random variable is clearly measurable, so $\mathbf{1} \in \mathbb{X}$. For $X \in \mathrm{dom}(\rho)$

$$\begin{aligned}
\rho^*(v) &\geq \sup_{m\in\mathbb{R}}\{\langle X + m\mathbf{1}, v\rangle - \rho(X + m\mathbf{1})\} \\
&= \sup_{m\in\mathbb{R}}\{m\langle \mathbf{1}, v\rangle + \langle X, v\rangle - \rho(X) + m\} \\
&= \sup_{m\in\mathbb{R}}\{m(v(\Omega) + 1) + \langle X, v\rangle - \rho(X)\}.
\end{aligned}$$

Hence, $\rho^*(v) = +\infty$, unless $v(\Omega) = \langle \mathbf{1}, v\rangle = -1$.

Conversely, if $v(\Omega) = -1$, then $\langle X+m\mathbf{1}, v\rangle = \langle X, v\rangle - v(\Omega)m = \langle X, v\rangle - m$. (where the first equality follows from linearity of the pairing). Hence, translation invariance follows from $\rho(X) = \sup_{v\in\mathrm{dom}(\rho^*)}\{\langle X, v\rangle - \rho^*(v)\}$.

□

Föllmer and Schied proved a version of Theorem 3.11 for $\mathbb{X} = L^\infty(\Omega, \mathcal{F}, P)$, $V = L^1(\Omega, \mathcal{F}, P)$ in *Convex measures of risk and trading constraints* (see [11]). In this case, it is sufficient to assume that the acceptance set $\mathcal{A}_\rho$ of ρ is weak*-closed (i.e., closed with respect to the coarsest topology that makes all the linear functionals originating from the inner product, $\langle \cdot, v\rangle$ continuous) in order to derive a representation of ρ as above.

This chapter has introduced convex risk measures and proved dual representation theorems for such measures using the conjugate duality theory of Chapter 2. Hence, the chapter illustrates one application of duality in mathematical finance. Convex risk measures, and their dual representation theorems will be used in a utility maximization problem in Section 5.9. The following chapter will introduce a model for the financial market, and also some central terms of mathematical finance.

Chapter 4

Mathematical finance

Mathematical finance is the mathematical theory of financial markets. In this field, one tries to model the financial market in order to learn more about how it works. Mathematical finance is used widely in "the real world", and uses plenty of advanced mathematics. The goal of the following chapter is to introduce a mathematical model of the financial market and define terms characterizing the market such as equivalent martingale measures, arbitrage and completeness. This is done in Section 4.2. The model introduced in this chapter will be used in the remaining part of thesis.

Section 4.1 gives a brief summary of Itô integration with respect to Brownian motion, and also defines the important term martingale. Both Itô integration and martingales will be crucial in the model of the financial market defined in Section 4.2. Section 4.3 simplifies the situation by considering the financial market in the case where the scenario space Ω is finite and the time t is discrete. In this case, the fundamental theorem of mathematical finance (Theorem 4.3) can be proved using linear programming duality, and this is the main goal of Section 4.3. This is an application of duality methods to mathematical finance.

This chapter is based on Kramkov and Schachermayer [22], [21], Øksendal [27], Pliska [32], and Schachermayer [40].

4.1 Martingales and Itô integration with respect to Brownian motion

Let $(\Omega, \mathcal{F}, P)$ be a given probability space, so Ω is a scenario space, $\mathcal{F}$ is a filtration on Ω and P is a probability measure on the measurable space $(\Omega, \mathcal{F})$. A *filtration* is a family of σ-algebras $(\mathcal{G}_t)_{0 \leq t \leq T}$ such that $\mathcal{G}_s \subseteq \mathcal{G}_t$ for $s \leq t$, where the time $t \in [0, T]$ and the terminal time $T \leq \infty$. The space $(\Omega, \mathcal{F}, (\mathcal{G}_t)_t, P)$ is called a *filtered probability space*. A *stochastic process* is a family of random variables on the (same) probability space $(\Omega, \mathcal{F}, P)$. Filtrations and stochastic processes will be essential in the following. Another term which will become very useful is martingales.

Definition 4.1 *(Martingale) A stochastic process $(M_t)_{0 \leq t \leq T}$ on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{G}_t), P)$ is called a* martingale *with respect to the filtration $(\mathcal{G}_t)_{0 \leq t \leq T}$ if:*

(i) M_t is $\mathcal{G}_t$-measurable for all t.

(ii) $E_P[|M_t|] < +\infty$ for all t.

(iii) $E_P[M_t|\mathcal{G}_u] = M_u$ for all $u \leq t$.

Condition (iii) can be interpreted: given all that as happened up to time $u < t$, including the value of the martingale at time u, one does not expect (under the P-measure) this value to change up to time t.

Note also that for any $(\mathcal{G}_t)_t$-martingale $(M_t)_{0 \leq t \leq T}$, where $\mathcal{G}_0 = \{\emptyset, \Omega\}$,

$$E_P[M_t] = E_P[M_t|\mathcal{G}_0] = M_0.$$

Here, the first equality follows from $\mathcal{G}_0$ being the trivial σ-algebra, and the second equality follows from the definition of a martingale.

Note that both of the terms $E_P[\cdot]$ and $E[\cdot]$ will be used to denote the expectation with respect to the given probability measure P. The expectation with respect to some other probability measure R will always be denoted by $E_R[\cdot]$, with the subscript.

Itô-integration is integration of a stochastic process with respect to another stochastic process. We will consider Itô-integration with respect to a specific stochastic process called *Brownian motion*. A (standard) Brownian motion $B(t)$ on probability space $(\Omega, \mathcal{F}, P)$ is a stochastic process, which satisfies the following properties:

- $B(0) = 0$
- B has independent increments, i.e. $B(t) - B(s)$ is independent of $B(u) - B(z)$ for $z < u \leq s < t$.
- Each increment $B(t) - B(s)$, $s < t$ has a normal distribution with expectation 0 and variance $t - s$.

For the precise definition and comments on the existence of Brownian motion, see Øksendal [27].

Given a Brownian motion $B = \{B(t)\}$, $t \in [0, T]$, there is a natural filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$ which is generated by B. This filtration is such that for every $t \in [0, T]$, $\mathcal{F}_t$ is the smallest σ-algebra which makes $B(t)$ measurable. It can be shown that the Brownian motion is a martingale with respect to this filtration.

Let $\mathcal{V}(0, T)$ be the class of functions $f : [0, \infty) \times \Omega \to \mathbb{R}$ such that

- $(t, \omega) \mapsto f(t, \omega)$ is $\mathcal{B} \times \mathcal{F}$-measurable, where $\mathcal{B}$ denotes the Borel σ-algebra.

- $\omega \mapsto f(t, \omega)$ is $\mathcal{F}_t$-measurable for every $t \in [0, T]$.
- $E[\int_0^T f(t, \omega)^2 dt] < \infty$.

For functions $f(t, \omega) \in \mathcal{V}(0, T)$ and $t \in [0, T]$ one can define the *Itô integral* $\int_0^t f(s, \omega) dB(s, \omega) = \lim_{n \to \infty} \int_0^t f_n(s, \omega) dB(s, \omega)$ (limit in $\mathcal{L}^2$-sense), where $(f_n)_{n=1}^{\infty}$ is a sequence of simple functions which converge to f (in an $\mathcal{L}^2$-sense). A *simple function* $g(t, \omega)$ is of the form $g(t, \omega) = \sum_{j=1}^{J} e_j(\omega) \mathbf{1}_{[t_j, t_{j+1})}(t)$, where e_j is $\mathcal{G}_{t_j}$-measurable for all $j = 1, \ldots, J$, $\mathbf{1}_A(\cdot)$ denotes the indicator function with respect to the set A and $t_0, t_1, ..., t_J$ is a partition of the interval $[0, t]$. Here $\int_0^t f_n(s, \omega) dB(s, \omega) := \sum_{j=1}^{J} e_j^n(\omega)(B(t_{j+1}, \omega) - B(t_j, \omega))$ for a simple function $f_n(s, \omega) := \sum_{j=1}^{J} e_j^n(\omega) \mathbf{1}_{[t_j, t_{j+1}]}(s)$. The existence of such an approximating sequence of functions is proved in Øksendal [27]. Note that the Itô integral is a random variable.

From this definition, one can prove properties of the Itô integral. To simplify notation, $\int_0^t f(s, \omega) dB(s, \omega)$ is sometimes denoted by $\int_0^t f(s, \omega) dB(s)$, or even $\int_0^t f dB$. Below is a list of some of the most important properties of the Itô-integral. Let $f, g \in \mathcal{V}(0, T)$, $t \in [0, T]$ and $a, b \in \mathbb{R}$ be constants.

(i) The Itô integral is linear:

$$\int_0^t (af(s, \omega) + bg(s, \omega)) dB(s) = \quad a \int_0^t f(s, \omega) dB(s) + b \int_0^t g(s, \omega) dB(s).$$

(ii) $\int_0^t f(s, \omega) dB(s)$ is $\mathcal{F}_t$-measurable for all $t \in [0, T]$.

(iii) $E[\int_0^t f(s, \omega) dB(s)] = 0$.

(iv) The Itô integral $\int_0^t f(s, \omega) dB(s)$ is a martingale with respect to the filtration $(\mathcal{G}_t)_{0 \leq t \leq T}$.

The Itô integral is very useful when modeling the financial market, as will be shown in Section 4.2.

4.2 A model of the financial market

In order to apply the theory of the previous chapters to mathematical finance, one must construct a mathematical model of the financial market. There are many possible ways to do this, but the following general model is a good foundation to build on.

The financial market model is based on a probability space $(\Omega, \mathcal{F}, P)$ consisting of a space of possible scenarios of the world Ω, a σ-algebra $\mathcal{F}$, and a probability measure P on the measurable space $(\Omega, \mathcal{F})$.

The financial market consists of $N + 1$ assets: N risky assets (stocks) and one bond (the bank). The assets each have a *price process* $S_n(t, \omega)$, for $n = 0, ..., N$, $\omega \in \Omega$ and $t \in [0, T]$, where $S_0(t, \omega)$ is the price process of the bond.

The price processes S_n, $n = 1, ..., N$, are stochastic processes. We denote by $S(t, \omega) = (S_0(t, \omega), S_1(t, \omega) ..., S_N(t, \omega))$, the composed price process of all the assets. Here, the time $t \in [0, T]$ where the final time T may be infinite. Though the final time may be infinite, one often assumes $T < \infty$ in mathematical finance.

Let $(\mathcal{F}_t)_{0 \leq t \leq T}$ be a filtration. Usually, one assumes that the price process S is adapted to the filtration $(\mathcal{F}_t)_t$ (S being *adapted* to $(\mathcal{F}_t)_t$ means that $S_n(t)$ is $\mathcal{F}_t$-measurable for all $t \in [0, T]$), and it is very common to let $(\mathcal{F}_t)_t$ be the filtration generated by the price process S. Actually, the price process is often assumed to be an $(\mathcal{F}_t)_t$-semimartingale. The precise definition of a semimartingale will not be discussed in this thesis, but the main point is that when S_n is a semimartingale, one can form the Itô-integral with respect to this stochastic process. See for example Shilling [41], or Øksendal [27], for more on filtrations and σ-algebras.

One often assumes that $S_0(t, \omega) = 1$ for all $t \in [0, T]$, $\omega \in \Omega$. This corresponds to having divided through all the other prices by S_0, and hence turning the bank into *the numeraire* of the market. The altered market is a *discounted* (or *normalized*) market. This text mainly considers discounted markets, but whether the market is discounted or not will be mentioned when necessary. To simplify notation, the price processes in the discounted market are denoted by S as well. Note that for a given probability space $(\Omega, \mathcal{F}, P)$, the price process S describe the market. Therefore the market is sometimes denoted simply by S.

An example of a security market as the one above is the model used in Øksendal [27]:

$$S_0(t, \omega) = 1 + \int_0^t r(s, \omega) S_0(s, \omega) ds \tag{4.1}$$

and for $n = 1, ..., N$

$$S_n(t, \omega) = x_n + \int_0^t \mu_n(s, \omega) ds + \int_0^t \sigma_n(s, \omega) dB(s, \omega) \tag{4.2}$$

or, written in brief form

$$dS_0(t) = r(t) S_0(t) dt, \;\; S_0(0) = 1 \tag{4.3}$$

and for $n = 1, ..., N$

$$dS_n(t) = \mu_n(t) dt + \sigma_n(t) dB(t), \;\; S_n(0) = x_n. \tag{4.4}$$

where $B(t)$ is a D-dimensional Brownian motion.

Here, the process r called *the interest rate process*, μ_n is called *the drift process* (of asset n) and σ_n is called *the volatility process* (of asset n). Note that sometimes, as above, the dependence on ω is suppressed for notational convenience. The Brownian motion B may be more than one-dimensional, in general, B is D-dimensional. Therefore, σ_n is also D-dimensional (for all n). If $D > 1$, the stochastic integral (Itô-integral) $\sigma_n(t)dB(t) := \sum_{i=1}^{D} \sigma_n^i(t)dB_i(t)$, where $\sigma_n^i(t)$ and $B_i(t)$ denotes the i'th components of $\sigma_n(t)$ and $B(t)$ respectively. This Brownian motion driven market model will be central in this text.

As mentioned, the space Ω represents the possible scenarios of the world, and the probability measure P gives the probabilities for each of the measurable subsets of Ω to occur. As in real financial markets, we are interested in modeling how information reveals itself to the investors, and that is where the filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$ is useful. This filtration represents what the investors know at time t. For instance, if $(\mathcal{F}_t)_t$ is the filtration generated by the price process S, the agents in the market know the asset prices at any time t, but nothing more. For a finite Ω and discrete time, using filtrations to model information can be explained by the bijection between partitions and σ-algebras (the σ-algebra consists of every union of elements in the partition). The nested chain of partitions is a convenient way to model how the world scenarios reveal themselves, see Figure 4.1. When Ω is infinite, the correspondence between partitions and filtrations disappears, so the use of a filtration is seen as a generalization of the finite Ω case.

Trading of assets is essential in real financial markets, so one must also model the concept of a portfolio. A *portfolio* (or a *trading strategy*) is an $N+1$-dimensional, predictable S-integrable stochastic process $H(t,\omega) = (H_0(t,\omega), H_1(t,\omega), ..., H_N(t,\omega))$. $H_n(t,\omega)$ represents the amount of asset n held by the investor at time t in state $\omega \in \Omega$. In general, a *predictable* stochastic process is defined as a stochastic process which is measurable (when the process is viewed as a function from $\mathbb{R}_+ \times \Omega$ into $\mathbb{R}$) with respect to a special σ-algebra $\mathcal{PR}$, called the *predictable σ-algebra*, on the product space $\mathbb{R}_+ \times \Omega$. $\mathcal{PR}$ is constructed from predictable rectangles. A *predictable rectangle* is a subset of $\mathbb{R}_+ \times \Omega$ of the form $(s,t] \times F$, where $s < t$ and $F \in \mathcal{F}_s$, or of the form $\{0\} \times F_0$, where $F_0 \in \mathcal{F}_0$. $\mathcal{PR}$ is the σ-algebra generated by the collection of all predictable rectangles. In discrete time, being predictable means that $H(t,\cdot)$ is $\mathcal{F}_{t-1}$-measurable for all $t \in \{1,2,...,T\}$, where $H_n(t,\omega)$ represents the amount of asset number n held from time $t-1$ to time t. Hence, predictability means that the investor has to choose how much to hold of asset n between times $t-1$ and t based on what she knows at time $t-1$, that is $\mathcal{F}_{t-1}$. The abstract notion of a predictable process is a generalization of this concept. Note that all predictable processes are adapted. However, it turns out that it does not matter whether one considers predictable or adapted trading strategies as long as the price process S is not a jump processes.

A portfolio H is called *admissible* if there exists a constant $C < 0$ such that $\int_0^t H(u)dS(u) := \int_0^t H(u) \cdot dS(u) > C$ P-almost surely for all $t \in [0,T]$ (where the stochastic processes $H(t,\omega) = (H_0(t,\omega), H_1(t,\omega), ..., H_N(t,\omega))$ and $S(t,\omega) = (S_0(t,\omega), S_1(t,\omega), ..., S_N(t,\omega))$ are viewed as vectors in $\mathbb{R}^{N+1}$, and $\cdot$ denotes the standard Euclidean inner product). The family of admissible trading strategies will be denoted by $\mathcal{H}$.

The *value of a portfolio* H at time t is the random variable denoted by $X^H(t)$

and defined

$$X^H(t,\omega) := \sum_{n=0}^{N} H_n(t,\omega) S_n(t,\omega). \tag{4.5}$$

That is, the value of the investor's portfolio at a time t is the sum of the amount she owns of each asset, times the price of that asset at time t. When the underlying portfolio is clear, we sometimes denote $X = X^H$, to simplify notation.

A portfolio H is called *self-financing* if it is an S-integrable process which satisfies

$$X^H(t) = X^H(0) + \int_0^t H(u) \cdot dS(u) \text{ for all } 0 \leq t \leq T$$

A portfolio being self-financing means that no money is taken in or out of the system. That is, all trading of assets after time $t = 0$ is financed by the price changes in the market, which affect the value of the portfolio.

An *arbitrage* in a financial market is, roughly speaking, a riskless way of making money. More formally, one says that the market $(S_t)_{0 \leq t \leq T}$ has an arbitrage if there exists an admissible portfolio H such that

(i) $X^H(0) = 0$,

(ii) $X^H(T) \geq 0$ almost surely,

(iii) $P(X^H(T) > 0) > 0$.

If there was an arbitrage in the market, all investors would want to execute the arbitrage, and (possibly) cash in a riskless profit. Economic principles then suggest that the demand for the arbitrage opportunity would cause the prices to adjust such that the arbitrage disappears. Therefore, an arbitrage is a sign of lack of equilibrium in the market.

For a potential investor, it is interesting to determine whether a market has an arbitrage or not. In order to do this, the notion of equivalent martingale measure is useful.

An *equivalent martingale measure* is a probability measure Q on the measurable space $(\Omega, \mathcal{F})$ such that

(i) Q is equivalent to P, and

(ii) The price process S is a martingale with respect to Q (and w.r.t. the filtration $(\mathcal{F}_t)_t$).

Denote by $\mathcal{M}^e(S)$ the *set of all equivalent martingale measures* for the market S. Then one has the following theorem, which is Lemma 12.1.6 in Øksendal [27] (for the proof of this theorem, see [27]).

Theorem 4.2 *Let the market be defined as in equations (4.3) and (4.4). If the set of equivalent martingale measures in nonempty, i.e. $\mathcal{M}^e(S) \neq \emptyset$, then there is no arbitrage.*

If Ω is finite and time is discrete, one can attempt to find equivalent martingale measures by solving the set of linear equations (4.6) for Q

$$E_Q[S(t)|\mathcal{F}_u] \quad = \quad S(u) \text{ for all } 0 \leq u \leq t \leq T \tag{4.6}$$

When this system is solved, one must check if any of the solutions are probability measures on $(\Omega, \mathcal{F})$ equivalent to P (i.e. such that P and Q have the same null sets).

It turns out that one can have either $\mathcal{M}^e(S) = \emptyset$, $\mathcal{M}^e(S) = \{Q\}$ (a one-element set) or $|\mathcal{M}^e(S)| = \infty$. Why is this? Assume there are at least two distinct equivalent martingale measures $Q, Q^* \in \mathcal{M}^e(S)$. Let $\lambda \in (0,1)$, then $\bar{Q} = \lambda Q + (1-\lambda)Q^* \in \mathcal{M}^e(S)$: Clearly, $\bar{Q}$ is a probability measure. It is a martingale measure because $E_{\bar{Q}}[S_t|\mathcal{F}_u] = E_{\lambda Q+(1-\lambda)Q^*}[S_t|\mathcal{F}_u] = \lambda E_Q[S_t|\mathcal{F}_u] + (1-\lambda)E_{Q^*}[S_t|\mathcal{F}_u] = \lambda S_u + (1-\lambda)S_u = S_u$ for $u < t$. Finally, $\bar{Q}$ is equivalent to P because if A is a P-null set, then, A is a Q- and Q^*-null set, and therefore also a $\bar{Q}$-null set. Conversely, if A is a $\bar{Q}$-null set, then it must be both a Q- and a Q^*-null set. Hence, since Q and Q^* are equivalent to P, A must also be a P-null set. Note that this also proves that $\mathcal{M}^e(S)$ is a convex set, see Definition 2.1 (actually, $\mathcal{M}^e(S)$ being convex implies that $\mathcal{M}^e(S)$ is either the empty set, a one-element set or an infinite set).

In mathematical finance, one usually assumes that $\mathcal{M}^e(S) \neq \emptyset$, since this (from Theorem 4.2) guarantees that the market is arbitrage free. Actually, Theorem 4.2 can be proved to "almost" hold with if and only if. If the market satisfies a condition similar to the no arbitrage condition called "no free lunch with vanishing risk" (abbreviated NFLVR), then there exists an equivalent martingale measure (see Øksendal [27]). This is sometimes called *the fundamental theorem of asset pricing*.

Assume $\mathcal{M}^e(S) \neq \emptyset$, so the market has no arbitrage from Theorem 4.2. A *contingent claim* (or just *claim*) B is a financial contract where the seller promises to pay the buyer a random amount of money $B(\omega)$ at time T, depending on which state of the world, $\omega \in \Omega$ occurs. Formally, a claim is a lower bounded, $\mathcal{F}_T$-measurable random variable. If all contingent claims can be replicated by trade in the market, the market is called *complete*. More formally, this means that for any claim B there exists a self-financing, admissible trading strategy H such that $B(\omega) = H(0) \cdot S(0) + \int_0^T H(t,\omega)dS(t,\omega)$ for all $\omega \in \Omega$. This can be shown to be equivalent to $\mathcal{M}^e(S) = \{Q\}$ (this is due to Harrison and Pliska (1983), and Jacod (1979), see Øksendal [27]). If there exists contingent claims which cannot be replicated by trade, the market is called *incomplete*. In this case, $|\mathcal{M}^e(S)| = \infty$. In the real world, financial markets tend to be incomplete, because in order to have a complete market one must, among other things, have full information and no trading costs. This is not a realistic situation. However, incomplete markets are mathematically more difficult to handle than complete markets.

4.3 The discrete case

To simplify things, consider the case of finite scenario space Ω and discrete time. This is a useful simplification because:

- It may be sufficient in practical situations, since one often envisions only a few possible world scenarios, and has a finite amount of times where one wants to trade.
- It gives a good understanding of the methods used in mathematical finance.
- To apply mathematical financial theory, discretization is necessary in order to use computers.

Hence, let $\Omega = \{\omega_1, \omega_2, ..., \omega_M\}$, $t = 0, 1, ..., T$, and consider the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$. In this case, the σ-algebras in the filtration correspond to partitions of Ω, as mentioned previously. One can construct a scenario-tree illustrating the situation, with the tree branching according to the information partitions. Hence, there is one node (vertex) for each block (set) in the partition. Each $\omega \in \Omega$ represents a specific development in time, ending up in the particular world scenario at the final time T. Denote the nodes at time t by $\mathcal{N}_t$, and let the nodes themselves be indexed by $k = 1, ..., K$.

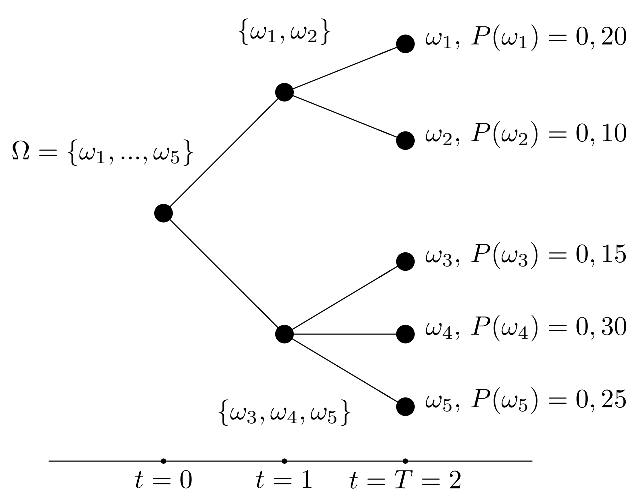

Figure 4.1: A scenario tree.

In the model illustrated in Figure 4.1, $K = 8$ and $M = 5$. The filtration $(\mathcal{F}_t)_{t=0,1,2}$ corresponds to the partitions $\mathcal{P}_1 = \Omega$, $\mathcal{P}_2 = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4, \omega_5\}\}$, $\mathcal{P}_2 = \{\{\omega_1\}, ..., \{\omega_5\}\}$.

A little more notation is useful: The *parent* $a(k)$ ($a(\cdot)$ for ancestor, $p(\cdot)$ is not practical since it often denotes probability measures) of a node k is the unique node $a(k)$ preceding k in the scenario tree. Every node apart from the first one has a parent. Each node k, except the terminal nodes $\mathcal{N}_T$, have *child nodes* $\mathcal{C}(k)$, that is a set of nodes immediately succeeding the node k in the scenario tree. For each non-terminal node k, the probability of ending up in node k is called p_k, and $p_k = \sum_{m \in \mathcal{C}(k)} p_m$. Hence, from the original probability measure P, which gives probabilities to each of the terminal nodes, one can work backwards, computing probabilities for all the nodes in the scenario tree.

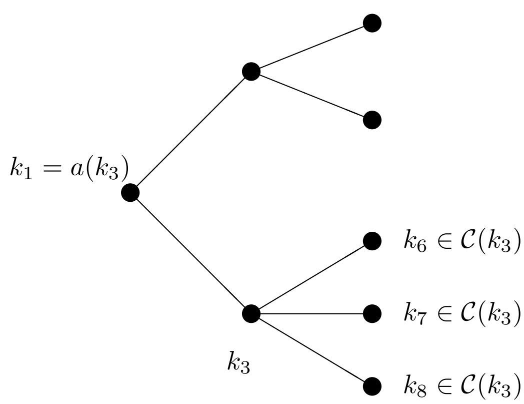

Figure 4.2: Illustration of parent and children nodes in a scenario tree.

As in the previous section, there are N risky assets and one bond with a composed price process $S = (S_0, S_1, ..., S_N)$. As S is adapted to the filtration $(\mathcal{F}_t)_t$, for each asset n, there is one value for the price S_n of asset n in each node k of the scenario tree. This value is denoted by S_k^n. H_k^n denotes the amount of asset n held by the investor at node k. Hence the value of the portfolio at node k is $X_k = S_k \cdot H_k$ (where $\cdot$ denotes Euclidean inner product).

In this finite Ω and discrete time case, one may assume that $P(\omega) > 0$ for all $\omega \in \Omega$. If not, the states with probability 0 can be removed without any consequences.

For a node $m \in \mathcal{C}(k)$, one can compute the conditional probability of m happening, given that k has happened. Let F_i be the subset of Ω where node i happens. Then: $P(F_m|F_k) = \frac{P(F_m \cap F_k)}{P(F_k)} = \frac{P(F_m)}{P(F_k)} = \frac{p_m}{p_k}$. Here, the equalities follow from the definition of conditional expectation and that $F_m \subseteq F_k$ (since $m \in \mathcal{C}(k)$, i.e. m can only happen if k has happened).

As mentioned, the portfolio value at time t is $X(t)$, and the value of this depends on which node $k \in \mathcal{N}(t)$ the world is in. Note that $E_P[X(t)] = \sum_{k \in \mathcal{N}(t)} p_k X_k$. The conditional expectation of the portfolio value $X(t+1)$ given $\mathcal{F}_t$, that is given $\mathcal{N}(t)$, is $E_P[X(t+1)|\mathcal{F}(t)] = E_P[X(t+1)|\mathcal{N}(t)] =$

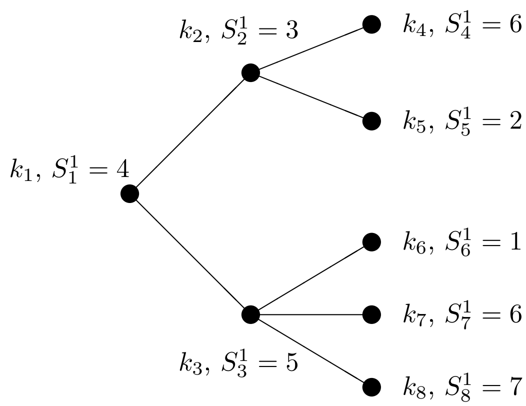

Figure 4.3: Discounted price process in a scenario tree: One risky asset.

$\sum_{m \in \mathcal{C}(k)} \frac{p_m}{p_k} X_m$, and this is a random variable which takes one value for each $k \in \mathcal{N}(t)$.

In this case, where Ω is finite and time is discrete, a stronger version of Theorem 4.2, namely Theorem 4.3, holds. Note that in Theorem 4.3, the condition of Q being equivalent to P, means that $Q > 0$, i.e. $Q(\omega) > 0$ for all $\omega \in \Omega$ (as $P(\omega) > 0$ for all $\omega \in \Omega$ by assumption). The proof of the following theorem is due to King [18] and uses a duality argument which shows how martingales are introduced via duality theory.

Theorem 4.3 *Let Ω be finite, and time discrete; $t = 0, 1, ..., T$. Then, the financial market has no arbitrage if and only if $\mathcal{M}^e(S) \neq \emptyset$.*

Proof: $\diamondsuit$

First, assume the market is normalized. It can be shown that if the normalized market is arbitrage free, so is the regular market (see Exercise 12.1 in Øksendal [27]).

In the finite Ω setting, an arbitrage is a portfolio H such that:

- $S_0 \cdot H_0 = 0$.
- $S_k \cdot H_k \geq 0$ for all $k \in \mathcal{N}_T$.
- There exists an $\bar{k} \in \mathcal{N}_T$ such that $S_{\bar{k}} \cdot H_{\bar{k}} > 0$.
- (Self-financing) $S_k \cdot H_k = S_k \cdot H_{a(k)}$ for all k.

where $S_k = (S_k^0, ..., S_k^N)$, $H_k = (H_k^0, ..., H_k^N)$, and $S_k \cdot H_k = \sum_{n=1}^N S_k^n H_k^n$.

That is, there exists a portfolio that costs nothing, has no risk of losing money, but a positive probability of earning money, and the portfolio is self-financing in the sense that all investments are financed by previous gains in the market, due to price changes.

One must solve the following optimization problem in order to determine whether there exists an arbitrage, and actually find one (if it exists).

$$\begin{array}{lrll}
\max_H & \sum_{k\in\mathcal{N}_T} p_k(S_k\cdot H_k) & & \\
\text{subject to} & & & \\
& S_0\cdot H_0 & = 0, & \\
& S_k\cdot(H_k - H_{a(k)}) & = 0 & \text{for } k\in\mathcal{N}_t, t\geq 1, \\
& S_k\cdot H_k & \geq 0 & \text{for all } k\in\mathcal{N}_T.
\end{array} \tag{4.7}$$

If problem (4.7) has a positive optimal value, there exists (from the definition) an arbitrage, and a maximizing portfolio H^* is an arbitrage.

Also, note that if problem (4.7) has a positive optimal value, then the problem is unbounded. Actually, if H generates a positive optimal value, then mH, for $m > 0$, also generates a positive optimal value, and letting $m \to \infty$, problem (4.7) is unbounded.

A closer look at problem (4.7) reveals that it is a linear optimization problem (LP-problem)! It is of course also a stochastic optimization problem, but all the randomness has been put directly into the problem by introducing the scenario tree.

Problem (4.7) will be studied using of linear programming duality. The first task is finding the dual problem. This is done by standard linear programming techniques. This is postponed to right after the proof (to avoid confusion).

We get a dual problem of the form:

$$\begin{array}{lrcll}
\min_{(x,y)} & 0 & & & \\
\text{subject to} & & & & \\
& x_k & \leq & 0 & \text{for } k\in\mathcal{N}_T, \\
& (p_k - y_k - x_k)S_k & = & 0 & \text{for } k\in\mathcal{N}_T, \\
& y_kS_k - \sum_{m\in\mathcal{C}(k)} y_mS_m & = & 0 & \text{for } k\in\mathcal{N}_t,\ t\leq T-1.
\end{array} \tag{4.8}$$

Note that the equation $y_kS_k = \sum_{m\in\mathcal{C}(k)} y_mS_m$ would be a martingale condition on the price process S if the y_k's were positive probabilities.

Now, for the actual proof. First, assume that there is no arbitrage. We want to show that then there must exist an equivalent martingale measure Q. Since there is no arbitrage, we know that problem (4.7) is bounded (it must have value 0). Also, (4.7) always has a feasible solution ($H = 0$), hence (4.7) has an optimal solution. Therefore, the LP duality theorem (see Vanderbei [42]) implies that the dual problem (4.8) also has an optimal solution (x, y).

From the dual problem (4.8), $(p_k - y_k - x_k)S_k = 0$ for all $k \in \mathcal{N}_T$. This is a vector equality, so it holds for every asset n, in particular for the bond $n = 0$. Hence, $(p_k - y_k - x_k)S_k^0 = 0$ for $k \in \mathcal{N}_T$. $S_k^0 \neq 0$ for all k, so $p_k - y_k - x_k = 0$

for all $k \in \mathcal{N}_T$. But then $p_k - y_k = x_k$ for all $k \in \mathcal{N}_T$. From (4.8), $x_k \leq 0$ for all k, hence

$$y_k \geq p_k \text{ for all } k \in \mathcal{N}_T.$$

Also, from (4.8)

$$y_k S_k^0 - \sum_{m \in \mathcal{C}(k)} y_m S_m^0 = 0$$

which implies, since $S_k^0 = 1$ for all $k = 1, ..., K$,

$$y_k = \sum_{m \in \mathcal{C}(k)} y_m.$$

By induction, this proves that $(y_k)_k$ is a strictly positive process such that $\sum_{k \in \mathcal{N}_T} y_k = y_0$. Hence, one can define $q_k = \frac{y_k}{y_0}$ for all $k \in \mathcal{N}_T$, so that $Q(\omega_k) := q_k$, for all $\omega_k \in \Omega$, is a probability measure.

Since $(y_k)_k$ satisfies the constraints of the dual problem (4.8), so will $(q_k)_k$, therefore

$$\sum_{m \in \mathcal{C}(k)} q_m S_m = q_k S_k \text{ for } k \in \mathcal{N}_t,\ t \leq T-1.$$

Hence, the price process S is a martingale with respect to Q and the filtration $(\mathcal{F}_t)_t$ (from the definition of conditional expectation), and Q is an equivalent martingale measure (since $Q > 0$).

Conversely, assume that there exists an equivalent martingale measure Q. Define $q_k := Q(\omega_k)$ for all $k \in \mathcal{N}_T$. Let $y_0 = \max\{\frac{p_k}{q_k} : k \in \mathcal{N}_T\}$, $y_k = q_k y_0$ for all $k \in \Omega$ and $x_k = p_k - y_k$ for all $k \in \mathcal{N}_T$.

Then, $x_k = p_k - y_k = p_k - q_k y_0 \leq p_k - q_k \frac{p_k}{q_k} = 0$. So (x, y) will be feasible for the dual problem (4.8) (from the definitions). Hence, by the weak duality theorem of LP (see Vanderbei [42]), problem (4.7) must be bounded, and hence there exists no arbitrage.

□

It remains to show how to find the dual of the LP-problem (4.7) in the proof of Theorem 4.3. The primal problem (4.7) is as follows

$$\begin{array}{lrll}
\max_H & \sum_{k \in \mathcal{N}_T} p_k S_k \cdot H_n & & \\
\text{subject to} & & & \\
 & S_0 \cdot H_0 & = 0 & \\
 & S_k \cdot (H_k - H_{a(k)}) & = 0 & \text{for } k \in \mathcal{N}_t, t \geq 1 \\
 & S_k \cdot H_k & \geq 0 & \text{for all } k \in \mathcal{N}_T
\end{array}$$

This problem is of the form $\max_H\{e^T H : EH = b, KH \geq c\}$, where e, b, c are vectors and E, K are matrices of suitable dimensions, where $\geq$ is meant componentwise (and H is a vector in $\mathbb{R}^{|\mathcal{N}|\times N+1}$ representing the trading strategy, where $|\mathcal{N}|$ denotes the number of nodes in the scenario tree). Therefore, we begin by proving the following lemma:

Lemma 4.4 $\diamondsuit$ *The dual problem of the linear programming problem*

$$\max_H \{e^T H : EH = b, KH \geq c\}$$

is

$$\min_{(v,w)} \{b^T v + c^T w : E^T v + H^T w = e, w \leq 0\}.$$

Proof: $\diamondsuit$

The idea of this proof is to reduce the primal problem to the original primal LP-problem $\max\{c^T x : Ax \leq b, \ x \geq 0\}$, which is known to have the dual $\min\{b^T y : A^T y \geq c, \ y \geq 0\}$ (see Vanderbei [42]).

The following maximization problems, denoted (i)-(iv), are equivalent in the sense of one having a solution if and only if the other has a solution, and having the same optimal solution

$$\text{(i)} \quad \max_H \quad e^T H \quad \text{subject to} \quad \begin{aligned} EH &= b, \\ KH &\geq c. \end{aligned}$$

$$\text{(ii)} \quad \max_H \quad e^T H \quad \text{subject to} \quad \begin{aligned} EH &\leq b, \\ EH &\geq b, \\ KH &\geq c. \end{aligned}$$

$$\text{(iii)} \quad \max_{H^+,H^-} \quad e^T(H^+ - H^-) \quad \text{subject to} \quad \begin{aligned} E(H^+ - H^-) &\leq b, \\ -E(H^+ - H^-) &\leq -b, \\ -K(H^+ - H^-) &\leq -c. \end{aligned}$$

$$\text{(iv)} \quad \max_{H^+,H^-} \quad \begin{bmatrix} e & -e^T \end{bmatrix} \begin{bmatrix} H^+ \\ H^- \end{bmatrix} \quad \text{subject to} \quad \begin{aligned} \begin{bmatrix} E & -E \\ -E & E \\ -K & K \end{bmatrix} \begin{bmatrix} H^+ \\ H^- \end{bmatrix} &\leq \begin{bmatrix} b \\ -b \\ -c \end{bmatrix}, \\ \begin{bmatrix} H^+ \\ H^- \end{bmatrix} &\geq 0. \end{aligned}$$

This is a standard form of the primal LP-problem, see [42]. Hence, duality theory gives that the dual has the form:

$$
\begin{array}{ll}
\min_{v_1,v_2,\bar{w}} & \begin{bmatrix} b^T & -b^T & -c^T \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \bar{w} \end{bmatrix} \\
\text{subject to} & \\
& \begin{bmatrix} E^T & -E^T & -H^T \\ -E^T & E^T & H^T \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \bar{w} \end{bmatrix} \geq \begin{bmatrix} e \\ -e \end{bmatrix}, \\
& v_1, v_2, \bar{w} \geq 0,
\end{array}
$$

i.e.

$$
\begin{array}{lrcl}
\min_{v_1,v_2,\bar{w}} & b^T v_1 - b^T v_2 - c^T \bar{w} & & \\
\text{subject to} & E^T v_1 - E^T v_2 - H^T \bar{w} & \geq & e, \\
& -E^T v_1 + E^T v_2 + H^T \bar{w} & \geq & -e, \\
& v_1, v_2, \bar{w} & \geq & 0.
\end{array}
$$

Note that the two first inequality constraints in the problem above are equivalent with $E^T(v_1 - v_2) - H^T \bar{w} = e$, and insert this into the problem.

$$
\begin{array}{lrcl}
\min_{v_1,v_2,\bar{w}} & b^T(v_1 - v_2) - c^T \bar{w} & & \\
\text{subject to} & E^T(v_1 - v_2) - H^T \bar{w} & = & e, \\
& v_1, v_2, \bar{w} & \geq & 0.
\end{array}
$$

Defining $v := v_1 - v_2$, v is a free variable, since $v_1, v_2 \geq 0$

$$
\begin{array}{lrcl}
\min_{v,\bar{w}} & b^T v - c^T \bar{w} & & \\
\text{subject to} & E^T v - H^T \bar{w} & = & e, \\
& \bar{w} & \geq & 0.
\end{array}
$$

Finally, let $w := -\bar{w}$, so

$$
\begin{array}{lrcl}
\min_{v,w} & b^T v + c^T w & & \\
\text{subject to} & E^T v + H^T w & = & e, \\
& w & \leq & 0.
\end{array}
$$

□

By noting that the primal problem (4.7) is exactly of the form of the primal problem in Lemma 4.4 and writing problem (4.7) in component form, one sees that the dual of (4.7) is precisely problem (4.8) (writing this out is quite technical and notation-packed, so it is omitted).

This chapter has introduced the financial market model, and the basic terms of mathematical finance, which will be used in the remaining part of the thesis.

In addition, the proof of Theorem 4.3 which uses linear programming duality, is an example of an application of duality theory to mathematical finance. The next chapter will consist of more examples of duality methods in mathematical finance, but in relation to utility maximization problems.

Chapter 5

Utility maximization

This chapter considers an investor in the financial market. The investor is assumed to have certain preferences, represented by a utility function U. This utility function is assumed to have some economically reasonable properties, as will be discussed in Section 5.1.

Throughout this chapter it is assumed that the investor wants to maximize her expected utility of terminal wealth. This problem is solved in Section 5.3 using the text book technique of Pliska [32] for a simple one step model and a complete market.

The direct method of Section 5.3 cannot be applied to markets where Ω is arbitrary and time is continuous. Section 5.5 introduces a more general method (in the finite Ω, complete market setting), a *duality method*, which is based on convexity arguments and Lagrange duality. Lagrange duality is the topic of Section 5.4. In Sections 5.6 and 5.7 this duality method is generalized to incomplete, but finite Ω, markets and finally to completely general markets.

Section 5.9 is devoted to a twist on the utility maximization problem, where a condition on how much risk the investor is willing to take is added to the problem. This connects the utility maximization problem to the convex risk measures of Chapter 3.

Section 5.2 shows how the previous utility maximization problem is equivalent to another utility maximization problem where the trading strategy involved is essential.

This chapter is based on Kramkov and Schachermayer [22], [21], Pham [31], Pliska [32] and Schachermayer [40].

5.1 Utility functions and utility maximization

It is natural to consider the preferences of an agent in the financial market. This is done by the introduction of utility functions.

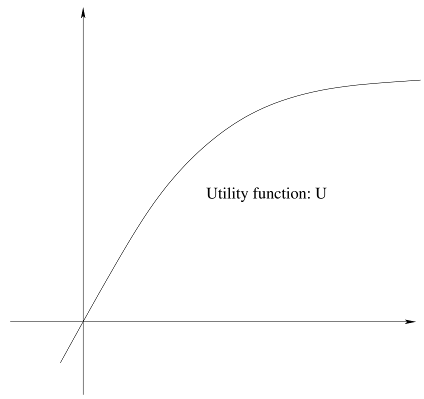

Figure 5.1: A utility function: U

Definition 5.1 *(Utility function, $U(\cdot)$)*
A utility function is a function $U : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ which is increasing (on $\{U > -\infty\}$), continuous, differentiable and strictly concave (on the interior of $\{U > -\infty\}$) and which satisfies:

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

The economic interpretation of the final condition is that the marginal utility converges towards 0 as the wealth goes to infinity. Hence, when the agent gets a lot of money, a little more means almost nothing. The other conditions on the utility function also have a natural economic explanation: U should be increasing, because more money is a good thing. Differentiability and continuity are more technical properties, but one may interpret these as well. Continuity, for instance, is natural because it does not make sense to be a lot more happy about having $10 + 0,00001$ NOK, than about having 10 NOK. The utility function U being concave reflects that the more money the agent gets, the less a little extra means: For a student, who makes 50 000 NOK a year, an extra 5000 NOK means a lot, but Bill Gates, who makes 1 billion NOK a year, may not even notice those extra 5000 NOK.

When it comes to utility functions, there are two cases to consider:

1. *Case 1: Not allowing negative wealth*: Assume that $U(x) = -\infty$ for $x < 0$ and that $U(x) > -\infty$ for $x \geq 0$. Also, assume the *Inada condition*:

$$U'(0) = \lim_{x \to 0^+} U'(x) = +\infty.$$

2. *Case 2: Allowing negative wealth:* In this case, assume that $U(x) > -\infty$ for all $x \in \mathbb{R}$ and that $U'(-\infty) = \lim_{x \to -\infty^+} U'(x) = +\infty$.

Often, there is no reason to distinguish between the two cases. However, in Section 5.7 the arguments depend on which case one is considering.

An example of a utility function of the first type is:

$$U(x) = \begin{cases} \ln(x) & \text{for } x > 0 \\ -\infty & \text{otherwise} \end{cases}$$

An example of the second type is $U(x) = -e^{-\gamma x}$ for all $x \in \mathbb{R}$, where $\gamma \geq 0$.

So, consider an investor in the market S (see Chapter 4) with preferences given by the utility function U. How does the investor decide which assets to buy? One possibility is that she wants to maximize her expected utility of terminal wealth. One must consider, for instance, the expectation since terminal wealth is a random variable, and it is not possible to maximize a random variable directly. Hence, the investor wants to solve the following problem

$$\max E[U(X^H(T))] \tag{5.1}$$

where the maximization is done over the set of self-financing, admissible portfolios H that are affordable to the investor.

How do we solve a problem of this type? Since U is concave, expectation is linear and a linear function composed with a concave function gives another concave function, $E_P[U(\cdot)]$ is a concave function. Thus problem (5.1) is a constrained concave maximization problem. One possible strategy for attacking such a problem is the following:

- Try to solve problem (5.1) directly.
- If this fails, find a "dual" to the problem.
- Try to solve this dual problem, and prove that the dual problem and problem (5.1) have the same value.

How to find this dual problem will be discussed later.

Note that this is the same as a basic idea of linear programming (LP). The original (primal) problem may be difficult to solve, but the linear programming duality theorem says that solving the dual problem is just as good. Duality is a "second chance" at solving the problem.

Note that problem (5.1) is also an optimal control problem. This is a type of problem which has been studied extensively, and this has resulted in Hamilton-Jacobi-Bellman equations. For more on this, see Øksendal [27].

5.2 Transformation of the utility maximization problem ◇

Assume $\mathcal{M}^e(S) \neq \emptyset$, so (from Theorem 4.2) the market has no arbitrage.

In the next sections the following problem will be considered

$$\begin{array}{ll} \max_{Y \in \mathcal{L}^0(\mathcal{F}_T)} & E[U(Y)] \\ \text{subject to} & \\ & E_Q[Y] \leq x \text{ for all } Q \in \mathcal{M}^e(S). \end{array} \tag{5.2}$$

Problem (5.2) is of interest because it turns out to be equivalent to problem (5.1) for an investor with initial capital x, that is

$$\max_H E[U(X_T^H)] \tag{5.3}$$

where the maximization is done over the set $\mathcal{H} := \{H : H$ is a self-financing, admissible, predictable portfolio process such that $H(1) \cdot S(0) \leq x\}$, where $H(t)$ denotes the composed amounts of assets held from time $t-1$ to time t, $S(t)$ is the composed price process at time t, and X_T^H denotes the terminal value of the portfolio corresponding to the portfolio process H. Since $H \in \mathcal{H}$ is self-financing, $X_T^H = H(1) \cdot S(0) + \int_0^T H(u) dS(u)$. Problem (5.3) is a natural utility maximization problem for an investor in the market. Note that if $X_T^{H^*}$ is an optimal solution of problem (5.3), then $H^*(1) \cdot S(0) = x$, since H^* is self-financing and $E[U(\cdot)]$ is an increasing function.

This section proves that in some cases, one may solve problem (5.2) instead of problem (5.3). That is, problems (5.2) and (5.3) have the same optimal value, and if one has found the optimal solution of problem (5.2), then one can also find the optimal solution of problem (5.3). It turns out that this holds for a market where Ω is finite (complete or incomplete), as well as for a complete market where Ω is arbitrary. In the remaining situation of an incomplete market where Ω is arbitrary, the original utility maximization problem (5.3) will be tackled directly.

Now, begin by considering the situation where the scenario space Ω is finite, but the market may be complete or incomplete. In this setting, Proposition 2.10 in Schachermayer [40] states:

An $\mathcal{F}_T$-measurable random variable $Y(\omega)$ can be dominated by a random variable of the form $x + \int_0^T H_s dS_s$, where $H \in \mathcal{H}$, iff. $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^e(S)$.

This proposition is proved via convexity arguments in $\mathbb{R}^M$ (where $|\Omega| = M$), the bipolar theorem (which is a finite-dimensional version of the biconjugate theorem, Theorem 2.39) and the fundamental theorem of asset pricing (which can also be proved from convexity arguments in $\mathbb{R}^M$).

Theorem 5.2 $\diamondsuit$

(Transformation, finite Ω) In the setting above, the following holds:

(i) The optimal values of problems (5.2) and (5.3) coincide.

(ii) If one has found the optimal solution Y^* of problem (5.2), one can find the optimal solution H^* of problem (5.3).

Proof: $\diamondsuit$

(i) Take a feasible solution of problem (5.2), i.e. a $Y \in \mathcal{L}_0(\mathcal{F}_T)$ such that $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^e(S)$. Then there exists, from Proposition 2.10 in Schachermayer, an $H \in \mathcal{H}$ such that $Y \leq x + \int_0^T H_s dS_s$. Hence, since $E[U(\cdot)]$ is increasing, $E[U(Y)] \leq E[U(x + \int_0^T H_s dS_s)]$. Therefore, because H is self-financing, the optimal value of problem (5.2) is less than or equal the optimal value of problem (5.3).

Conversely, let $H \in \mathcal{H}$ be a feasible solution of problem (5.3). Then, $X_T^H \leq x + \int_0^T H_s dS_s \in \mathcal{L}^0(\mathcal{F}_T)$ (the terminal value of the value process corresponding to H), hence $E_Q[U(X_T)] \leq E_Q[x + \int_0^T H_s dS_s] = x \leq x$ for all $Q \in \mathcal{M}^e(S)$, since S is a martingale w.r.t. Q for all $Q \in \mathcal{M}^e(S)$. Hence, X_T^H is feasible for problem (5.2), and therefore the optimal value of problem (5.3) is less than or equal the optimal value of problem (5.2).

Hence, the two optimal values must coincide.

(ii) Assume Y^* is the optimal solution of problem (5.2). From Schachermayer's proposition there exists an $H^* \in \mathcal{H}$ such that $Y^* \leq x + \int_0^T H_s^* dS_s$. But then

$$\begin{aligned}
\sup_{\{Y: E_Q[Y] \leq x \ \forall \ Q \in \mathcal{M}^e(S)\}} E[U(Y)] &= E[U(Y^*)] \\
&\leq E[U(x + \textstyle\int_0^T H_s^* dS_s)] \\
&\leq \sup_{H \in \mathcal{H}} E[U(X_T^H)] \\
&= \sup_{\{Y: E_Q[Y] \leq x \ \forall \ Q \in \mathcal{M}^e(S)\}} E[U(Y)].
\end{aligned}$$

Hence, all the inequalities hold with equality, so H^* must give the optimal solution of problem (5.3).

□

Therefore, solving problem (5.2) is useful in the case of finite Ω.

Now, consider an arbitrary scenario space Ω and a complete market with the Brownian motion driven market model of Section 4.2. Then, there is only one equivalent martingale measure (from the comments after Theorem 4.2). From Theorem 1.6.6 in Karatzas and Shreve [17] the market model must satisfy $N = D$ (where N is the number of risky assets and D is the dimension of the Brownian motion) and the volatility matrix σ must be invertible (P-a.s.). Hence, *the Girsanov theorem* holds for the Brownian motion driven market model.

Theorem 5.3 *(Girsanov theorem)*

Let $Y(t) \in \mathbb{R}^n$ be an Itô-process (see Øksendal [27] for a definition of Itô-process) on a filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{0 \leq t \leq T})$ of the form

$$dY(t) = \beta(t, \omega)dt + \theta(t, \omega)dB(t) \ , \ 0 \leq t \leq T$$

where $B(t) \in \mathbb{R}^m$ is a Brownian motion, $\beta(t,\omega) \in \mathbb{R}^n$ and $\theta(t,\omega) \in \mathbb{R}^{n\times m}$. Suppose there exist generalized Itô-integrable (see [27]) processes $u(t,\omega) \in \mathbb{R}^m$ and $\alpha(t,\omega) \in \mathbb{R}^n$ such that

$$\theta(t,\omega)u(t,\omega) = \beta(t,\omega) - \alpha(t,\omega).$$

Define

$$M_t = e^{-\int_0^t u(s,\omega)dB_s - \frac{1}{2}\int_0^t u^2(s,\omega)ds} \quad \text{for} \quad 0 \leq t \leq T$$

and

$$dQ(\omega) = M_T(\omega)dP(\omega).$$

Assume that M_t is a martingale (this holds, for instance, under a boundedness condition called the Novikov condition; see [27]). Then Q is a probability measure which is equivalent to P (since $M_T(\cdot) > 0$ per definition). Also, the process defined by

$$\bar{B}(t) = \int_0^t u(s,\omega)ds + B(t) \text{ for } 0 \leq t \leq T$$

is a Brownian motion with respect to Q. Finally

$$dY(t) = \alpha(t,\omega)dt + \theta(t,\omega)d\bar{B}(t).$$

See [27] for more on the Girsanov theorem. Since the Girsanov theorem holds, and the Girsanov measure Q is an equivalent martingale measure (from the theorem), the single equivalent martingale measure in the market must be the Girsanov measure. Karatzas and Shreve [17] prove a replication result for this kind of market (the theorem is adapted to the present notation and setting):

Theorem 5.4 *(Replication result) Consider the setting above. Let the initial endowment $x > 0$ be given, and let Y be an $\mathcal{F}_T$-measurable random variable such that $E_Q[Y] = x$. Then, there exists a portfolio process $H \in \mathcal{H}$ such that $x + \int_0^T H_s dS_s = Y$ (P-a.s.).*

In Karatzas and Shreve [17], this theorem is proved in a more general setting using a "budget constraint" (derived from the definition of admissible portfolios, using that a lower bounded local martingale is a supermartingale) and a more general version of the martingale representation theorem (see Øksendal [27]) called the local martingale representation theorem (to get the existence of a replicating portfolio H).

Now, one can prove the following theorem:

Theorem 5.5 $\diamondsuit$

(Transformation, arbitrary Ω) Consider the setting above, then:

(i) Problems (5.2) and (5.3) have the same optimal value.

(*ii*) *If one has found the optimal solution, Y^*, of problem (5.2), then there exists an optimal solution H^* of problem (5.3) which replicates Y^*.*

Proof: $\diamondsuit$

(*i*) Note that $\sup_{\{Y:E_Q[Y]\leq x\}} E[U(Y)] = \sup_{\{Y:E_Q[Y]=x\}} E[U(Y)]$ since $E[U(\cdot)]$ is increasing. Choose $Y \in \mathcal{L}^0(\mathcal{F}_T)$ such that $E_Q[Y] = x$. Then Y is feasible for problem (5.2). The replication result, Theorem 5.4, implies that there exists an $H \in \mathcal{H}$ that replicates Y, hence

$$\begin{aligned} \sup_{\{Y:E_Q[Y]\leq x\}} E[U(Y)] &= \sup_{\{Y:E_Q[Y]=x\}} E[U(Y)] \\ &\leq \sup_{\{H\in\mathcal{H}\}} E[U(x + \int_0^T H_s dS_s)] \\ &= \sup_{\{H\in\mathcal{H}\}} E[U(X_T^H)]. \end{aligned}$$

So the optimal value of problem (5.2) is less than or equal the optimal value of problem (5.3).

Conversely, take $H \in \mathcal{H}$ (i.e. a feasible solution of problem (5.3)), and consider $X_T^H \leq x + \int_0^T H_s dS_s \in \mathcal{L}^0(\mathcal{F}_T)$ (where X_T^H denotes the terminal value of the portfolio H). Then, $E_Q[X_T^H] \leq E_Q[x + \int_0^T H_s dS_s] = x$, since Q is an equivalent martingale measure. Hence, X_T^H is feasible for problem (5.2). Therefore, by the same type of calculations as above, the optimal value of problem (5.3) is less than or equal the optimal value of problem (5.2).

Hence, the optimal values coincide.

(*ii*) Finally, assume that one has found an optimal solution Y^* of problem (5.2). Then, $E_Q[Y^*] = x$ (if not, Y^* could not be optimal since $E[U(\cdot)]$ is increasing), and from the replication result (Theorem 5.4) there exists an $H^* \in \mathcal{H}$ that replicates Y^*. Since the two optimal values coincide (from (*i*)), this H^* must be optimal for problem (5.3).

$\square$

Hence, one may solve problem (5.2) instead of problem (5.3) in the arbitrary Ω, complete market situation as well.

The problem with all of the arguments above is that there is no direct formula for finding the replicating portfolio H when one has derived the optimal terminal value Y^*. The core theorems of the replication results; the bipolar theorem and the local martingale representation theorem, merely give the *existence* of a replicating portfolio. Hence, even after solving problem (5.2), there is a substantial task in actually *finding* this portfolio. However, this problem will not be considered in this thesis.

5.3 A direct approach to utility maximization

The utility maximization problem (5.1) will now be solved for an example market using a direct method. This method is introduced in Pliska's book [32], and is a very intuitive way to solve the problem. The market considered is very simple, and this is why the direct approach works. In more complicated market models, such as arbitrary scenario space models, one needs a more advanced solution strategy. This will be considered in the next sections, and is where the equivalence of problems (5.2) and (5.3) is useful.

Example 5.6 $\diamondsuit$

Consider the following complete market in a one step model. Let $\Omega = \{\omega_1, \omega_2\}$*,* $t = 0, 1$*. The market is assumed to be normalized, so the price process of the bond satisfies* $S_0(0) = 1$*,* $S_0(1, \omega_1) = S_0(1, \omega_2) = 1$*. The market only has one risky asset with price process* $S_1(0) = 4$*,* $S_1(1, \omega_1) = 7$*,* $S_1(1, \omega_2) = 2$*. The probability measure* P *is such that* $P(\omega_1) = \frac{1}{3}$ *and* $P(\omega_2) = \frac{2}{3}$*.*

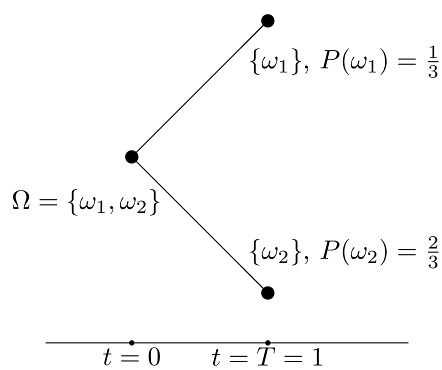

Figure 5.2: Scenario tree illustration of the model.

The set of equations

$$
\begin{aligned}
E_Q[\Delta S_1] &= 0 \\
Q(\omega_1) + Q(\omega_2) &= 1
\end{aligned}
$$

only has one solution, $(Q(\omega_1), Q(\omega_2)) = (\frac{2}{5}, \frac{3}{5})$*, which satisfies* $Q(\omega_1), Q(\omega_2) > 0$*, and hence the probability measure* Q *is an equivalent martingale measure. Therefore, the market is arbitrage free (from Theorem 4.3) and complete (from the comments after Theorem 4.3).*

Consider an investor in this market with utility function $U(y) = \ln(y)$*. The initial endowment of the investor is* $x > 0$*. Now, consider the utility maximization problem of the investor, as discussed previously*

$$
\max_{H \in \mathcal{H}} E_P[U(X^H(1))] = \max_{H \in \mathcal{H}} \{\frac{1}{3} \ln(X^H(1, \omega_1)) + \frac{2}{3} \ln(X^H(1, \omega_2))\}
$$

where $\mathcal{H}$ *is the set of all self-financing, predictable portfolio processes* H *such that* $X^H(0) = x$*.*

Note that since the time is discrete, stochastic integrals are reduced to sums, so

$$\begin{aligned}\int_0^t H(u)dS(u) &= \sum_{u=1}^{t} H(u)\cdot\Delta S(u)\\ &= \sum_{u=1}^{t} H(u)\cdot(S(u)-S(u-1))\end{aligned}$$

where $\cdot$ *denotes the Euclidean inner product, and* $H=(H_0,H_1,\ldots,H_N)$ *and* $S=(S_0,S_1,\ldots,S_N)$ *are viewed as vectors in* $\mathbb{R}^{N+1}$.

Hence

$$\begin{aligned}\textstyle\int_0^1 H_u dS_u &= H_1(1)(S_1(1)-4)\\ &= \begin{cases} 3H_1(1) & \text{for } \omega=\omega_1\\ -2H_1(1) & \text{for } \omega=\omega_2\end{cases}\end{aligned} \tag{5.4}$$

Here, $H_n(1)$, *for* $n=0,1$, *denotes the number of shares in asset* n *the investor decides to hold between time* 0 *and time* 1. *From equation (5.4) the self-financing condition becomes* $X^H(1,\omega_1)=x+3H_1(1)$ *and* $X^H(1,\omega_2)=x-2H_1(1)$.

Define the value function

$$u(x) := \max\nolimits_{H\in\mathcal{H}} E_P[U(X^H(1))].$$

From equation (5.4) and the self-financing condition

$$u(x) = \max\nolimits_H\{\tfrac{1}{3}\ln(x+3H_1(1)) + \tfrac{2}{3}\ln(x-2H_1(1))\}.$$

This is a calculus maximization problem, which can be solved by differentiating with respect to $H_1(1)$, *and finding* $H_0(1)$ *from the condition that* $X^H(0)=x$. *Differentiating and setting equal to zero gives* $\frac{1}{x+3H_1(1)} - \frac{4}{3(x-2H_1(1))} = 0$. *Therefore* $H_1(1) = -\frac{x}{18}$. *This is the optimal amount of asset* 1 *for the investor to buy. Determine* $H_0(1)$, *the amount of money the investor should keep in the bank, from the initial endowment* $X^H(0)=H_0(1)+4H_1(1)=x$. *Hence* $H_0(1)=\frac{11x}{9}$. *Therefore, the optimal portfolio of the investor is* $H_1=(H_0(1),H_1(1))=(\frac{11x}{9},-\frac{x}{18})$.

The following is a summary of the direct method for a complete model where $|\Omega|=M$ and there are T time steps:

- Consider the function $\sum_{m=1}^{M} p_m U(x+\sum_{t=1}^{T} H(t,\omega_m)(S(t,\omega_m)-S(t-1,\omega_m)))$.

- For each $m = 1, ..., M$, $n = 1, ..., N$ and each $t \in \{1, ..., T\}$, differentiate this function with respect to $H_n(t, \omega_m)$, and set the result equal to zero. Here, one must do MTN differentiations (this is OK since U is differentiable). This gives a maximum since $E[U(\cdot)]$ is concave (because U is concave and $E[\cdot]$ is linear). Also, since the maximization is over all of $\mathbb{R}$, there is no boundary to check (see Section 2.2).

- Solve this system of MTN equations for $H_n(t, \omega_m)$ for each n, t, m.

- Compute $H_0(1)$ from that $X^H(0) = x$ and then compute $H_0(t, \omega)$ using the self-financing condition.

What are the advantages of this direct method? First of all, it is very intuitive and uses only calculus. In addition, constraints on the trading strategies are easy to model. However, when there are constraints, the utility maximization problem must be solved using other techniques than the one illustrated above. For instance, the Lagrange multiplier method can be applied to tackle the constraints. The disadvantages of the method are that as the number of time-steps and the number of $\omega \in \Omega$ increase, many computations become necessary. Also, the system of equations in the method above can be very difficult to solve because it is not necessarily a linear system of equations. Another weakness of the direct method is that it cannot be generalized to a model where Ω is arbitrary and time is continuous, which makes it less useful in theoretical mathematical finance. Because of these disadvantages, it is interesting to consider an alternative method for solving the utility maximization problem. An example of such a method is the duality method. Before introducing this method, it is useful to consider some background theory, called Lagrange duality.

5.4 Lagrange duality

As mentioned, the method of Section 5.3 has some negative sides. Therefore, we need an alternative method for solving the utility maximization problem, and duality methods are such an alternative. Before considering the duality method presented in Schachermayer [40], some background theory, which is used to derive the method, will be covered. This background theory is called *Lagrange duality*. When reading the note Schachermayer [40], the duality method can be difficult to understand and various clever functions seem to appear out of thin air. However, the method becomes clearer by bringing in Lagrange duality theory.

The method of Lagrange duality can be described as follows: Let X be a general inner product space with inner product $\langle \cdot, \cdot \rangle$. Assume there is a function $f : X \to \mathbb{R}$ to be maximized under certain constraints.

Consider a problem of the following, very general, form

$$\text{maximize } f(x) \text{ subject to } g(x) \leq 0, \; x \in S \tag{5.5}$$

where g is a function such that $g : X \to \mathbb{R}^N$ and $S \neq \emptyset$ (to exclude a trivial case). This will be called *the primal problem*. Note that if one has a problem with equality constraints, one can rewrite this in the form of problem (5.5) by writing each equality as two inequalities. Also, $\geq$ can be turned into $\leq$ by multiplying with -1, and by basic algebra, one can always make sure there is 0 on one side of the inequality. Note that there are no constraints on f or S and only one (weak) constraint on g. Hence, many problems can be written in the form (5.5).

Let $\lambda \in \mathbb{R}^N$ be such that $\lambda \geq 0$ (componentwise), and assume that $g(x) \leq 0$ (componentwise) for all $x \in S$. Then:

$$f(x) \leq f(x) - \lambda \cdot g(x) \tag{5.6}$$

because $\lambda \cdot g(x) \leq 0$ (where $\cdot$ denotes the Euclidean inner product). This motivates the definition of the *Lagrange function*, $L(x, \lambda)$

$$L(x, \lambda) = f(x) - \lambda \cdot g(x).$$

Hence, $L(x, \lambda)$ is an upper bound on the objective function for each $\lambda \in \mathbb{R}^N$, $\lambda \geq 0$ and $x \in X$ such that $g(x) \leq 0$. By taking supremum on each side of the inequality in (5.6), for each $\lambda \geq 0$,

$$\begin{array}{rcl} \sup\{f(x) : g(x) \leq 0,\ x \in S\} & \leq & \sup\{f(x) - \lambda \cdot g(x) : g(x) \leq 0, x \in S\} \\ & = & \sup\{L(x, \lambda) : x \in S, g(x) \leq 0\} \\ & \leq & \sup_{x \in S} L(x, \lambda) \\ & := & L(\lambda) \end{array} \tag{5.7}$$

where the second inequality follows because we are maximizing over a larger set, hence the optimal value cannot decrease.

This implies that for all $\lambda \geq 0$, $L(\lambda)$ is an upper bound for the optimal value function. We want to find the smallest upper bound. This motivates the definition of the *Lagrangian dual problem*

$$\inf_{\lambda \geq 0} L(\lambda). \tag{5.8}$$

Therefore, the following theorem is proven (by taking the infimum on the right hand side of equation (5.7)).

Theorem 5.7 *(Weak Lagrange duality)*
In the setting above, the following inequality holds

$$\sup\{f(x) : g(x) \leq 0,\ x \in S\} \leq \inf\{L(\lambda) : \lambda \geq 0\}.$$

This theorem shows that the Lagrangian dual problem gives the smallest upper bound on the optimal value of problem (5.5) generated by the Lagrange

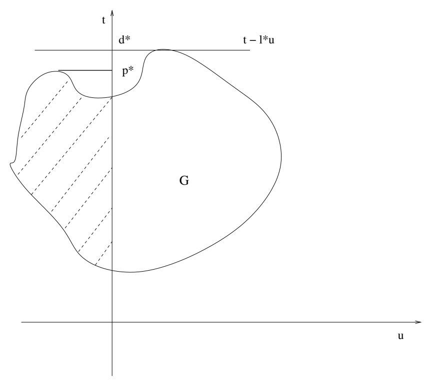

Figure 5.3: Illustration of Lagrange duality with duality gap.

function. The Lagrangian dual problem has only one, quite simple, constraint, namely $\lambda \geq 0$, and this may mean that the dual problem is easier to solve than the original problem.

In some special cases, one can proceed to show duality theorems, proving that $\sup\{f(x) : g(x) \leq 0,\ x \in S\} = \inf_{\lambda \geq 0} L(\lambda)$. If this is the case, one says that there is no *duality gap*. This typically happens in convex optimization problems under certain assumptions. However, often there actually is a duality gap, but the Lagrangian dual problem still gives us an upper bound, and hence some idea of the optimal value of our problem.

An example where Lagrangian duality is applied, and a duality theorem is derived, is linear programming (LP) duality. The calculation of the LP dual from the primal is omitted here, but it is fairly straight-forward.

The definition of conjugate functions (see Section 2.5) also has its roots in Lagrangian duality. The conjugate function shows up naturally when finding the Lagrangean dual of a minimization problem with linear inequalities as constrains. Section 5.5 will illustrate a version of this.

One can illustrate Lagrange duality such that it is simple to see graphically whether there is a duality gap. Consider problem (5.5) where $S = X$, and define the set $\mathcal{G} = \{(g(x), f(x)) \in \mathbb{R}^{N+1} : x \in X\}$. The optimal value of problem (5.5), denoted p^*, can then be written as $p^* = \sup\{t : (u,t) \in \mathcal{G}, u \leq 0\}$ (from the definitions). This can be illustrated for $g : X \to \mathbb{R}$ (i.e. for only one inequality) as in Figures 5.3 and 5.4.

Figure 5.4 shows the set $\mathcal{G}$, the optimal primal value p^* and the Lagrange-function for two different Lagrange multipliers. The value of the function $L(l) = \sup_{x \in X}\{f(x) - lg(x)\}$ is given by the intersection of the line $t - lu$ and the t-axis. Note that the shaded part of $\mathcal{G}$ corresponds to the feasible solutions of problem (5.5). Hence, to find the optimal primal solution p^* in the figure, find the point (u^*, t^*) in the shaded area of $\mathcal{G}$ such that t^* is as large as possible.

How can one find the optimal dual solution in Figure 5.4? Fix an $l \geq 0$,

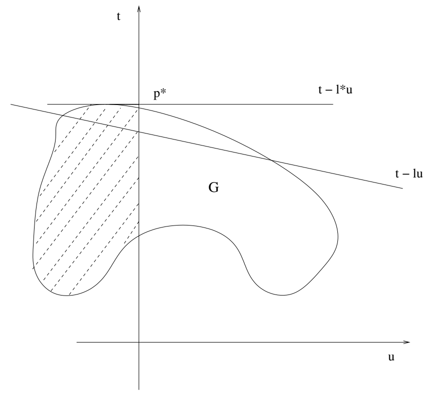

Figure 5.4: Illustration of Lagrange duality with no duality gap.

and draw the line $t - lu$ into the figure. Now, find the function $L(l)$ by parallel-adjusting the line so that the intersection of $t - lu$ is as large as possible, while making sure that the line still intersects $\mathcal{G}$. Having done this, tilt the line such that l is still greater than or equal 0, but such that the intersection of the line and the t-axis becomes as big as possible. The final intersection is the optimal dual solution.

Actually, there is no duality gap in the problem of Figure 5.4, since the optimal primal value corresponds to the optimal dual value, given by the intersection of the line $t - l^*u$ and the t-axis.

In Figure 5.3 there is a duality gap, since the optimal dual value, denoted d^* is greater than the optimal primal value, denoted p^*. What goes wrong? By examining the two figures above, one sees that the absence of a duality gap has something to do with the set $\mathcal{G}$ being "locally convex" near the t-axis. Bertsekas [2] formalizes this idea, and shows a condition for the absence of a duality gap (in the Lagrange duality case), called the Slater condition.

The Slater condition, in the case where $X = \mathbb{R}^n$ (see Boyd and Vandenberghe [3]), states the following: Assume there is a problem of the form (5.5). If f is concave, $S = X$, each component function of g is convex and there exists $x \in \text{rint}(D)$ (see Definition 2.6), where D is defined as the set of $x \in X$ where both f and g are defined, such that $g(x) < 0$, then there is no duality gap.

Actually, (from Boyd and Vandenberghe [3]) this condition can be weakened in the case where the component-functions g are actually affine (and f is still concave) and $\text{dom}(f)$ is open. In this case it is sufficient that there exists a feasible solution for the absence of a duality gap. Note that for a minimization problem, the same condition holds as long as f is convex (since a maximization problem can be turned into a minimization problem by using that $\sup f = -\inf(-f)$).

There is also an alternative version of the Slater condition, where $X = \mathbb{R}^n$. This is from Bertsekas et. al [2, p.371]: If the optimal value of the primal problem (5.5) is finite, S is a convex set, f and g are convex functions and there exists $x' \in S$ such that $g(x) < 0$, then there is no duality gap.

There is also a generalized version of the Lagrange duality method. The previous Lagrange duality argument can be done for $g : X \to Z$, where Z is some normed space (see Rynne and Youngston [39] for more on normed spaces) with an ordering that defines a non-negative orthant. From this, one can derive a slightly more general version of the Slater condition (using the separating hyperplane theorem). This version of the Slater condition is Theorem 5 in Luenberger [24] (adapted to the notation of this section): Let X be a normed space and let f be a concave function, defined on a convex subset C of X. Also, let g be a convex function which maps into a normed space Z (with some ordering). Assume there exists some $x' \in C$ such that $g(x') < 0$. Then the optimal value of the Lagrange primal problem equals the optimal value of the Lagrange dual problem, i.e. there is no duality gap.

In particular, since $\mathbb{R}^m$ is a normed space with an ordering that defines a non-negative orthant (componentwise ordering), this generalized Slater condition applies to the Lagrange problem at the beginning of this section.

Finally, note that the Lagrange duality method is quite general, since it holds for an arbitrary vector space X.

5.5 Utility maximization via duality: Complete market, finite Ω $\diamondsuit$

The results of this section hold for general semi-martingale models, not just Brownian motion driven models. This section is based on Schachermayer [40], but some alterations to the approach have been made.

We will now look at an alternative way of solving optimization problems in a complete financial market where $\Omega = \{\omega_1, ..., \omega_M\}$ is finite, and the time T is finite as well. Since the market is complete, there is only one equivalent martingale measure, so $\mathcal{M}^e(S) = \{Q\}$. This new method is a duality method, based on Lagrangean duality (see Section 5.4) and is more general than the straight-forward maximization of Section 5.3. We will derive Theorem 5.8 which summarizes our results, and this theorem can be used to solve the investor's pricing problem in all situations where Ω is finite and the time is discrete and finite.

In the following, define $p_m := P(\omega_m)$ and $q_m := Q(\omega_m)$ for $m = 1, ..., M$.

Consider the following utility maximization problem:

$$u(x) := \sup_{Y \in \mathcal{L}^0(\mathcal{F}_T)} \{E_P[U(Y)] : E_Q[Y] \leq x\} \tag{5.9}$$

Recall from Section 5.2 that if one can solve problem (5.9), one can find a solution to the utility maximization problem over all self-financing portfolio value processes.

Note that the inequality $E_Q[Y] \leq x$ will be replaced by equality in an optimal solution. The reason for this is the following: Assume that Y is an optimal solution of problem (5.9), but that $E_Q[Y] = \sum_m q_m Y(\omega_m) < x$. Define $\epsilon_m := Y(\omega_m)$ for $m = 1, ..., M$. Then, the investor may increase one of the ϵ_m's, until $E_Q[Y] = \sum_m q_m \epsilon_m = x$, and hence increase the objective function, $E_P[U(Y)]$ (because U is increasing, and $E_P[\cdot]$ is linear, so $E_P[U(\cdot)]$ is increasing). Therefore, the original solution Y could not have been optimal.

Our approach for solving this problem can be summarized as follows

(i) Write the problem in component form by defining $\epsilon_m := Y(\omega_m)$ for all $\omega_m \in \Omega$:

$$\begin{array}{ll} \max_{\epsilon} & \sum_m p_m U(\epsilon_m) \\ \text{subject to} & \\ & \sum_m q_m \epsilon_m \leq x. \end{array}$$

(ii) Set up the Lagrange function

$$L(\epsilon_1, ..., \epsilon_M, y) \;=\; \sum_{m=1}^{M} p_m U(\epsilon_m) - y(\sum_{m=1}^{M} q_m \epsilon_m - x).$$

Here, the Lagrange multiplier is denoted by y.

(iii) Find and solve the Lagrange dual problem $\inf_{y \geq 0} L(y)$ by introducing the *KS-conjugate* function $V(y) := \sup_x \{U(x) - xy\}$ and using the properties of this function.

We now execute our plan, using the method from Schachermayer [40].

Define $\epsilon_m := Y(\omega_m)$ for $m = 1, ..., M$ and denote $\epsilon := (\epsilon_1, ..., \epsilon_M)$. Then, problem (5.9) can be rewritten

$$\begin{array}{ll} \sup_{\epsilon \in \mathbb{R}^M} & \sum_m p_m U(\epsilon_m) \\ \text{subject to} & \\ & \sum_m q_m \epsilon_m \leq x. \end{array}$$

Note that this is a concave maximization problem with one constraint and variable $\epsilon \in \mathbb{R}^M$.

Define, for $y \geq 0$, the Lagrange function $L : \mathbb{R}^{M+1} \to \mathbb{R}$

$$L(\epsilon_1, ..., \epsilon_M, y) \;=\; \sum_{m=1}^{M} p_m U(\epsilon_m) - y(\sum_{m=1}^{M} q_m \epsilon_m - x).$$

As in Section 5.4, denote by $L(y)$ the supremum of the Lagrange function

$$L(y) := \sup_{\epsilon} \{\sum_m p_m U(\epsilon_m) - y(\sum_m q_m \epsilon_m - x)\}.$$

From Section 5.4, $L(y)$ gives an upper bound for the optimal value of problem (5.9).

The Lagrangean dual problem is the problem of finding the smallest upper bound. Hence, the dual problem is

$$\begin{aligned}\inf_{y\geq 0} L(y) &= \inf_{y\geq 0}\sup_{\epsilon} L(\epsilon_1,...,\epsilon_M,y)\\ &= \inf_{y\geq 0}\sup_{\epsilon}\{\sum p_m U(\epsilon_m) - y(\sum q_m\epsilon_m - x)\}.\end{aligned}$$

Another definition is useful. The *KS-conjugate* V of a function U is defined by $V(y) = \sup_x\{U(x) - xy\}$, $y > 0$. The KS-conjugate is a technical tool used in duality arguments, and it is defined as it is because it naturally appears in the Lagrange duality argument that follows. The name KS-conjugate is chosen (by me) because V is the version of a conjugate function used by Kramkov and Schachermayer in their papers [21] and [22], as well as by Schachermayer in [40]. The relationship between the KS-conjugate V and the regular conjugate U^* of Chapter 2 is that $V(y) = -U^*(-y)$.

Continuing the computations

$$\begin{aligned}\inf_{y\geq 0} L(y) &= \inf_{y\geq 0}\sup_{\epsilon}\{\sum_m p_m U(\epsilon_m) - y(\sum q_m\epsilon_m - x)\}\\ &= \inf_{y\geq 0}\sup_{\epsilon}\{\sum_m p_m (U(\epsilon_m) - y\frac{q_m}{p_m}\epsilon_m) + xy\}\\ &= \inf_{y\geq 0}\{\sum p_m \sup_{\epsilon_m}\{U(\epsilon_m) - y\frac{q_m}{p_m}\epsilon_m\} + xy\}\\ &= \inf_{y\geq 0}\{\sum p_m V(y\frac{q_m}{p_m}) + xy\}.\end{aligned}$$

Note that the Lagrangean dual problem above decomposes into M different optimization problems, each of the form $\sup_{\epsilon_m}\{U(\epsilon_m) - y\frac{q_m}{p_m}\epsilon_m\}$. These problems are quite easy to solve, and this decomposition is a reason why the duality approach is useful. Also, the decomposed problems above have a very convenient form, since $\sup_{\epsilon_m}\{U(\epsilon_m) - y\frac{q_m}{p_m}\epsilon_m\} = V(y\frac{q_m}{p_m})$, where V is the KS-conjugate of U. Note how the term $\frac{q_m}{p_m}$ arises in the calculations above. This may be viewed as the Radon-Nikodym derivative of the probability measure Q with respect to P. We return to this in Section 5.7.

It is possible to find the ϵ which attains the maximum in V, because U is assumed to be strictly concave and differentiable (recall the definition of a utility function, Definition 5.1). Differentiating and setting equal to zero implies that for all $m = 1,...,M$

$$U'(\epsilon_m^*) = \frac{yq_m}{p_m}.$$

This equation has a unique solution $\epsilon_m^* = (U')^{-1}(\frac{yq_m}{p_m})$ since U' is strictly increasing (recall that U is strictly concave). This is a maximum since U is strictly concave, and there is no boundary to check since the maximization is done over all of $\mathbb{R}$. In order to simplify notation, this will not be inserted directly into $L(\epsilon_1, ..., \epsilon_M, y)$, instead general properties of the KS-conjugate V will be used to find $\inf_{y \geq 0} L(y)$, and then the ϵ_m^*'s will be inserted into the final expression.

One can prove that V has the following properties (see Schachermayer [40]):

- V is finite-valued.
- V is differentiable.
- V is convex.
- $V'(0) = -\infty$

This implies that $v(y) := \sum p_m V(\frac{y q_m}{p_m})$, is also differentiable in y. Now, find $\inf_{y \geq 0} L(y)$ by differentiating and setting equal to zero:

$$v'(y) + x = 0 \text{ so } v'(y) = -x.$$

This is a minimum since v is convex (because V is convex). Hence, the smallest upper bound for our original problem is $L(\epsilon_1^*, ..., \epsilon_M^*, y^*)$.

We now wish to show that there is no *duality gap*, i.e., that

$$u(x) := \sup_{\sum q_m \epsilon_m \leq x} \sum p_m U(\epsilon_m) = L(\epsilon_1^*, ..., \epsilon_M^*, y^*).$$

From Section 5.4, $u(x) \leq L(y^*) = L(\epsilon_1^*, ..., \epsilon_M^*, y^*)$ ($L(y^*)$ is an upper bound).

The Slater condition from Section 5.4 will be used to show that there is no duality gap. Assume that the utility function U is defined for all $x \in \mathbb{R}^M$. The primal problem, problem (5.9), is a concave maximization problem with one affine constraint. Also, $Y(\omega) = 0$ for all $\omega \in \Omega := \{\omega_1, \ldots, \omega_m\}$ is a feasible solution, because $E_Q[Y] = E_Q[0] = 0 < x$, and $0 \in \text{rint}(\mathbb{R}^m)$, (note that $D = \mathbb{R}^M$). (Actually, one can work around assuming that U is defined for all $x \in \mathbb{R}^M$ by defining a new function $\bar{U}$ to be equal to U where it is defined, and very negative where U is not defined, and considering the corresponding utility maximization problem. Then, the optimal values of the two problems will be the same since $Y = 0$ is feasible for both problems, and this gives a non-negative primal value function. Then, the Slater condition can be applied to the new utility maximization problem.)

Hence, $u(x) = L(\epsilon_1^*, ..., \epsilon_M^*, y^*) = \sum p_m U(\epsilon_m^*)$. Also, the optimal solution is $Y^*(\omega_m) = \epsilon_m^* = I(\frac{y^* q_m}{p_m})$ where $I := (U')^{-1}$, and y^* is determined from the equation $v'(y^*) = -x$. From the correspondence between problems (5.2) and (5.3) of Section 5.2, it follows that the optimal solution $Y^* = X_T^{H^*}$, the optimal terminal portfolio value of the utility maximization problem.

There is a connection between the primal optimal value function u and the function v,

$$\begin{aligned} u(x) &= L(\epsilon_1^*, ..., \epsilon_M^*, y^*) \\ &= \inf_{y \geq 0} \{v(y) + xy\} \end{aligned}$$

where the last equality follows from the calculations below:

$$\begin{aligned} \inf_{y>0} \{v(y) + yx\} &= v(y^*(x)) + y^*(x)x \\ &= E_P[V(y^*(x))\frac{dQ}{dP}] + y^*(x)x \\ &= E_P[\sup_{\epsilon_m} \{U(\epsilon_m) - y^*(x)\frac{q_m}{p_m}\epsilon_m\}] + y^*(x)x \\ &= \sum_m p_m U(\epsilon_m^*) - \sum_m q_m y^*(x)\epsilon_m^* + y^*(x)x \\ &= E_P[U(\epsilon^*)] - y^*(x)(E_Q[\epsilon^*] - x) \\ &= L(\epsilon_1^*, ..., \epsilon_M^*, y^*(x)). \end{aligned}$$

Hence, u and v are KS-conjugate by Proposition 2.14 in [40].

We summarize our results in the following theorem, which is a part of Theorem 2.16 in Schachermayer [40]:

Theorem 5.8 *Consider an agent in a complete financial market based on a finite probability space* $(\Omega, \mathcal{F}, P)$*, where* $\mathcal{M}^e(S) = \{Q\}$*. Let* U *be the agent's utility function and* x *her initial endowment. Also, let*

$$u(x) = \sup_{\{Y \in \mathcal{L}^0(\mathcal{F}_T): E_Q[Y] \leq x\}} E_P[U(Y)]$$

and let

$$v(y) = E_P[V(y\frac{dQ}{dP})]$$

where V *is the KS-conjugate of* U*.*

Then:

- $u(x) = \inf_y \{v(y) + xy\}$*, hence (from Proposition 2.14 in [40])* u *and* v *are KS-conjugate.*

- *The optimal terminal portfolio value* Y^* *exists and is uniquely determined by*
$$Y^*(x) = I(y^*\frac{dQ}{dP})$$
where $I := (U')^{-1}$*,* $\frac{dQ}{dP}$ *is the Radon-Nikodym derivative of* Q *w.r.t.* P *and* y^* *is uniquely determined by* $v'(y^*) = -x$*.*

- *In the optimum, the constraint holds with equality:* $E_Q[Y^*] = x$*.*

What is the practical use of this theorem? Theorem 5.8 gives an alternative method for solving the investor's problem (5.9): If one cannot solve the investor's problem directly, Theorem 5.8 says that one can find the function $v(y) := E_P[V(y\frac{dQ}{dP})]$ instead, and use that $u(x) = \inf_y\{v(y) + xy\}$ to find u. The optimal terminal value can then be determined from the formula $Y^*(x) = I(y^*\frac{dQ}{dP})$.

Example 5.9 $\diamondsuit$

Now, consider this duality method for solving the investor's problem (5.9) in the market of Example 5.6 in Section 5.3. This is pearls before swine, but useful for achieving an understanding of the method.

We already know that $\mathcal{M}^e(S) = \{Q\} = \{(\frac{2}{5}, \frac{3}{5})\}$ in our market.

Writing (5.9) for Example 5.6

$$\begin{array}{ll} \max_{\epsilon_1,\epsilon_2} & \frac{1}{3}\ln(\epsilon_1) + \frac{2}{3}\ln\epsilon_2 \\ \textit{such that} & \\ & \frac{2}{5}\epsilon_1 + \frac{3}{5}\epsilon_2 \leq x. \end{array}$$

Theorem 5.8 implies that the optimal terminal value is given by $Y^*(\omega) = I(y^*\frac{dQ}{dP}(\omega))$. $U(x) = \ln(x)$, so $U'(x) = \frac{1}{x}$, hence

$$I(y) := (U')^{-1}(y) = \frac{1}{y}$$

Also $\frac{dQ}{dP}(\omega_1) = \frac{6}{5}$ and $\frac{dQ}{dP}(\omega_1) = \frac{9}{10}$. y^* is determined from $v'(y) = -x$: $V(z) := \sup_x\{\ln(x) - zx\}$, by differentiating and setting equal to zero $V(z) = -(1 + \ln(z))$. Now $v(y) := E[V(y\frac{dQ}{dP})]$, so

$$v(y) = -1 - \ln(\tfrac{6}{5})\tfrac{1}{3} - \ln(\tfrac{9}{10}) - \ln(y).$$

Hence, $-\frac{1}{y} = v'(y) = -x$, so $y^* = \frac{1}{x}$. Therefore, the optimal solution is

$$\begin{aligned} Y^*(x) &= I(y^*(x)\frac{dQ}{dP}) \\ &= \begin{cases} \frac{5x}{6} &, \omega_1 \\ \frac{10x}{9} &, \omega_2 \end{cases} . \end{aligned}$$

As mentioned, $Y^* = X_T^{H^*}$, the optimal terminal portfolio value of the utility maximization problem. Hence, $X_T^{H^*}(\omega_1) = \frac{5x}{6}$ and $X_T^{H^*}(\omega_2) = \frac{10x}{9}$ are the optimal values of the portfolio in ω_1, ω_2 respectively. From this, one can derive the optimal trading strategy (since the market is complete) by solving the linear system of equations

$$\begin{aligned} H_0(1) + 4H_1(1) &= x \\ H_0(1) + 7H_1(1) &= \tfrac{5x}{6} \\ H_0(1) + 2H_1(1) &= \tfrac{10x}{9}. \end{aligned}$$

where $H_n(1)$ denotes the amount of asset n the agent should choose to hold from time 0 to time 1. This system of equations has solution $H(1) = (H_0(1), H_1(1)) =$

$(\frac{11x}{9}, -\frac{x}{18})$. *Note that this is the same as the previously derived optimal trading strategy.*

Now, two different approaches for solving the utility maximization problem have been presented: The direct method of Section 5.3, and the duality method of this section. One of the major advantages of the duality method is that there is no system of (possibly) non-linear equations to solve, which was an issue with the direct approach. Another advantage with the duality method, even for small examples, is that there are far fewer computations than the direct method. The duality method is also simple to program. Appendix A contains two Matlab-programs. The first one checks if a financial market is complete (since the market must be complete in order to apply Theorem 5.8). The second program applies Theorem 5.8 to a complete financial market where the investor has utility function $U(x) = \ln(x)$, and hence derives the optimal terminal portfolio value for the agent. Then, the program uses the optimal terminal portfolio value to find the optimal trading strategy (by computing backwards). Both the programs work for markets with finite, discrete time and finite scenario space. One can use the program for different utility functions as well, by computing the alternative formula for y^* and inserting this into the program. A final advantage with the duality method is that it generalizable, as will be shown.

5.6 Utility maximization via duality: Incomplete market, finite Ω $\diamondsuit$

The results in this section hold for general semi-martingale models. It is based on Schachermayer [40], but some alterations have been made to the approach.

Now that we know how to solve the investor's problem for a complete market where Ω is finite, it is natural to look at what happens in an incomplete market (also for finite Ω). So, assume $|\mathcal{M}^e(S)| = \infty$, i.e. that the set of equivalent martingale measures, contain infinitely many elements. Hence, there is no arbitrage and the market is incomplete.

The investor's problem now takes the form

$$\begin{array}{lllll} \max & E[U(Y)] & & & \\ \text{subject to} & & & & \\ & E_Q[Y] & \leq x & \text{for all} & Q \in \mathcal{M}^e(S). \end{array} \tag{5.10}$$

Recall that from Section 5.2, this problem can be solved instead of the original utility maximization problem $\sup_{H \in \mathcal{H}} E[U(X_T^H)]$.

The difficulty in the investor's problem (5.10) is that there are infinitely many constraints. If one could reduce these infinitely many constraints to a finite number, the Lagrange duality theory of Section 5.4 would apply. Luckily, it is actually possible to reduce the number of constraints. In order to do this, observe the following lemma:

Lemma 5.10 *For $Y \in \mathcal{L}^0(\mathcal{F}_T)$, $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^e(S)$ if and only if $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^a(S)$, where $\mathcal{M}^a(S)$ is the set of all probability measures that are absolutely continuous w.r.t. to P and are such that the price process S is a Q-martingale.*

Proof: $\diamondsuit$
One direction is clear, since $\mathcal{M}^e(S) \subseteq \mathcal{M}^a(S)$.

As for the other direction: Assume that $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^e(S)$. Take $Q \in \mathcal{M}^a(S)$, and take $Q' \in \mathcal{M}^e(S)$. Then, for every $\lambda \in (0,1)$, $Q_\lambda := \lambda Q' + (1-\lambda)Q \in \mathcal{M}^e(S)$. Hence, $E_Q[Y] = \lim_{\lambda \to 0} E_{Q_\lambda}[Y] \leq x$. Therefore $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^a(S)$.

$\square$

Because of Lemma 5.10, consider the constraints $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^a(S)$ from now on. Why is this a better set to work with? The following Lemma 5.11 is the reason. Note that in Lemma 5.11 the set $\mathcal{M}^a(S)$ is identified with the corresponding set of probability vectors in $\mathbb{R}^m$, and this set is denoted by $\mathcal{M}^a(S)$ as well.

Lemma 5.11 *$\mathcal{M}^a(S)$ is a polytope in $\mathbb{R}^M$. Hence, $\mathcal{M}^a(S)$ is the convex hull of its finitely many extreme points, and it is a compact set.*

Proof: $\diamondsuit$

Denote by $q_m := Q(\omega_m)$ for all $\omega_m \in \Omega$, and $q := (q_1, ..., q_m)$. We want to show that $\mathcal{M}^a(S)$ is a polytope. From Theorem 2.5 in Chapter 2, we know that $\mathcal{M}^a(S)$ is a polytope if and only if it is a bounded polyhedron. Hence, it suffices to show that $\mathcal{M}^a(S)$ is bounded and can be described as the intersection of a finite number of closed halfspaces (from Definition 2.2). That is, we want to prove that $\mathcal{M}^a(S)$ is bounded and that it is the solution set of a finite number of (non-strict) linear inequalities (see comment after Definition 2.2). $\mathcal{M}^a(S)$ is bounded because it is contained in the unit ball of $\mathbb{R}^M$. It is clear from the definition of $\mathcal{M}^a(S)$ that it is the solution set of finitely many non-strict linear inequalities (because the martingale condition can be written as a finite set of linear equalities in q, which can be rewritten as a finite set of linear inequalities in q).

Hence, $\mathcal{M}^a(S)$ is a polytope, and therefore it is compact, and it is the convex hull of its finitely many extreme points, from Lemma 2.4 and Definition 2.3.

$\square$

Hence, all of $\mathcal{M}^a(S)$ can be described by a finite subset, namely its extreme points. This is very good news, because we are trying to reduce the infinitely many constraints $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^a(S)$, to a finite number of constraints.

Lemma 5.12 *Let $Q_1, ..., Q_K$ be the extreme points of $\mathcal{M}^a(S)$ and let $Y \in \mathcal{L}^0(\mathcal{F}_T)$. Then $E_{Q_k}[Y] \leq x$ for $k = 1, ..., K$ if and only if $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^a(S)$.*

Proof: $\diamondsuit$
We show this by combining convexity and linearity of the expectation operator.

- Assume $E_{Q_k}[Y] \leq x$ for $k = 1, ..., K$. We want to show that $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^a(S)$. So take $Q \in \mathcal{M}^a(S)$. Then Q can be written as a convex combination of the Q_k's, so $Q = \sum_{k=1}^K \lambda_k Q_k$, where $\sum_k \lambda_k = 1$, $\lambda_k \geq 0$ for all $k = 1, ..., K$. But then:

$$\begin{aligned} E_Q[Y] &= \textstyle\sum_{m=1}^M q_m Y(\omega_m) \\ &= \textstyle\sum_m (\sum_k \lambda_k q_m^k) Y(\omega_m) \\ &= \textstyle\sum_k \lambda_k (\sum_m q_m^k Y(\omega_m)) \\ &= \textstyle\sum_k \lambda_k E_{Q_k}[Y] \\ &\leq \textstyle\sum_k \lambda_k x \\ &= x \textstyle\sum_k \lambda_k \\ &= x \end{aligned}$$

  where $q_m^k := Q_k(\omega_m)$ for all $\omega_m \in \Omega$, $k = 1, \ldots, K$ and the inequality follows from that $Q_k \in \{Q_1, ..., Q_K\}$ for $k = 1, ..., K$ and our assumption. Hence, $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^a(S)$.

- Assume $E_Q[Y] \leq x$ for all $Q \in \mathcal{M}^a(S)$. Then, clearly, $E_{Q_k}[Y] \leq x$ for $k = 1, ..., K$, since $Q_k \in \mathcal{M}^a(S)$ for $k = 1, ..., K$.

$\square$

Lemma 5.12 implies that the investor's problem can be written as a concave maximization problem over $\mathbb{R}^M$ with a *finite* number of constraints

$$\begin{array}{lll} \max & E[U(Y)] & \\ \text{subject to} & & \\ & E_{Q_k}[Y] & \leq \; x \text{ for } k = 1, ..., K. \end{array}$$

Or, in component form, defining $\epsilon_m := Y(\omega_m)$ for $m = 1, \ldots, M$

$$\begin{array}{lll} \max & \sum_m p_m U(\epsilon_m) & \\ \text{subject to} & & \\ & \sum_m q_m^k \epsilon_m & \leq \; x \text{ for } k = 1, ..., K. \end{array} \tag{5.11}$$

The form of this problem is suitable for the Lagrange duality method of Section 5.4. Therefore, define the Lagrange function

$$\begin{aligned} L(\epsilon_1, ..., \epsilon_M, \eta_1, ..., \eta_K) &= \textstyle\sum_m p_m U(\epsilon_m) - \sum_{k=1}^K \eta_k (\sum_m q_m^k \epsilon_m - x) \\ &= \textstyle\sum_m p_m (U(\epsilon_m) - \sum_k \frac{\epsilon_m \eta_k q_m^k}{p_m}) + x \sum_k \eta_k \end{aligned}$$

where $\epsilon_m \in \text{dom}(U)$ for $m = 1, ..., M$ and $\eta_k \in \mathbb{R}_+$ for $k = 1, ..., K$ are the Lagrange dual variables.

Some notation is useful in order to make the Lagrange function similar to the one in the complete case of Section 5.5. Let $y := \sum_k \eta_k$, $\mu_k := \frac{\eta_k}{y}$ for $k = 1, ..., K$, $\mu := (\mu_1, ..., \mu_K)$ and $Q = \sum_k \mu_k Q^k$.

Note that if $(\eta_1, ..., \eta_K)$ runs through $\mathbb{R}_+^K$, then the couple (y, Q) runs through $\mathbb{R} \times \mathcal{M}^a(S)$ (since $\mathcal{M}^a(S)$ is the convex hull of its extreme points $Q_1, ..., Q_K$).

Hence

$$\begin{aligned} L(\epsilon_1, ..., \epsilon_M, \eta_1, ..., \eta_K) &= E[U(Y)] - y(E_Q[Y] - x) \\ &= \textstyle\sum_m p_m(U(\epsilon_m) - \frac{yq_m\epsilon_m}{p_m}) + yx. \end{aligned}$$

where $\epsilon_m \in \mathrm{dom}(U)$ for all m, and the first equation is true because

$$\begin{aligned} E[U(Y)] - y(E_Q[Y] - x) &= \textstyle\sum_{m=1}^M p_m U(\epsilon_m) - \\ &\quad (\eta_1 + ... + \eta_K)(\textstyle\sum_{m=1}^M q_m\epsilon_m - x) \\ &= \textstyle\sum_m p_m U(\epsilon_m) - \\ &\quad (\textstyle\sum_{k=1}^K \eta_k)(\sum_m(\sum_k q_m^k \mu_k)\epsilon_m - x) \\ &= \textstyle\sum_m p_m U(\epsilon_m) - \\ &\quad (\textstyle\sum_k \eta_k)(\sum_m(\sum_k q_m^k \frac{\eta_k}{\sum_k \eta_k})\epsilon_m - x) \\ &= \textstyle\sum_m p_m U(\epsilon_m) - \sum_m(\sum_k q_m^k \eta_k)\epsilon_m + \\ &\quad x(\textstyle\sum_k \eta_k) \\ &= \textstyle\sum_m p_m(U(\epsilon_m) - \sum_k \frac{\eta_k q_m^k \epsilon_m}{p_m}) + \sum_k \eta_k x \end{aligned}$$

which is the previous expression for the Lagrange function.

Hence

$$L(\epsilon_1, ..., \epsilon_M, y, Q) = \textstyle\sum_m p_m(U(\epsilon_m) - \frac{yq_m}{p_m}\epsilon_m) + xy$$

where $\epsilon_m \in \mathrm{dom}(U)$ for all m, $y \geq 0$ and $Q = (q_1, ..., q_M) \in \mathcal{M}^a(S)$. This is the same expression as in the complete case, except that Q is not a fixed measure, but varies in $\mathcal{M}^a(S)$.

From Lagrange duality (see Section 5.4), $u(x) \leq \sup_\epsilon L(\epsilon_1, ..., \epsilon_M, y, Q)$ for all $y \geq 0$ and $Q \in \mathcal{M}^a(S)$.

Note that

$$\begin{aligned} \sup_{\epsilon_1,...,\epsilon_M} L(\epsilon_1, ..., \epsilon_M, y, Q) &= \sup_{\epsilon_1,...,\epsilon_M} \textstyle\sum_m p_m(U(\epsilon_m) - \frac{yq_m}{p_m}\epsilon_m) + yx \\ &= \textstyle\sum_m p_m \sup_{\epsilon_m}\{U(\epsilon_m) - \frac{yq_m}{p_m}\epsilon_m\} + yx \\ &= \textstyle\sum_m p_m V(\frac{yq_m}{p_m}) + yx. \end{aligned}$$

for $y \geq 0, Q \in \mathcal{M}^a(S)$, where V is the KS-conjugate of U.

Now consider the Lagrangian dual problem of finding the smallest upper bound on the optimal primal value generated by the Lagrange function

$$\begin{aligned} \inf_{y\geq 0, Q\in\mathcal{M}^a(S)} &\sup_{\epsilon_1,...,\epsilon_M} L(\epsilon_1, ..., \epsilon_M, y, Q) = \\ &\inf_{y\geq 0, Q\in\mathcal{M}^a(S)}\{\textstyle\sum_m p_m V(\frac{q_m y}{p_m}) + yx\} \end{aligned}$$

First consider $\inf_{Q\in\mathcal{M}^a(S)}\sum_m p_m V(\frac{q_m y}{p_m})+yx$. The function

$$Q\mapsto\sum_m p_m V(\frac{q_m y}{p_m})+yx$$

is a continuous function, and the set $\mathcal{M}^a(S)$ is compact. Hence, the extreme value theorem (see any book on multi-variable calculus) implies that the function $Q\mapsto\sum_m p_m V(\frac{q_m y}{p_m})+yx$ achieves its minimum $Q^*(y)$ on $\mathcal{M}^a(S)$. Since V is a strictly convex function, the minimum is unique. Actually, one can prove that $Q^*(y)\in\mathcal{M}^e(S)$. This is done in Schachermayer [40].

Now, define $v(y):=\sum_m p_m V(\frac{yq^*_m(y)}{p_m})=\inf_{Q\in\mathcal{M}^a(S)}\sum_m p_m V(\frac{yq_m}{p_m})$, and consider the Lagrange dual problem

$$\inf_{y\geq 0}\inf_{Q\in\mathcal{M}^a(S)}\sum_m p_m V(\frac{q_m y}{p_m})+yx \quad = \quad \inf_{y\geq 0}\{v(y)+xy\}.$$

v is a differentiable function (from properties of V and the definition of v). Hence, this minimization problem can be solved by differentiating and setting equal to zero. Therefore y^* is defined by $v'(y^*)=-x$, as in the complete case, and for the same reasons as before $y^*\geq 0$.

Let $Q^*:=Q^*(y^*)$ and $\epsilon^*_m:=I(\frac{y^*q^*_m}{p_m})$ for all $m=1,...,M$. We want to show that there is no duality gap, i.e. that $u(x)=L(\epsilon^*_1,...,\epsilon^*_M,y^*,Q^*)$. From Lagrange duality, $u(x)\leq L(\epsilon^*_1,...,\epsilon^*_M,y^*,Q^*)$. Again, the Slater condition proves that there is no duality gap: Problem (5.11) is a concave maximization problem with finitely many affine inequalities as constraints. As in Section 5.5, $Y=0$ is a strictly feasible solution in $\text{rint}(D)=\text{rint}(\mathbb{R}^M)$. Hence, the Slater condition implies that there is no duality gap. Therefore, $u(x)=L(\epsilon^*_1,...,\epsilon^*_M,y^*,Q^*)$.

Finally, we get theorem similar to Theorem 5.8 for the incomplete case:

Theorem 5.13 *(Utility maximization: finite Ω, incomplete market)*
Consider a financial market S based on a finite probability space $(\Omega,\mathcal{F},P)$ and let $\mathcal{M}^e(S)\neq\emptyset$. Consider an agent with a utility function U. Let:

$$\begin{aligned} u(x) &:= \sup_{Y\in\mathcal{L}^0}\{E[U(Y)]: E_Q[Y]\leq x \text{ for all } Q\in\mathcal{M}^e(S)\} \\ v(y) &:= \inf_Q\{E[V(y\tfrac{dQ}{dP})]: Q\in\mathcal{M}^a(S)\} \end{aligned}$$

Then:

(i) u and v are KS-conjugate.

(ii) Y^* and Q^* that optimize respectively u and v above exist, are unique, and satisfy $Q^*\in\mathcal{M}^e(S)$ as well as:

$$Y^*(x) \quad = \quad I(y^*\tfrac{dQ^*}{dP})$$

where $I=(U')^{-1}$ and y^* is defined as the unique solution to the equation $v'(y^*)=-x$.

Proof: Everything in (ii) follows from the previous derivation. As for (i):

$$\begin{aligned} u(x) &= L(\epsilon_1^*, ..., \epsilon_M^*, y^*, Q^*) \\ &= \inf_{y \geq 0}\{v(y) + xy\} \end{aligned}$$

so from Proposition 2.14 in [40], u and v are KS-conjugate. $\square$

5.7 Utility maximization via duality: The general case

As in the two previous sections, the results of this section hold for a general semi-martingale model. This section is based on the papers [21] and [22] by Kramkov and Schachermayer.

Until now it has been assumed that the scenario space Ω is finite when working with the utility maximization problem. The goal of this section is to generalize the theory of the previous sections to the case where Ω is arbitrary.

In the following, assume that $\mathcal{M}^e(S) \neq \emptyset$, so that the market has no arbitrage, but it may be incomplete. This is a very general, but also very realistic situation. The question is: *What, if any, extra conditions need to be imposed in order to get a theorem analogous to Theorem 5.13 for the case where Ω is arbitrary?*

Recall from Section 5.1 that the utility function $U : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ is assumed to be increasing on $\mathbb{R}$, continuous on $\{U > -\infty\}$, differentiable and strictly concave on the interior of $\{U > -\infty\}$ and that $\lim_{x\to\infty} U'(x) = 0$. In that section, we also distinguished between two different cases: negative wealth not allowed and negative wealth allowed, and defined some extra conditions on the utility function depending on which of these cases applied. So far, there has really been no difference between the two cases, but in the situation where Ω is arbitrary, the cases separate. Here, only case 1 will be considered. This is the case where negative wealth is not allowed. Hence we assume (as in Section 5.1) that $\lim_{x\to 0} U'(x) = \infty$. The idea for solving case 2 is the same, and can be found in Kramkov and Schachermayer [22].

It turns out that one does not need to add a lot of extra conditions in order to get a theorem similar to Theorem 5.13 also in the present, more general case. Actually, the answer to our question is that we have to choose some clever sets for the terminal portfolio value and Q to vary in, and these sets depend on whether we are in case 1 or case 2. Also the utility function U must satisfy what Kramkov and Schachermayer [22] refer to as reasonable asymptotic elasticity.

A utility function U satisfies *reasonable asymptotic elasticity* (in case 1) if

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

One can show that this condition holds if for instance $U(x) = \ln(x)$.

The economic principles behind the concept of reasonable asymptotic elasticity is that $\frac{xU'(x)}{U(x)} = \frac{U'(x)}{\frac{U(x)}{x}}$ is the *marginal utility* divided by the *average utility.*

Since a utility function has decreasing marginal utility (see Definition 5.1), intuition suggests that as x gets bigger, the marginal utility should become smaller than the average utility. This is what the concept of reasonable asymptotic elasticity states.

Even though one does not need a lot of extra conditions on the market in the general case, things are a lot more complicated than in our previous finite Ω situation. Therefore, we begin with some intuition of how to proceed with the proof of the generalization of Theorem 5.13, before actually embarking on the proof.

The proof uses similar ideas as the previous sections. However, a difficulty is that many of the previous arguments are based on results in multivariate calculus which do not generalize to infinite dimension. Instead, one must use functional analysis and an infinite dimensional version of the MiniMax theorem.

The MiniMax theorem in infinite dimension states (see Kramkov and Schachermayer [22] or Komiya [20]):

Theorem 5.14 *(MiniMax theorem)*

Let E and F be a pair of locally convex paired vector spaces and let $C \subseteq E$, $D \subseteq F$ be convex subsets of these spaces. Also, let $L(x, y)$ be a function defined on $C \times D$ which is concave and upper semi-continuous in the x-variable and convex and lower semi-continuous in the y-variable. If in addition, at least one of the sets C, D is compact, there exists a saddle point $(\epsilon^*, \mu^*) \in C \times D$ such that

$$L(\epsilon^*, \mu^*) = \sup_{\epsilon \in C} \inf_{\mu \in D} L(\epsilon, \mu) = \inf_{\mu \in D} \sup_{\epsilon \in C} L(\epsilon, \mu).$$

The idea of the following proof is to first prove things in an abstract setting, using spaces that satisfy the MiniMax theorem. Then, one chooses appropriate spaces for the primal and dual variables to vary in, so that one falls into the abstract setting.

In order to formulate the theorem, some more notation is needed.

Let $\mathcal{X}(x)$ be the set of all value processes that are greater than or equal to 0 with start value x, i.e.

$$\mathcal{X}(x) := \{X : X_t = x + \int_0^t H_s dS_s \geq 0 \text{ for all } t, H \in \mathcal{H}\}$$

where $\mathcal{H}$ consists of all predictable stochastic processes that are integrable with respect to S.

Let also $\mathcal{X} := \mathcal{X}(1)$.

The investor wants to solve the following version of problem (5.1)

$$u(x) = \sup_{X \in \mathcal{X}(x)} E_P[U(X_T)]. \tag{5.12}$$

Note that this is the original utility maximization problem, problem (5.3). Here, the function u is the *value function* of the investor's problem.

As usual, the KS-conjugate in the concave sense (KS for Kramkov and Schachermayer) of the utility function U is defined by $V(y) = \sup_{x>0}\{U(x) - xy\}$ for $y > 0$.

Define

$$\mathcal{Y} := \{Y \geq 0 : Y_0 = 1, XY \text{ is a supermartingale for all } X \in \mathcal{X}\} \tag{5.13}$$

and let $\mathcal{Y}(y) := y\mathcal{Y} = \{yY : Y \in \mathcal{Y}\}$ for $y > 0$.

Consider the following dual optimization problem

$$v(y) = \inf_{Y \in \mathcal{Y}(y)} E_P[V(Y_T)]. \tag{5.14}$$

So v is the value function of the dual problem.

Then, the following theorem, which is a combination of Theorem 1 and Theorem 2 in Kramkov and Schachermayer [22], holds.

Theorem 5.15 *Assume that $u(x) < \infty$ for some x and $v(y) < \infty$ for all $y > 0$. Then*

(i) $v(y) = \sup_{x>0}\{u(x) - xy\}$, for all $y > 0$, so u and v are KS-conjugate (see Proposition 2.14 in Schachermayer [40]).

(ii) $v(y) = \inf_{Q \in \mathcal{M}^e(S)} E_P[V(y\frac{dQ}{dP})]$ for all $y > 0$. Here $\frac{dQ}{dP}$ denotes the Radon-Nikodym derivative of Q w.r.t. P (that is, $dQ = (\frac{dQ}{dP})dP$).

(iii) For a fixed x^*, $y^* = u'(x^*)$ and Q^* that optimizes the infimum in (ii), the optimal terminal value $X_x^*(T)$ in the primal problem will satisfy

$$U'(X_x^*(T)) = y^* \frac{dQ^*}{dP}. \tag{5.15}$$

The proof is from Kramkov and Schachermayer [21], [22], some details have been filled in.

Proof: $\diamondsuit$

(i) : First, define an abstract setting: Let $\mathcal{C}$ and $\mathcal{D}$ be two subsets of $\mathcal{L}^0_+(\Omega, \mathcal{F}, P)$ (the set of all non-negative P-integrable functions from Ω into $\mathbb{R}$) such that

(a) $\mathcal{C}$ and $\mathcal{D}$ are convex, solid (i.e. $h \in \mathcal{D}$ and $0 \leq g \leq h$ implies that $g \in \mathcal{C}$) and closed in the convergence of measure topology.

(b)

$$\begin{aligned} g \in \mathcal{C} \quad &\iff \quad E_P[gh] \leq 1 \text{ for all } h \in \mathcal{D} \text{ and} \\ h \in \mathcal{D} \quad &\iff \quad E_P[gh] \leq 1 \text{ for all } g \in \mathcal{C}. \end{aligned}$$

(c) $\mathcal{C}$ is a bounded subset of $\mathcal{L}^0(\Omega, \mathcal{F}, P)$ such that $1 \in \mathcal{C}$ (where 1 denotes the function identically equal to 1).

Let $\mathcal{C}(x) = x\mathcal{C}$ and $\mathcal{D}(y) = y\mathcal{D}$.

Let, as before, U be the utility function and V its KS-conjugate. Consider the following abstract versions of the previous primal and dual optimization problems

$$\begin{aligned} u(x) &= \sup_{g \in \mathcal{C}(x)} E_P[U(g)] \\ v(y) &= \inf_{h \in \mathcal{D}(y)} E_P[V(h)]. \end{aligned}$$

Abstract version of statement (i): $v(y) = \sup_{x>0}\{u(x) - xy\}$.

Proof of abstract version of statement (i): For $n > 0$, let $\mathcal{B}_n$ be the positive elements in the ball with radius n in $\mathcal{L}^\infty(\Omega, \mathcal{F}, P)$, so $\mathcal{B}_n = \{g : 0 \le g \le n\}$. $\mathcal{B}_n$ is a $\sigma(\mathcal{L}^\infty, \mathcal{L}^0)$-compact set (this follows from Alaoglu's theorem plus the fact that a closed subset of a compact set is compact).

Since $1 \in \mathcal{C}$, $E_P[g] = E_P[g1] \le y < \infty$ for all $yg \in y\mathcal{D} = \mathcal{D}(y)$. This, plus the assumptions on $\mathcal{D}$, imply that $\mathcal{D}(y)$ is a closed, convex subset of $\mathcal{L}^\infty(\Omega, \mathcal{F}, P)$. Therefore, the MiniMax theorem holds. This implies

$$\sup_{g \in \mathcal{B}_n} \inf_{h \in \mathcal{D}(y)} E_P[U(g) - gh] = \inf_{h \in \mathcal{D}(y)} \sup_{g \in \mathcal{B}_n} E_P[U(g) - gh].$$

From the dual relations (assumption (b)) between $\mathcal{C}$ and $\mathcal{D}$

$$g \in \mathcal{C} \iff E_P[gh] \le 1 \text{ for all } h \in \mathcal{D}.$$

So,

$$g \in \mathcal{C} \iff \sup_{h \in \mathcal{D}} E_P[gh] \le 1.$$

Hence,

$$\bar{g} \in x\mathcal{C} = \mathcal{C}(x) \iff \sup_{h \in \mathcal{D}} E_P[\frac{\bar{g}}{x} h] \le 1.$$

Therefore, this happens if and only if

$$\sup_{h \in \mathcal{D}} E_P[\bar{g}h] \le x.$$

So,

$$\bar{g} \in \mathcal{C}(x) \iff \sup_{\bar{h} \in \mathcal{D}(y)} E_P[\bar{g}\frac{\bar{h}}{y}] \leq x.$$

Finally,

$$\bar{g} \in \mathcal{C}(x) \iff \sup_{\bar{h} \in \mathcal{D}(y)} E_P[\bar{g}\bar{h}] \leq xy.$$

Hence,

$$\begin{aligned} \lim_{n\to\infty} \sup_{g \in \mathcal{B}_n} \inf_{h \in \mathcal{D}(y)} E_P[U(g) - gh] &= \lim_{n\to\infty} \sup_{g \in \mathcal{B}_n} \{E_P[U(g)] - \sup_{h \in \mathcal{D}(y)} E_P[gh]\} \\ &= \sup_{x>0} \sup_{g \in \mathcal{C}(x)} E_P[U(g) - xy] \\ &= \sup_{x>0} \{u(x) - xy\}. \end{aligned}$$

The second to last equality follows from $\sup_{h \in \mathcal{D}(y)} E_P[gh] \leq xy$ in $\mathcal{C}(x)$, and since we want to make this expression small, it is sufficient to consider $g \in \mathcal{C}(x)$. The last equality follows from the definition of u, plus the fact that xy is deterministic.

On the other hand,

$$\begin{aligned} \inf_{h \in \mathcal{D}(y)} \sup_{g \in \mathcal{B}_n} E_P[U(g) - gh] &= \inf_{h \in \mathcal{D}(y)} E_P[V_n(h)] \\ &:= v_n(y). \end{aligned}$$

where $V_n(y) := \sup_{0 \leq x \leq n}\{U(x) - xy\}$.

Hence, it suffices to prove that

$$\lim_{n\to\infty} v_n(y) = \lim_{n\to\infty} \inf_{h \in \mathcal{D}(y)} E_P[V_n(h)] = v(y), \text{ for } y > 0. \tag{5.16}$$

(because $\sup_g \inf_h E_P[U(g) - gh] = \inf_h \sup_g E_P[U(g) - gh]$).

Note that $v_n \leq v$ because

$$\begin{aligned} v_n(y) &= \inf_{h \in \mathcal{D}(y)} E_P[V_n(h)] \\ &= \inf_{h \in \mathcal{D}(y)} E_P[\sup_{0 \leq x \leq n}\{u(x) - xy\}] \end{aligned}$$

and

$$\begin{aligned} v(y) &= \inf_{Y \in \mathcal{Y}(y)} E_P[V(Y_T)] \\ &= \inf_Y E_P[\sup_{x>0}\{u(x) - xY_T\}] \end{aligned}$$

and we also have $\mathcal{D}(y) \subseteq \mathcal{Y}(y)$, $\{0 < x \leq n\} \subseteq \{x > 0\}$.

Let $(h_n)_{n\geq 1}$ be a sequence in $\mathcal{D}(y)$ such that

$$\lim_{n\to\infty} E_P[V_n(h_n)] = \lim_{n\to\infty} v_n(y).$$

(such a sequence exists from the definition of infimum).

The following lemma holds.

Lemma (i): Let $(f_n)_{n\geq 1}$ be a sequence of positive random variables. Then there exists a sequence $(g_n)_{n\geq 1} \in \text{co}\{f_n, f_{n+1}, ...\}$, $n \geq 1$ which converges almost everywhere towards a variable g with values in $[0, \infty]$.

For a proof of Lemma (i), see Kramkov and Schachermayer [21].

Lemma (i) implies that there exists a sequence $(f_n)_n$ in $\text{co}\{h_n, h_{n+1}, ...\}$ which converges P-a.s. towards a variable h. This h must be in $\mathcal{D}(y)$, since $\mathcal{D}(y)$ is closed under convergence of measure (by assumption).

Since $V_n(y) = V(y)$ for $y \geq I(1) \geq I(n)$, where $I := (U')^{-1}$, Lemma 3.2 in [21] implies that the sequence $\{(V_n(f_n))^-\}_{n\geq 1}$ is uniformly integrable (here, $a^- := \max\{-a, 0\}$).

Convexity of V_n and Fatou's Lemma now implies that

$$\begin{aligned}
\lim_{n\to\infty} E[V_n(h_n)] &\geq \liminf_{n\to\infty} E[V_n(f_n)] \\
&\geq E[V(h)] \\
&\geq v(y).
\end{aligned}$$

This gives equation (5.16), since

$$v_n \leq v \Rightarrow \lim_{n\to\infty} v_n \leq v$$

and

$$\begin{aligned}
v(y) &\leq \lim_{n\to\infty} E_P[V_n(h_n)] \\
&= \lim_{n\to\infty} v_n(y).
\end{aligned}$$

This completes the proof of the abstract version of (i).

The next step is to *prove (i) from its abstract version*:

Let

$$\begin{aligned}
\mathcal{C}(x) &= \{g \in \mathcal{L}^0(\Omega, \mathcal{F}, P) : 0 \leq g \leq X_T \text{ for some } X \in \mathcal{X}(x)\} \\
\mathcal{D}(y) &= \{h \in \mathcal{L}^0(\Omega, \mathcal{F}, P) : 0 \leq h \leq Y_T \text{ for some } Y \in \mathcal{Y}(y)\}.
\end{aligned}$$

Then

$$u(x) = \sup_{g\in\mathcal{C}(x)} E_P[U(g)] v(y) = \inf_{h\in\mathcal{D}(y)} E_P[V(h)]$$

where u, v are the original optimal value functions of the primal and dual problems.

From Proposition 3.1 in Kramkov and Schachermayer [21], $\mathcal{C}(x)$ and $\mathcal{D}(y)$ will satisfy the conditions (a)-(c). Hence, the claim in (i) is proved.

(ii) : Moving on to *claim (ii)*, we want to prove that $v(y) = \inf_{Q \in \mathcal{M}^e(S)} E[V(y\frac{dQ}{dP})]$.

Proof of claim (ii): We will use the following proposition from Kramkov and Schachermayer [22], see this article for the proof.

Proposition 1: Assume that U is such that $U'(0) = \infty, U'(\infty) = 0$. Assume also that $\mathcal{C}, \mathcal{D}$ are sets such that $\mathcal{C}$ is bounded in $\mathcal{L}^0(\Omega, \mathcal{F}, P)$, $1 \in \mathcal{C}$, $g \in \mathcal{C}$ if and only if $E_P[gh] \leq 1$ for all $h \in \mathcal{D}$, $h \in \mathcal{D}$ if and only if $h \geq 0$ and $E_P[gh] \leq 1$ for all $g \in \mathcal{C}$ and that $\bar{\mathcal{D}}$ is a convex subset of $\mathcal{D}$ such that:

- For all $g \in \mathcal{C}$ we have $\sup_{h \in \bar{\mathcal{D}}} E_P[gh] = \sup_{h \in \mathcal{D}} E_P[gh]$.
- $\bar{\mathcal{D}}$ is closed under countable convex combinations.

Then, $v(y) = \inf_{h \in \bar{\mathcal{D}}} E_P[V(yh)]$.

Proposition 1 is used to prove (ii). Let $\bar{\mathcal{D}}$ be the set of Radon-Nikodym derivatives for equivalent martingale measures

$$\bar{\mathcal{D}} = \{h = \frac{dQ}{dP} : Q \in \mathcal{M}^e(S)\}. \tag{5.17}$$

$\bar{\mathcal{D}}$ is closed under countable convex combinations: If $(h_n)_n \subseteq \bar{\mathcal{D}}$, $(\lambda_n)_n$, $\lambda_n \geq 0$ for all n, $\sum_n \lambda_n = 1$, then there exists a sequence $(Q_n)_n \in \mathcal{M}^e(S)$ such that $h_n = \frac{dQ_n}{dP}$ for $n = 1, 2, ...$ Then

$$\begin{aligned}
\sum_n \lambda_n h_n &= \sum_n \lambda_n \frac{dQ_n}{dP} \\
&= \frac{d}{dP} \sum_n \lambda_n Q_n \\
&= \frac{d\bar{Q}}{dP}
\end{aligned}$$

where the middle equality follows from linearity of the integral, and the last equality follows from that $\sum_n \lambda_n Q_n \in \mathcal{M}^e(S)$.

In addition

$$g \in \mathcal{C} \iff g \geq 0 \text{ and } E_Q[g] \leq 1 \text{ for all } Q \in \mathcal{M}^e(S).$$

Hence, all the conditions of Proposition 1 are satisfied, and claim (ii) follows from the proposition plus the definition of $\bar{\mathcal{D}}$.

(iii) : Finally, we prove claim (iii). Assume again the generalized setting of (i). From the same type of argument as it (i), it will suffice to prove the following generalized version of claim (iii):

Abstract version of (iii): Under the same assumptions on $U, \mathcal{C}, \mathcal{D}, v$ and u as in (i), we have: The optimal solution $\bar{g}(x) \in \mathcal{C}(x)$ of $\sup_{g \in \mathcal{C}(x)} E_P[U(g)]$ exists for $x > 0$ and it is unique. If $\bar{y} = u'(\bar{x})$ for $\bar{x} > 0$, we have

$$U'(\bar{g}(x)) = \bar{h}(\bar{y})$$

where $\bar{h}$ is the optimal solution to $\inf_{h \in \mathcal{D}(\bar{y})} E_P[V(h)]$.

Proof of the abstract version of (iii): Existence of solution follows from Lemma 1 and Theorem 3 in Kramkov and Schachermayer [22]. Strict concavity of $U \Rightarrow E_P[U(\cdot)]$ is strictly concave $\Rightarrow E_P[U(\cdot)]$ has a unique maximum $\Rightarrow \bar{g}$ is unique.

Let $\bar{x} > 0$, $\bar{y} = u'(\bar{x})$ and $\bar{g}, \bar{h}$ be optimal solutions of $\sup_{g \in \mathcal{C}(\bar{x})} E_P[U(g)]$ and $\inf_{h \in \mathcal{D}(\bar{y})} E_P[V(h)]$ respectively. Then

$$\begin{aligned} 0 &\leq E_P[|V(\bar{h}(\bar{y})) + \bar{g}(x)\bar{h}(\bar{y}) - U(\bar{g}(x))|] \\ &= E_P[V(\bar{h}(\bar{y})) + \bar{g}(x)\bar{h}(\bar{y}) - U(\bar{g}(x))] \\ &\leq v(\bar{y}) + \bar{x}\bar{y} - u(\bar{x}) \\ &= 0. \end{aligned}$$

Here, the first equality follows from properties of $\bar{g}$ and $\bar{h}$. The second inequality follows from the definitions of u, v, that $\bar{g} \in \mathcal{C}(\bar{x}), \bar{h} \in \mathcal{D}(\bar{y})$ and $g \in \mathcal{C}(x)$ if and only if $E_P[gh] \leq xy$ for all $h \in \mathcal{D}(y)$. The last equality follows from the definition of v, regular maximization by differentiation and setting equal to 0, and that $\bar{y} = u'(\bar{x})$.

Hence

$$U(\bar{g}(\bar{x})) = V(\bar{h}(\bar{y})) + \bar{g}(\bar{x})\bar{h}(\bar{y}).$$

Denoting $\bar{g}(\bar{x})$ by x' and differentiating the expression above w.r.t. x' gives

$$U'(x') = \bar{h}(\bar{y}) \text{ a.s .}$$

So $U'(\bar{g}(\bar{x})) = \bar{h}(\bar{y})$ P-a.s.

This proves the abstract version of (iii), and hence (iii) is also proven.

□

5.8 Duality in a complete market ♢

The arguments in the previous section are fairly complicated. However, it turns out that for a complete market model, even for Ω arbitrary and continuous time, it is possible to derive a dual problem and quickly show that there is no duality gap using Lagrange duality, the Slater condition and the transformation of Section 5.2.

Consider a financial market based on a probability space $(\Omega, \mathcal{F}, P)$, where Ω is arbitrary. Assume, as usual, that there are N risky assets in the market, and one non-risky asset (bond). The composed price process of these assets is denoted S_t, where the time $t \in [0,T]$ (where the final time T may be infinite). Let $(\mathcal{F}_t)_t$ be the filtration generated by the price processes. Assume that the market is complete, so there is only one equivalent martingale measure (see Section 4.2).

From Section 5.2, it is reasonable to consider the following utility maximization problem

$$\begin{array}{ll} \max & E[U(Y)] \\ \text{subject to} & \\ & E_Q[Y] \leq x \end{array}$$

where Q is the single equivalent martingale measure, $x \in \mathbb{R}, x > 0$ is a constant, and the maximization is done over all $Y \in \mathcal{L}^0(\Omega, \mathcal{F}_T, P)$.

Let $\lambda \geq 0$ be the Lagrange dual variable. The Lagrange function is $L(Y, \lambda) = E[U(Y)] - \lambda(E_Q[Y] - x)$, so the Lagrange dual problem takes the form

$$\begin{aligned} \inf_{\lambda \geq 0} \sup_{Y \in \mathcal{L}^0(\mathcal{F}_T)} & \{E[U(Y)] - yE_Q[Y] + x\lambda\} \\ &= \inf_{\lambda \geq 0}\{\sup_Y \{E[U(Y) - \lambda \tfrac{dQ}{dP} Y]\} + x\lambda\} \\ &= \inf_{\lambda \geq 0}\{\sup_Y \{\textstyle\int (U(Y) - \lambda \tfrac{dQ}{dP} Y) dP\} + x\lambda\} \\ &= \inf_{\lambda \geq 0}\{\textstyle\int \sup_{Y(\omega)} (U(Y(\omega)) \\ &\qquad -\lambda \tfrac{dQ}{dP}(\omega) Y(\omega)) dP\} + x\lambda\} \\ &= \inf_{\lambda \geq 0}\{\textstyle\int V(\lambda \tfrac{dQ}{dP}) dP + x\lambda\} \\ &= \inf_{\lambda \geq 0}\{E[V(\lambda \tfrac{dQ}{dP})] + x\lambda\} \end{aligned}$$

where V is the KS-conjugate of the utility function U and $\frac{dQ}{dP}$ denotes the Radon-Nikodym derivative of Q with respect to P. In the calculations above the first equality follows from a change of measure and the third equality follows from the interchange rule between expectations and supremum of Theorem 14.60 in Rockafellar and Wets [36] ($h(x, \omega) = U(x) - y\frac{dQ}{dP}(\omega)x$ is a normal integrand from Definition 14.27 in [36]).

Now, to show that there is no duality gap, we check that the generalized Slater condition of Section 5.4 is satisfied. The objective function $E[U(\cdot)]$ is a concave function, the constraint $E_Q[\cdot] - x \leq 0$ is affine (hence convex) and $Y(\omega) = 0$ for all $\omega \in \Omega$ is a strictly feasible solution. Hence, the generalized Slater condition of section 5.4 implies that there is no duality gap, so

$$\sup\{E[U(Y)] : E_Q[Y] \leq x, X_T \in \mathcal{L}^0(\mathcal{F}_T)\} = \inf_{\lambda \geq 0}\{E[V(\frac{dQ}{dP}\lambda)] + x\lambda\}.$$

Actually, by solving $\sup_{Y(\omega)}\{U(Y(\omega)) - \lambda \frac{dQ}{dP}(\omega) Y(\omega)\}$, one finds that the optimal solution is $Y^*_\lambda(\omega) = I(\lambda \frac{dQ}{dP}(\omega))$ for each $\omega \in \Omega$ (by differentiating and setting equal to zero, where $I := (U')^{-1}$). This is allowed since U is differentiable, it is a maximum because U is concave, and since the maximization is done over all of $\mathbb{R}$ there is no boundary to check.

One must also find the optimal λ from $\inf_{\lambda \geq 0}\{E[V(\lambda \frac{dQ}{dP})] + x\lambda\}$. Note that

$$\inf_{\lambda \geq 0}\{E[V(\lambda \tfrac{dQ}{dP})] + x\lambda\} \quad = \inf_{\lambda \geq 0}\{v(\lambda) + x\lambda\}$$

where $v(\lambda) := E[V(\lambda \frac{dQ}{dP})]$. By differentiating ($v$ is differentiable) and setting equal to zero one sees that the optimal λ^* is determined as the solution of $v'(\lambda) = -x$.

Summarizing, the following theorem holds

Theorem 5.16 *Consider the primal optimization problem*

$$\sup\{E[U(Y)] : E_Q[Y] \leq x, Y \in \mathcal{L}^0(\mathcal{F}_T)\}$$

in the complete market setting above. The Lagrange dual problem is

$$\inf_{\lambda \geq 0}\{E[V(\frac{dQ}{dP}\lambda)] + x\lambda\}.$$

There is no duality gap, so

$$\sup\{E[U(Y)] : E_Q[Y] \leq x, Y \in \mathcal{L}^0(\mathcal{F}_T)\} = \inf_{\lambda \geq 0}\{E[V(\frac{dQ}{dP}\lambda)] + x\lambda\}.$$

Also, the optimal terminal portfolio value $Y^*(\omega) = I(\lambda \frac{dQ}{dP})$, where λ is determined from $v'(\lambda) = -x$.

Note how much simpler it is to prove the main result of Section 5.7 when the market is assumed to be complete, and Lagrange duality applies.

5.9 Utility maximization under risk constraints ♢

Consider a financial market where the scenario space Ω is finite. Assume that the market is arbitrage free and complete, so there is only one equivalent martingale measure $Q \in \mathcal{M}^e(S)$ (see comments after Theorem 4.2).

Consider the utility maximization of Section 5.5, but with a twist

$$\begin{array}{lll} \max_{Y \in \mathcal{L}^0(\Omega, \mathcal{F}_T, P)} & E_P[U(Y)] & \\ \text{subject to} & & \\ & E_Q[Y] & \leq x \text{ for all } Q \in \mathcal{M}^e(S) \\ & \rho(Y - x) & \leq c \end{array} \tag{5.18}$$

where $\rho(\cdot)$ is a convex risk measure, and c is some constant determined by the investor. One can assume that $c \geq 0$, since $c < 0$ corresponds to a negative risk, i.e. no chance of losing money, which any investor would be willing to accept.

This problem is actually equivalent to the problem

$$\begin{array}{lll} \max & E_P[U(X_T)] & \\ \text{subject to} & & \\ & \rho(X_T - x) & \leq c \end{array} \tag{5.19}$$

where the maximization is done over all admissible, predictable, self-financing trading strategies H such that $X_T = x + \int_0^T H_s dS_s$, (where the market is determined by the price process S). The interpretation of problem (5.19) is: Maximize the expected utility of terminal wealth over all admissible, predictable, self-financing trading strategies such that the risk of the investment is not greater than what the investor can tolerate.

These two problems are equivalent from arguments similar to those of Section 5.2. The following theorem is useful:

Theorem 5.17 $\diamondsuit$ *In the setting above, the following holds:*

(i) Problems (5.18) and (5.19) have the same optimal value.

(ii) If Y^* is an optimal solution of problem (5.18), then there exists an optimal solution H^* of problem (5.19) which replicates Y^*.

Proof: $\diamondsuit$

(*i*) To prove that the optimal value of problem (5.18) is less than or equal the optimal value of problem (5.19), assume that Y^* is an optimal solution of (5.18). Then there exists (from Proposition 2.10 in Schachermayer [40]) an $H \in \mathcal{H}$ such that $Y^* \leq x + \int_0^T H_s dS_s$. This implies that $c \geq \rho(Y^* - x) \geq \rho(\int_0^T H_s dS_s)$ (since Y^* must be feasible for problem (5.18)), hence H is feasible for problem (5.19). Also, since $E[U(\cdot)]$ is increasing $E[U(Y^*)] \leq E[U(x + \int_0^T H_s dS_s)]$, thus the optimal value of (5.18) is less than or equal the optimal value of (5.19).

Conversely, assume that H^* is an optimal solution of (5.19). Then, $X_T^{H^*} := x + \int_0^T H_S^* dS_s$, so $E_Q[X_T^{H^*}] = E_Q[x + \int_0^T H_S^* dS_s] = x \leq x$ for all $Q \in \mathcal{M}^e(S)$ (since S is a Q-martingale for all $Q \in \mathcal{M}^e(S)$). Also, since H^* must be feasible for problem (5.19), $\rho(X_T^{H^*} - x) \leq c$. Therefore, the optimal value of (5.19) must be less than or equal the optimal value of (5.18).

Therefore, the two optimal values coincide.

(*ii*) Now, assume that Y^* is an optimal solution of problem (5.18). Then there exists (from Proposition 2.10 in [40]) $H \in \mathcal{H}$ such that $Y^* \leq x + \int_0^T H_s dS_s$. Since Y^* is an optimal solution, $E[U(\cdot)]$ is an increasing function and the two problems have the same optimal value, this inequality must hold with equality. Hence there exists an $H \in \mathcal{H}$ which replicates Y^*, and this is therefore a feasible and optimal solution of problem (5.19).

□

Rewriting problem (5.18) by using Theorem 3.9

$$
\begin{array}{lrl}
\max_Y & E_P[U(Y)] & \\
\text{subject to} & & \\
& E_Q[Y] & \leq x \text{ for all } Q \in \mathcal{M}^e(S) \\
& \sup_{R \in \mathcal{P}}\{E_R[-Y+x] - \alpha(R)\} & \leq c
\end{array}
$$

where $\mathcal{P}$ is the set of all probability measures on $(\Omega, \mathcal{F})\}$ and $\alpha(\cdot) : \mathcal{P} \to \mathbb{R}$ is a convex and closed penalty function.

Ω is assumed to be finite, say $\Omega \subseteq \mathbb{R}^M$, so any $P \in \mathcal{P}$ can be uniquely identified with a $p \in \mathbb{R}^M$, by defining $p = (p_1, ..., p_M) = p(P) = (P(\omega_1), ..., P(\omega_M))$. This is a one-to-one correspondence. Let $\bar{\mathcal{P}} = \{p(P) : P \in \mathcal{P}\}$. This set is called the *standard simplex of* $\mathbb{R}^M$. This standard simplex is a polytope (can be proved similarly as Lemma 5.11), and hence it can be described as all the convex combinations of its finitely many extreme points, which will be denoted by $R_1, ..., R_S$. Note that the extreme points of the standard simplex are actually the coordinate vectors, so $R_1 = (1, 0, \ldots, 0), R_2 = (0, 1, 0, \ldots, 0), \ldots, R_S = R_M = (0, \ldots, 0, 1)$ (where all the vectors are in $\mathbb{R}^M$, and $S = M$ since there are M coordinate vectors of $\mathbb{R}^M$). This implies that $\sup_{R \in \mathcal{P}}\{E_R[-Y+x] - \alpha(R)\} \leq c$ if and only if $E_{R_i}[-Y+x] - \alpha(R_i) \leq c$ for $i = 1, ..., S$ (since $R \mapsto E_R[-Y+x]$ and $\alpha(\cdot)$ are convex functions).

Hence, problem (5.18) can be rewritten as

$$
\begin{array}{lrll}
\max_Y & E_P[U(Y)] & & \\
\text{subject to} & & & \\
& E_Q[Y] & \leq x & \text{for all } Q \in \mathcal{M}^e(S) \\
& E_{R_i}[-Y] - \alpha(R_i) & \leq c - x & \text{for } i = 1, ..., S.
\end{array}
$$

But from the same kind of arguments as in Section 5.6 this problem can be rewritten

$$
\begin{array}{lrll}
\max_Y & E_P[U(Y)] & & \\
\text{subject to} & & & \\
& E_{Q_k}[Y] & \leq x & \text{for } k = 1, ..., K \\
& E_{R_i}[-Y] - \alpha(R_i) & \leq c - x & \text{for } i = 1, ..., S.
\end{array}
\tag{5.20}
$$

where $Q_1, ..., Q_K$ are the extreme points of the polytope $\mathcal{M}^a(S)$.

This can be rewritten in component form as in Section 5.10, by defining $\epsilon_m := Y(\omega_m)$ for $m = 1, ..., M$, $\epsilon = (\epsilon_1, ..., \epsilon_M)$.

Let $y \in \mathbb{R}^K_+$ and $z \in \mathbb{R}^S_+$ be Lagrange dual variables. Then, the Lagrange function of problem (5.20) becomes

$$
L(\epsilon, y, z) = E_P[U(\epsilon)] - \sum_{k=1}^{K} y_k(E_{Q_k}[\epsilon] - x) - \sum_{i=1}^{S} z_i(E_{R_i}[-\epsilon] - \alpha(R_i) - (c - x)).
$$

Hence, the Lagrange dual problem is

$$\begin{array}{rcl}
\inf_{y,z\geq 0}\sup_{\epsilon} L(\epsilon,y,z) & = & \inf_{y,z\geq 0}\sup_{\epsilon}\{\sum_m (p_m U(\epsilon_m)- \\
& & \epsilon_m(\sum_k y_k q_m^{(k)}+\sum_i z_i r_m^{(i)})) + x\sum_k y_k \\
& & +(c-x)\sum_i z_i + \sum_i z_i\alpha(R_i)\} \\
& = & \inf_{y,z\geq 0}\{\sum_m p_m \sup_{\epsilon_m}\{U(\epsilon_m)- \\
& & \frac{\epsilon_m}{p_m}(\sum_k y_k q_m^{(k)}+\sum_i z_i r_m^{(i)})\} \\
& & +x\sum_k y_k + (c-x)\sum_i z_i + \sum_i z_i\alpha(R_i)\} \\
& = & \inf_{y,z\geq 0}\{\sum_m p_m V(\sum_k y_k \frac{dQ_k}{dP}+\sum_i z_i\frac{dR_i}{dP}) \\
& & +x\sum_k y_k + (c-x)\sum_i z_i + \sum_i z_i\alpha(R_i)\}
\end{array} \tag{5.21}$$

where the Radon-Nikodym derivatives must exist since Ω is assumed to be finite, so one can assume that $P>0$, and then the Radon-Nikodym derivative of any measure R w.r.t. P is just $\frac{dR}{dP}(\omega)=\frac{r_m}{p_m}$ where $r_m=R(\omega_m)$ for $m=1,...,M$. Note that for the measures $R_1,\dots,R_M$, corresponding to the extreme points of the standard simplex in $\mathbb{R}^M$, this implies that

$$\frac{dR_i}{dP}(\omega)=\begin{cases}\frac{1}{P(\omega_i)} & \text{for } \omega=\omega_i, \\ 0 & \text{otherwise.}\end{cases}$$

Since V (the KS-conjugate of U) is differentiable, the optimal ϵ_m^*'s in V are found by differentiating and setting equal to 0. Hence

$$\epsilon_m^* = I(\sum_k y_k\frac{dQ_k}{dP}+\sum_i z_i\frac{dR_i}{dP})$$

where $I:=(U')^{-1}$, as previously. Note that this is in fact a maximum from the properties of V, and that since the maximization is done over all of $\mathbb{R}$, there are no boundaries to check.

The Slater condition of Section 5.4 can again be used to show that there is no duality gap: Problem (5.20) is a concave maximization problem with linear constraints and $\epsilon_m := x-\frac{c}{2}$ for $m=1,...,M$ is a (strictly) feasible solution (in $\mathrm{rint}(\mathbb{R}^m)$) because $E_{Q_k}[\epsilon]=x-\frac{c}{2}<x$ for all k (since $c>0$), and $\rho(\epsilon-x)=\rho(0+\mathbf{1}(x-\frac{c}{2}-x))=\rho(0)-(-\frac{c}{2})=0+\frac{c}{2}<c$ (from Definition 3.1). Hence, the Slater condition implies that there is no duality gap, so problems (5.19) and (5.21) share the same optimal value.

Theorem 5.18 *(Utility maximization under risk constraints) Problem (5.19) has optimal solution*

$$\begin{array}{rcl}
X_T^{H^*}(\omega_m) & = & I(\sum_k y_k\frac{dQ_k}{dP}(\omega_m)+\sum_i z_i\frac{dR_i}{dP}(\omega_m)) \\
& = & I(\sum_k y_k\frac{dQ_k}{dP}(\omega_m)+z_m\frac{1}{P(\omega_m)})
\end{array}$$

where $I:=(U')^{-1}$, Q_k $(k=1,\dots,K)$ are the extreme points of the set of absolutely continuous martingale measures $\mathcal{M}^a(S)$, and we have used the form of R_i, $i=1,\dots,S$ (recall $S=M$), the probability measures corresponding to the extreme points of the standard simplex of $\mathbb{R}^M$.

Proof: The theorem follows from the derivation above. □

5.10 Final comments on utility maximization ♢

This chapter has shown how duality theory can be applied to solve utility maximization problems. The following is a summary of the techniques that have been used:

- For a complete market or finite Ω:
  - Consider the problem $\max\{E[U(X_T)]$: there exists a self-financing admissible trading strategy H such that $X_T = x + \int_0^T H_s dS_s\}$.
  - Rewrite the problem as in Section 5.2. The rewritten problem is a constrained concave maximization problem which only depends on the final portfolio value. This is called *Phase 1.*
  - Solve the rewritten problem using Lagrange duality, and use the Slater condition to show that there is no duality gap (as in e.g. Section 5.5). This is called *Phase 2.*
- For an incomplete market where Ω is arbitrary:
  - Skip Phase 1, and work with the original problem.
  - Try to find a dual problem and a space for the dual variables to vary in such that it is possible to apply the MiniMax theorem.

The problem with the first method, including Phase 1, is that it is difficult to generalize to the situation where there are, for example, portfolio constraints. The Phase 1 rewriting is clever, but it removes the portfolio process from the problem. Hence, the only kind of additional constraints one can add to this method are constraints on the terminal value. However, the Phase 1- Phase 2 method also works when the investor gets utility not only from terminal wealth, but also from consumption (as long as the market is complete). The objective function to be maximized is then of the form $E[\int_0^T U_1(c(u), u)du + U_2(X_T)]$, where U_1, U_2 are utility functions, c is a consumption process and X_T is the terminal portfolio value. This is illustrated in Karatzas and Shreve [17] (though without close explanation of the Lagrange duality principle, and no use of the Slater condition).

Karatzas and Shreve [17], also illustrate utility maximization in incomplete markets. For a utility maximization problem where the portfolio process is constrained to a closed, convex set $K \subseteq \mathbb{R}^N$, such that $0 \in K$, they introduce a family of new markets that are translations of the original market. In each new market, they solve the unconstrained utility maximization problem. A proposition is derived, which states that if one can find a translated market satisfying certain equations, then the solution of the utility maximization problem in the translated market, gives the solution to the problem in the original market as well, (see Proposition 6.3.8 in [17]). The weakness of this proposition is that it does not give any existence conditions for such a market. Therefore, the proposition may not provide much information.

The purpose of this chapter has been to show how duality methods can be used to solve utility maximization problems. Results from convexity- and duality theory, such as Lagrange duality, the Slater condition, and the Mini-Max theorem are combined with stochastic, real, and functional analysis to find dual problems, and show the absence of a duality gap. The next chapter will proceed along these lines, but will consider pricing problems instead of utility maximization problems.

Chapter 6

Pricing of claims via duality

The purpose of this chapter is to show (to my knowledge) new applications of duality to pricing problems in mathematical finance. Section 6.1 investigates (for a finite scenario space, discrete time model) the pricing problem of the seller of a contingent claim who has a general level of inside information. Lagrange duality and the linear programming duality theorem is used to derive a characterization of the seller's price of the claim (see Theorem 6.3). This result is applied to show that a seller with more information can offer the claim at a lower price than a seller with less information, see Theorem 6.5. The buyer's problem is analogous to the seller's problem, so this implies that there is a smaller probability of a seller and buyer agreeing on a price in a market with partial information than in a market with full information.

Section 6.2 generalizes the results of Section 6.1 to the case where the scenario space Ω is arbitrary.

Section 6.3 considers the seller's pricing problem with short-selling constraints in a financial model where the scenario space is finite and the time is discrete. Lagrange duality and the Slater condition are used to characterize the seller's price of a claim under short selling constraints, see Theorem 6.14. Section 6.4 generalizes the results of Section 6.3 to a model where the scenario space is arbitrary using conjugate duality.

Section 6.5 also considers the pricing problem of a seller, but under the constraint $H_{n^*}(t, \omega) \in [A, C]$ for all t, ω (where $0 \in [A, C] \subseteq \mathbb{R}$). This section works with a finite scenario space and discrete time, and a characterization of the seller's price is derived via Lagrange duality.

This chapter emphasizes the close connection between fundamental results of mathematical finance and the conjugate duality theory introduced by Rockafellar [34] (see Chapter 2). In the literature, there seems to be quite few papers considering this connection. Exceptions are the work of Henclova [15], King and Korf [19] and Pennanen [29].

A comment on notation: In this chapter, terms such as *full information* and *inside information* will be used. Full information means that the agent knows the prices of all the assets in the market at any given time. Inside information means that the agent knows something more than this. Though knowing the prices of the assets is referred to as having full information, this is also in a sense a minimal level of information, at least when working with a discrete time model. Why is this a minimal information? Assume that an agent does not know the actual prices of the assets in the market at some time t, she only knows the prices of the assets at, say, time $t-1$. Then, the agent will make a decision of how to trade in the market, based on incorrect information. However, when she tries to execute this trading strategy, she will be informed of the actual prices in the market, and may not wish to (or afford to) follow her plan. She will then make a new decision based on her updated information, and hence, the fact that she initially did not know the prices does not matter (where it is assumed that trading and decision-making happens instantaneously).

6.1 The pricing problem with inside information: Finite Ω $\diamondsuit$

The arguments of this section are inspired by King [18] (see Section 4.3).

This section considers the pricing problem of the seller of a contingent claim under a general filtration modeling full- or inside information.

Consider a financial market based on a probability space $(\Omega, \mathcal{F}, P)$ where the scenario space Ω is finite. There are N risky assets in the market with price processes $S_n(t)$, $n = 1, \ldots, N$ and one bond with price process $S_0(t)$. The time $t \in \{0, 1, \ldots, T\}$ where $T < \infty$. Let S denote the composed price process, and let $(\mathcal{F}_t)_t$ be the filtration generated by the price process. Hence, the composed price process S is adapted to the filtration $(\mathcal{F}_t)_t$. The market can be modeled by a scenario tree as in Section 4.3.

Recall that a contingent claim is a nonnegative, $\mathcal{F}_T$-measurable random variable on the probability space $(\Omega, \mathcal{F}_T, P)$. Consider a contingent claim B in the market. Consider also a seller of this claim with a *general* information modeling filtration $(\mathcal{G}_t)_{t=0}^T$ such that $\mathcal{G}_0 = \{\emptyset, \Omega\}$ and $\mathcal{G}_T$ is the algebra corresponding to the partition $\{\{\omega_1\}, ..., \{\omega_M\}\}$ (sometimes the filtration is denoted by $(\mathcal{G}_t)_t$ to simplify notation). We assume that the price process S is adapted to $(\mathcal{G}_t)_t$, in order for the seller to know the price of each asset at any given time. Hence the filtration $(\mathcal{F}_t)_t$ is nested in $(\mathcal{G}_t)_t$.

It will turn out that the set $\mathcal{M}^a(S, \mathcal{G})$ is important: $Q \in \mathcal{M}^a(S, \mathcal{G})$ if

- Q is a probability measure on $(\Omega, \mathcal{G}_T)$ (recall that $\mathcal{F}_T = \mathcal{G}_T$ by assumption).
- Q is absolutely continuous with respect to P.
- The price process S is a Q-martingale w.r.t. the filtration $(\mathcal{G}_t)_t$.

So $\mathcal{M}^a(S,\mathcal{G})$ is the set of absolutely continuous martingale measures w.r.t. $(\mathcal{G}_t)_t$. This is (to my knowledge) a new term.

The pricing problem of this seller is

$$\begin{array}{rrcll} \min & & v & & \\ \text{subject to} & & & & \\ & S_0 \cdot H_0 & \leq & v & \\ & B_k & \leq & S_k \cdot H_{a(k)} & \text{for all } k \in \mathcal{N}_T^{\mathcal{G}}, \\ & S_k \cdot H_k & = & S_k \cdot H_{a(k)} & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}},\ 1 \leq t \leq T-1, \end{array} \tag{6.1}$$

where the minimization is done with respect to $v \in \mathbb{R}$ and $H_k \in \mathbb{R}^{N+1}$ for $k \in \mathcal{N}_t^{\mathcal{G}}$, $t = 0, ..., T-1$. Here, $\mathcal{N}_t^{\mathcal{G}}$ denotes the set of time t-vertices (nodes) in the scenario tree representing the filtration $(\mathcal{G}_t)_t$, and B_k denotes the value of the claim B in the node $k \in \mathcal{N}_T$. Recall that $a(k)$ denotes the ancestor of vertex k, see Section 4.3.

Hence, the seller's problem is: Minimize the price v of the claim B such that the seller is able to pay B at time T from investments in a self-financing, adapted portfolio that costs less than or equal to v at time 0. Note that the portfolio process H has been translated, so that H is adapted to $(\mathcal{G}_t)_t$, not predictable (which is assumed in many papers in mathematical finance). This is just a simplification of notation.

Problem (6.1) is actually a linear programming (LP) problem, and one can find the dual of this problem using standard LP-duality techniques. However, it turns out to be easier to find the dual problem via Lagrange duality techniques (see Section 5.4). The LP-dual problem is a special case of the Lagrange dual problem, so LP-duality theory implies that there is no duality gap (this can also be shown via the Slater condition, since choosing $v := 1 + \sup_{\omega \in \Omega} B(\omega)$ and putting everything in the bank account is a strictly feasible solution of problem (6.1)). Note also that since problem (6.1) is a linear programming problem, the simplex algorithm is an efficient method for computing optimal prices and optimal trading strategies for specific examples.

Problem (6.1) is equivalent to

$$\begin{array}{rrcll} \min & & v & & \\ \text{subject to} & & & & \\ & S_0 \cdot H_0 - v & \leq & 0 & \\ & B_k - S_k \cdot H_{a(k)} & \leq & 0 & \text{for all } k \in \mathcal{N}_T^{\mathcal{G}} \\ & S_k \cdot (H_k - H_{a(k)}) & \leq & 0 & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}},\ 1 \leq t \leq T-1, \\ & -S_k \cdot (H_k - H_{a(k)}) & \leq & 0 & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}},\ 1 \leq t \leq T-1, \end{array} \tag{6.2}$$

which is of a form suitable for the Lagrange duality method.

Let $y_0 \geq 0$, $z_k \geq 0$ for all $k \in \mathcal{N}_T^{\mathcal{G}}$ and $y_k^1, y_k^2 \geq 0$ for all $k \in \mathcal{N}_t^{\mathcal{G}}$, $t \in \{1, \ldots, T-1\}$ be the Lagrange multipliers. Then, the Lagrange dual problem is

$$
\begin{aligned}
\sup_{y_0,z,y^1,y^2\geq 0}\inf_{v,H} \quad & \{v + y_0(S_0\cdot H_0 - v) + \textstyle\sum_{k\in\mathcal{N}_T^{\mathcal{G}}} z_k(B_k - S_k\cdot H_{a(k)}) \\
& + \textstyle\sum_{t=1}^{T-1}\sum_{k\in\mathcal{N}_t^{\mathcal{G}}}(y_k^1 - y_k^2)S_k(H_k - H_{a(k)})\} \\
= \sup_{y_0,z\geq 0,y}\{\inf_v \quad & \{v(1-y_0)\} + \inf_{H_0}\{y_0 S_0\cdot H_0 - \textstyle\sum_{m\in\mathcal{C}_{\mathcal{G}}(0)} y_m S_m\cdot H_0\} \\
& + \textstyle\sum_{t=1}^{T-2}\sum_{k\in\mathcal{N}_t^{\mathcal{G}}}\inf_{H_k}\{y_k S_k\cdot H_k - \sum_{m\in\mathcal{C}_{\mathcal{G}}(k)} y_m S_m\cdot H_k\} \\
& + \textstyle\sum_{k\in\mathcal{N}_{T-1}^{\mathcal{G}}}\inf_{H_k}\{y_k S_k\cdot H_k - \sum_{m\in\mathcal{C}_{\mathcal{G}}(k)} z_m S_m\cdot H_k\} \\
& + \textstyle\sum_{k\in\mathcal{N}_T^{\mathcal{G}}} z_k B_k\}
\end{aligned}
$$

where we have defined $y_k := y_k^1 - y_k^2$ for all $k\in\mathcal{N}_t^{\mathcal{G}}$, $t\in\{0,1,\ldots,T-1\}$. y (the vector of the y_k's) is a free variable (i.e. the sign of the components of y is not clear a priori).

Consider each of the minimization problems separately. In order to have a feasible dual solution, all of these minimization problems must have optimal value greater than $-\infty$.

- $\inf_v\{v(1-y_0)\} > -\infty$ (that is, there is a feasible dual solution) if and only if $y_0 = 1$.
- $\inf_{H_0}\{y_0 S_0\cdot H_0 - \sum_{m\in\mathcal{C}_{\mathcal{G}}(0)} y_m S_m\cdot H_0\} > -\infty$ if and only if $y_0 S_0 = \sum_{m\in\mathcal{C}_{\mathcal{G}}(0)} y_m S_m$.
- $\inf_{H_k}\{y_k S_k\cdot H_k - \sum_{m\in\mathcal{C}_{\mathcal{G}}(k)} y_m S_m\cdot H_k\} > -\infty$ if and only if $y_k S_k = \sum_{m\in\mathcal{C}_{\mathcal{G}}(k)} y_m S_m$. Therefore, in order to get a dual solution, this must hold for all $k\in\mathcal{N}_t^{\mathcal{G}}$, for $t = 1,\ldots,T-2$.
- Finally, $\inf_{H_k}\{y_k S_k\cdot H_k - \sum_{m\in\mathcal{C}_{\mathcal{G}}(k)} z_m S_m\cdot H_k\} > -\infty$ if and only if $y_k S_k = \sum_{m\in\mathcal{C}_{\mathcal{G}}(k)} z_m S_m$. In order to get a feasible dual solution this must hold for all $k\in\mathcal{N}_{T-1}^{\mathcal{G}}$.

Hence, the dual problem is

$$
\begin{array}{ll}
\sup_{y_0,z\geq 0,y} & \sum_{k\in\mathcal{N}_T^{\mathcal{G}}} z_k B_k \\
\text{subject to} &
\end{array}
$$

$$
\begin{aligned}
y_0 &= 1 \\
y_k S_k &= \textstyle\sum_{m\in\mathcal{C}_{\mathcal{G}}(k)} y_m S_m \quad \text{for all } k\in\mathcal{N}_t^{\mathcal{G}}, \\
& \qquad\qquad\qquad\qquad\quad t = 0,1,\ldots,T-2, \\
y_k S_k &= \textstyle\sum_{m\in\mathcal{C}_{\mathcal{G}}(k)} z_m S_m \quad \text{for all } k\in\mathcal{N}_{T-1}^{\mathcal{G}}.
\end{aligned}
$$

By considering the second equation above for the bond, i.e. for S_k^0, one sees that in order to have a feasible dual solution

$$y_0 S_0^0 = \sum_{m\in\mathcal{C}_{\mathcal{G}}(0)} y_m S_m^0$$

must hold. Since the market is assumed to be normalized, so $S_k^0 = 1$ for all k,

$$1 = y_0 = \sum_{m \in \mathcal{C}_{\mathcal{G}}(0)} y_m.$$

Also, from the second equation above, the same type of argument implies that

$$y_k = \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m$$

for all $k \in \mathcal{N}_t^{\mathcal{G}}$, $t = 1, ..., T-2$. Hence $\sum_{k \in \mathcal{N}_{T-1}^{\mathcal{G}}} y_k = y_0 = 1$ (by induction). From the last dual feasibility equation (considered for $n = 0$), $\sum_{k \in \mathcal{N}_T^{\mathcal{G}}} z_k = \sum_{k \in \mathcal{N}_{T-1}^{\mathcal{G}}} y_k = 1$ and $z_k \geq 0$ for all k (since z is a Lagrange multiplier). Therefore, $\{z_k\}_{k \in \mathcal{N}_T^{\mathcal{G}}}$ can be identified with a probability measure Q (on the terminal vertices of the scenario tree) such that the Q-probability of ending up in terminal vertex k is z_k. Then, the condition $y_k S_k = \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m S_m$ for all $k \in \mathcal{N}_t^{\mathcal{G}}$, is a martingale condition of the form $S(t-1) = E[S_t | \mathcal{G}_{t-1}]$ (from the definition of conditional expectation), which can be shown by induction to imply the general martingale condition in this discrete time case. This proves that any feasible dual solution is in $\mathcal{M}^a(S, \mathcal{G})$. The converse also holds: Take $Q \in \mathcal{M}^a(S, \mathcal{G})$, and define $z_m := Q(\omega_m)$ for $m = 1, ..., M$, $y_k := \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} z_m$ for $k \in \mathcal{N}_{T-1}^{\mathcal{G}}$ and $y_k := \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m$ for $k \in \mathcal{N}_t^{\mathcal{G}}$, $0 \leq t \leq T-2$. It can be checked (from these definitions) that this is a feasible dual solution.

Hence, the Lagrange dual problem can be rewritten

$$\sup_{Q \in \mathcal{M}^a(S, \mathcal{G})} E_Q[B]$$

where $\mathcal{M}^a(S, \mathcal{G})$ denotes the set of martingale measures with respect to $(\mathcal{G}_t)_t$ which are absolutely continuous with respect to the original measure P.

As explained previously, LP-duality (or the Slater condition) implies that there is no duality gap. Hence, the optimal primal value, i.e. the seller's price of the contingent claim B, is equal to the optimal dual value, that is $\sup_{Q \in \mathcal{M}^a(S, \mathcal{G})} E_Q[B]$.

From general pricing theory (see Karatzas and Shreve [17], Theorem 6.2), a seller who has the filtration $(\mathcal{F}_t)_t$ (the original filtration) will offer B at the price $\sup_{Q \in \mathcal{M}^e(S, \mathcal{F})} E_Q[B]$ (in a normalized market). Here, $\mathcal{M}^e(S, \mathcal{F})$ denotes the set of all equivalent martingale measures w.r.t. the filtration $(\mathcal{F}_t)_t$. This appears to be a different price than the one derived above. However, it turns out that

$$\sup_{Q \in \mathcal{M}^a(S, \mathcal{F})} E_Q[B] = \sup_{Q \in \mathcal{M}^e(S, \mathcal{F})} E_Q[B].$$

Why is this?

The following lemma is a slight adaptation of Proposition 2.10 in Schachermayer [40].

Lemma 6.1 *Assume there exists $Q^* \in \mathcal{M}^e(S, \mathcal{G})$. For any $g \in \mathcal{L}^\infty(\Omega, \mathcal{F}, P)$,*

$$E_Q[g] \leq 0 \text{ for all } Q \in \mathcal{M}^a(S, \mathcal{G}).$$

if and only if

$$E_Q[g] \leq 0 \text{ for all } Q \in \mathcal{M}^e(S, \mathcal{G}).$$

Proof:

By assumption, there exists at least one $Q^* \in \mathcal{M}^e(S, \mathcal{G})$ (and hence also a $Q \in \mathcal{M}^a(S, \mathcal{G})$, since $\mathcal{M}^e(S, \mathcal{G}) \subseteq \mathcal{M}^a(S, \mathcal{G})$). For any $Q \in \mathcal{M}^a(S, \mathcal{G})$ and $\lambda \in (0, 1)$, $\lambda Q^* + (1 - \lambda)Q \in \mathcal{M}^e(S, \mathcal{G})$. Hence, $\mathcal{M}^e(S, \mathcal{G})$ is dense in $\mathcal{M}^a(S, \mathcal{G})$, and the lemma follows. □

Lemma 6.2 ♢

Assume there exists a $Q \in \mathcal{M}^e(S, \mathcal{G})$. Then

$$\sup_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B] = \sup_{Q \in \mathcal{M}^e(S,\mathcal{G})} E_Q[B].$$

Proof: ♢

- To prove that $\sup_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B] \leq \sup_{Q \in \mathcal{M}^e(S,\mathcal{G})} E_Q[B]$: Define $x := \sup_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B]$. Then $E_Q[B] \leq x$ for all $Q \in \mathcal{M}^a(S, \mathcal{G})$. But from Lemma 6.11 (with $g = B - x$), this implies that $E_Q[B] \leq x$ for all $Q \in \mathcal{M}^e(S, \mathcal{G})$, so $\sup_{Q \in \mathcal{M}^e(S,\mathcal{G})} E_Q[B] \leq x$, and hence the inequality follows.
- The opposite inequality follows from that $\mathcal{M}^e(S, \mathcal{G}) \subseteq \mathcal{M}^a(S, \mathcal{G})$.

□

Hence, in particular

$$\sup_{Q \in \mathcal{M}^a(S,\mathcal{F})} E_Q[B] = \sup_{Q \in \mathcal{M}^e(S,\mathcal{F})} E_Q[B]. \tag{6.3}$$

if we assume that there is no arbitrage with respect to $(\mathcal{F}_t)_t$, so there exists $Q \in \mathcal{M}^e(S, \mathcal{F}) = \mathcal{M}^e(S)$.

Note that the arguments above would go through in the same way for the buyer's problem, so for a general filtration $(\mathcal{G}_t)_t$, the buyer's price of the claim B is $\inf_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B]$. Note also that by assuming that both seller and buyer in the market have the same full information filtration $(\mathcal{F}_t)_t$, and that the market is complete, we know that there is only one equivalent martingale measure (see Øksendal [27], chapter 12), and hence the buyer and seller will agree on $E_Q[B]$ (where Q is the single equivalent martingale measure) as the price of B.

Theorem 6.3 $\diamondsuit$

Consider a normalized financial market based on a finite scenario space. The seller of a contingent claim B, who has filtration $(\mathcal{G}_t)_t$ such that $\mathcal{F}_t \subseteq \mathcal{G}_t$ for all times $t \in \{0, 1, \ldots, T\}$ will sell B at the price

$$\sup_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B]$$

where $\mathcal{M}^a(S, \mathcal{G})$ denotes the set of probability measures (hence, absolutely continuous w.r.t. P) that turn the price process S into a martingale w.r.t. the filtration $(\mathcal{G}_t)_t$.

If there exists $Q \in \mathcal{M}^e(S, \mathcal{G})$,

$$\sup_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B] = \sup_{Q \in \mathcal{M}^e(S,\mathcal{G})} E_Q[B]$$

Hence, in particular, if we consider a seller with full information $(\mathcal{F}_t)_t$ in a market where there is no arbitrage (w.r.t. $(\mathcal{F}_t)_t$), then the price offered by the seller is $\sup_{Q \in \mathcal{M}^e(S,\mathcal{F})} E_Q[B]$.

Now, we will consider pricing with different levels of information. We now introduce two sellers of the claim B into the model. The first seller, seller A_1, has information corresponding to a filtration $(\mathcal{H}_t)_{t=0}^T$ (where $\mathcal{H}_0 = \{\emptyset, \Omega\}$ and $\mathcal{H}_T$ is the algebra corresponding to the partition $\{\{\omega_1\}, ..., \{\omega_M\}\}$). The price process S in the market is adapted to this filtration (in order for the seller to know the prices of each asset at any given time). Hence $(\mathcal{F}_t)_t$ is nested in $(\mathcal{H}_t)_t$. The second seller, seller A_2, has more information. His information is modeled by a filtration $(\mathcal{J}_t)_{t=0}^T$ (where $\mathcal{J}_0 = \{\emptyset, \Omega\}$ and $\mathcal{J}_T$ corresponds to $\{\{\omega_1\}, ..., \{\omega_M\}\}$), and this filtration is such that $\mathcal{H}_t \subseteq \mathcal{J}_t$ for all $t = 0, 1, ..., T$.

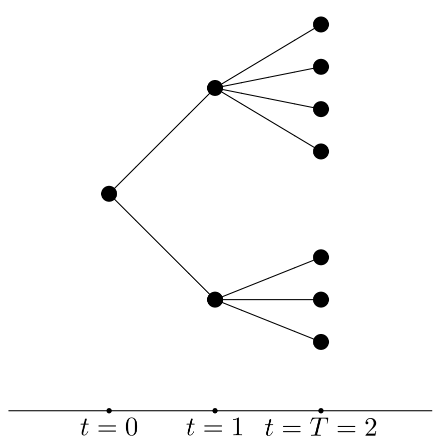

Figure 6.1: Possible scenario tree for seller A_1.

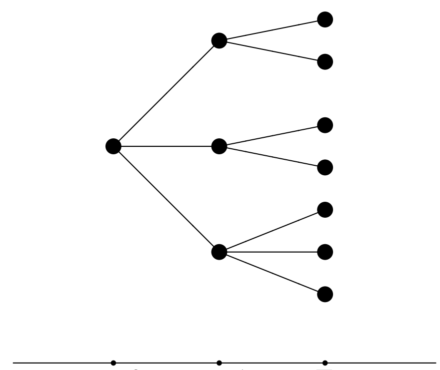

Figure 6.2: Possible scenario tree for seller A_2.

Possible scenario trees for seller A_1 and seller A_2 when $T=2$ are illustrated in Figures 6.1 and 6.2.

From Theorem 6.3, we know that seller A_1, with less information, will offer B at the price $\sup_{Q \in \mathcal{M}^a(S,\mathcal{H})} E_Q[B]$, while seller A_2, with more information can sell B for $\sup_{Q \in \mathcal{M}^a(S,\mathcal{J})} E_Q[B]$. What is the relationship between these two prices? (Recall that $\mathcal{M}^a(S,\mathcal{G})$ denotes the set of absolutely continuous probability measures Q such that S is a Q-martingale w.r.t. the filtration $(\mathcal{G}_t)$.)

Lemma 6.4 $\diamondsuit$
$\mathcal{M}^a(S,\mathcal{J}) \subseteq \mathcal{M}^a(S,\mathcal{H})$

Proof: $\diamondsuit$
Assume that $Q \in \mathcal{M}^a(S,\mathcal{J})$. Then, $E_Q[S_t|\mathcal{J}_s] = S_s$ for all $s < t$ and Q is a probability measure. But then

$$\begin{aligned} E_Q[S_t|\mathcal{H}_s] &= E_Q[E_Q[S_t|\mathcal{J}_s]|\mathcal{H}_s] \\ &= E_Q[S_s|\mathcal{H}_s] \\ &= S_s \end{aligned}$$

where the first equality follows from the rule of double expectation (since $\mathcal{H}_s \subseteq \mathcal{J}_s$) and the final equality follows from S_s being $\mathcal{H}_s$-measurable. Hence, S is a Q-martingale with respect to $(\mathcal{H}_t)_t$, so $Q \in \mathcal{M}^a(S,\mathcal{H})$. $\square$

Lemma 6.4 is illustrated in Figure 6.3.

The next theorem follows directly from Lemma 6.4.

Theorem 6.5 $\diamondsuit$
In the setting above

$$\sup_{Q \in \mathcal{M}^a(S,\mathcal{J})} E_Q[B] \leq \sup_{Q \in \mathcal{M}^a(S,\mathcal{H})} E_Q[B]$$

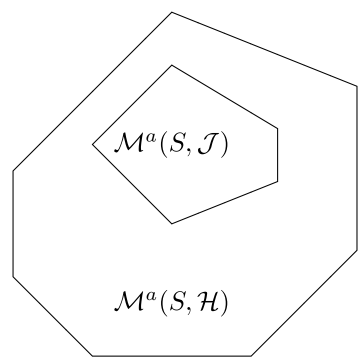

Figure 6.3: The polytopes $\mathcal{M}^a(S, \mathcal{H})$ and $\mathcal{M}^a(S, \mathcal{J})$.

Hence, a seller of claim B with less information will offer B at a price greater than or equal the price of a seller with more information.

Proof: $\diamondsuit$
This follows from Lemma 6.4 and that maximizing over a larger set must give a greater (or equal) value. $\square$

The result of Lemma 6.5 has a natural economic interpretation: The well informed seller can offer a better price for the claim than the less informed seller, because it is a smaller risk to offer a low price for a seller who knows a lot about the development of the market than it is for a seller with less information.

This type of argument can be done from a buyer's point of view as well. This leads to the same kind of result: $\inf_{Q \in \mathcal{M}^a(S,\mathcal{H})} E_Q[B] \leq \inf_{Q \in \mathcal{M}^a(S,\mathcal{J})} E_Q[B]$. That is: The well informed buyer is willing to pay more for the claim since he knows the market development better. Hence, there is a larger probability of the seller and buyer agreeing on a price in a market with more information than in one with less information.

From Section 5.6, the following holds for $\mathcal{M}^a(S, \mathcal{G})$ (for a general filtration $(\mathcal{G}_t)_t$):

- $\mathcal{M}^a(S, \mathcal{G})$ is a polytope, hence it is compact, convex and described by its finitely many extreme points.
- $\mathcal{M}^a(S, \mathcal{G})$ is contained in the standard simplex of $\mathbb{R}^M$ (the set of all probability measures on an M-dimensional scenario space).

Since $\mathcal{M}^a(S, \mathcal{G})$ is described by its finitely many extreme points, it follows that $\sup_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B]$ is actually attained in one of these extreme points denoted $Q_1^{\mathcal{G}}, ..., Q_L^{\mathcal{G}}$ of $\mathcal{M}^a(S, \mathcal{G})$ (since the function $Q \mapsto E_Q[B]$ is linear). Hence

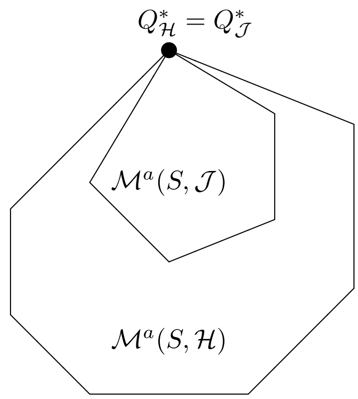

Figure 6.4: Possible optimal solution.

$$\max_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B] = \max_{Q \in \{Q_1^{\mathcal{G}}, \ldots, Q_L^{\mathcal{G}}\}} E_Q[B] = E_{Q^*_{\mathcal{G}}}[B].$$

Returning to the two sellers with different information. How large is the gap between the the prices offered by these sellers? That is, what is

$$\max_{\mathcal{M}^a(S,\mathcal{H})} E_Q[B] - \max_{\mathcal{M}^a(S,\mathcal{J})} E_Q[B]?$$

(note that the sup has been changed to a max since the optimum is actually attained) The price of the seller with less information, seller A_1, is $\max_{Q \in \{Q_1^{\mathcal{H}}, \ldots, Q_{L'}^{\mathcal{H}}\}} = E_{Q^*_{\mathcal{H}}}[B]$. Hence, from Lemma 6.5, if $Q^*_{\mathcal{H}} \in \mathcal{M}^a(S,\mathcal{J})$ then $\max_{Q \in \mathcal{M}^a(S,\mathcal{J})} = E_{Q^*_{\mathcal{H}}}[B]$, and hence the two sellers offer B at the same price. (Note that if $Q_l^{\mathcal{H}} \in \{Q_1^{\mathcal{H}}, \ldots, Q_{L'}^{\mathcal{H}}\}$ is an extreme point for $\mathcal{M}^a(S,\mathcal{H})$ and $Q_l^{\mathcal{H}} \in \mathcal{M}^a(S,\mathcal{J})$, then $Q_l^{\mathcal{H}}$ is an extreme point for $\mathcal{M}^a(S,\mathcal{J})$ as well).

Note that the optimal measure $Q^* \in \mathcal{M}^a(S)$ (with respect to either of the filtrations $(\mathcal{H}_t)_t$ or $(\mathcal{J}_t)_t$) may not be unique (depending on the structure of $\mathcal{M}^a(S)$ and the claim B). If there are two extreme points ("corners" of the polytope $\mathcal{M}^a(S)$) that achieve the same optimal value, all the measures represented by the points on the line connecting these two corners must also attain this optimal value, see Figure 6.5.

Regardless of whether the optimums are unique or not, the difference

$$\max_{\mathcal{M}^a(S,\mathcal{H})} E_Q[B] - \max_{\mathcal{M}^a(S,\mathcal{J})} E_Q[B]$$

will be equal to the (Euclidean) norm of B times the distance between the hyperplanes through $Q^*_{\mathcal{H}}$ and $Q^*_{\mathcal{J}}$ (respectively) with normal vector B.

We will now look at an example illustrating Theorem 6.5.

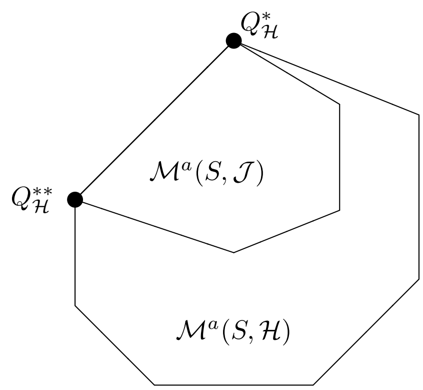

Figure 6.5: Potential optimal solutions.

Consider a market as in Section 4.3, where $N = 1$, so there is only one risky asset, plus a bond. The terminal time $T = 2$, and sellers A_1 and A_2 have information structures as in Figures 6.1 and 6.2, where (as previously) the filtration of seller A_1 is denoted by $(\mathcal{H}_t)_t$, and the information of seller A_2 is denoted $(\mathcal{J}_t)_t$, where $(\mathcal{H}_t)_t$ is nested in $(\mathcal{J}_t)_t$. Assume that $(\mathcal{H}_t)_t = (\mathcal{F}_t)_t$ (the original, full information, filtration). The market is assumed to be normalized, so $S^0_k = 1$ for all $k \in \mathcal{N}^{\mathcal{H}}$ and all $k \in \mathcal{N}^{\mathcal{J}}$ (i.e. for all the vertices in the scenario tree of both the agents). The price process of S^1 (the risky asset) is illustrated in Figures 6.6 and 6.7.

In Figures 6.6 and 6.7, $S^1_k(t)$ denotes the price of the risky asset at node number $k \in \mathcal{N}_t$ at time t.

Consider the contingent claim B given by $(B_1, B_2, B_3, B_4, B_5, B_6, B_7) = (2, 3, 7, 1, 3, 5, 1)$, where B_k is the value of B in node $k \in \mathcal{N}_2$ of the terminal vertices. B takes one value in each of the terminal vertices, and is therefore $\mathcal{F}_2$-measurable.

In order to compute the price seller A_1 and seller A_2 will demand for B, one must (from Theorem 6.3) determine $\mathcal{M}^a(S,\mathcal{H})$ and $\mathcal{M}^a(S,\mathcal{J})$ (respectively).

To determine $\mathcal{M}^a(S,\mathcal{J})$, one must solve

$$\begin{aligned}
q_1 + \ldots + q_7 &= 1 \\
7(q_1+q_2) + 7(q_3+q_4) + 2(q_5+q_6+q_7) &= 4 \\
9\tfrac{q_1}{q_1+q_2} + 6\tfrac{q_2}{q_1+q_2} &= 7 \\
4\tfrac{q_3}{q_3+q_4} + 8\tfrac{q_4}{q_3+q_4} &= 7 \\
3\tfrac{q_5}{q_5+q_6+q_7} + 5\tfrac{q_6}{q_5+q_6+q_7} + 1\tfrac{q_7}{q_5+q_6+q_7} &= 2.
\end{aligned}$$

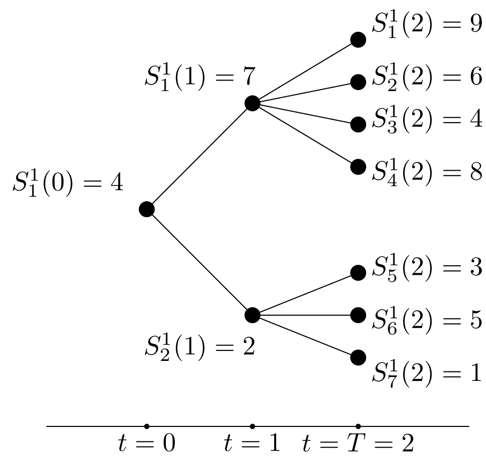

Figure 6.6: Seller A_1.

Solving this for q gives that the set

$$
\begin{aligned}
\mathcal{M}^a(S, \mathcal{J}) \quad &= \{q \in \mathbb{R}^7 : q_1 = 0.133 - 0.444q_4 \geq 0, \\
&\quad q_2 = 0.267 - 0.889q_4 \geq 0, q_3 = 0.333q_4, q_4 \geq 0, q_5 = 0.9 - 2q_7 \geq 0, \\
&\quad q_6 = -0.3 + q_7 \geq 0, q_7 \geq 0\} \\
&= \{q \in \mathbb{R}^7 : q_1 = 0.133 - 0.444q_4, q_2 = 0.267 - 0.889q_4, q_3 = 0.333q_4 \\
&\quad 0 \leq q_4 \leq 0.3, q_5 = 0.9 - 2q_7, \\
&\quad q_6 = -0.3 + q_7, 0.3 \leq q_7 \leq 0.45\}.
\end{aligned}
$$

Hence, the price of the claim B offered by seller A_2 will be

$$
\begin{aligned}
\sup_{q \in \mathcal{M}^a(S,\mathcal{J})} E_q[B] \quad &= \sup_{\{0 \leq q_4 \leq 0.3\}} \{2.27 - 0.22q_4\} \\
&= 2.27
\end{aligned}
$$

since the supremum is attained for $q_4 = 0$ (and for example $q_7 = 0.3$, and q_1, q_2, q_4, q_5, q_6 determined by the equations necessary to have $q \in \mathcal{M}^a(S, \mathcal{J})$).

By doing analogous calculations for seller A_1 and $\mathcal{M}^a(S, \mathcal{H})$, one finds that the price offered by seller A_1 is $\sup_{Q \in \mathcal{M}^a(S,\mathcal{H})} E_Q[B] = 3.20$. Hence, the price of seller A_1, the less informed seller, is greater than the price of seller A_2, the well informed seller. This confirms Theorem 6.5 for this example.

6.2 The pricing problem with inside information: Arbitrary Ω $\diamondsuit$

Consider again a financial market based on a probability space $(\Omega, \mathcal{F}, P)$ where Ω is *arbitrary* (as opposed to the situation in Section 6.1). As previously, there are $N+1$ assets: one bond and N risky assets, with the composed price process

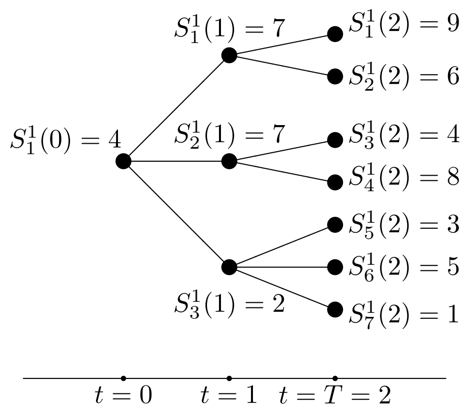

Figure 6.7: Seller A_2.

$S(t) = (S_0(t), S_1(t), ..., S_N(t))$ for discrete time $t \in \{0, 1, \dots, T\}$. Denote by $(\mathcal{F}_t)_t$ the filtration generated by the price process S.

Since Ω is arbitrary, the scenario tree model of Section 6.1 no longer applies. Also, neither linear programming or Lagrange duality work because the seller's pricing problem turns out to have infinitely many constraints. Fortunately, the conjugate duality theory of Rockafellar [34] can be applied in this setting.

Let $(\mathcal{G}_t)_{t \in \{0,1,\dots,T\}}$ be a filtration such that $\mathcal{G}_0 = \{\emptyset, \Omega\}$, $\mathcal{G}_T = \mathcal{P}(\Omega)$, where $\mathcal{P}(\Omega)$ denotes the set of all $\mathcal{F}$-measurable subsets of Ω. Assume that the price process S is adapted to $(\mathcal{G}_t)_t$, i.e. that $(\mathcal{F}_t)_t$ is nested in $(\mathcal{G}_t)_t$. As in Section 6.1, $(\mathcal{G}_t)_t$ represents the information being revealed to the seller at any time t. Also, assume that there exists there is no arbitrage with respect to $(\mathcal{G}_t)_t$. For example, if one is considering $(\mathcal{G}_t)_t := (\mathcal{F}_t)_t$, one may assume that there exists $Q \in \mathcal{M}^e(S, \mathcal{G})$ (so there is no arbitrage, from the fundamental theorem of mathematical finance, see Section 7.2).

The following lemma will be useful in what follows:

Lemma 6.6 ♢

Let f be any random variable w.r.t. $(\Omega, \mathcal{F}, P)$ and let $\mathcal{G}_t$ be a sub-σ-algebra of $\mathcal{F}$. Let $\mathcal{X}_t$ denote the set of all $\mathcal{G}_t$-measurable random variables. Then

$$\inf_{\{g \in \mathcal{X}_t\}} E[fg] > -\infty$$

if and only if $\int_A f dP = 0$ for all $A \in \mathcal{G}_t$.

Proof: ♢

- Assume there exists $A \in \mathcal{G}_t$ such that $\int_A f dP = K \neq 0$. Define $g(\omega) = M$ for all $\omega \in A$, where M is a constant. Also define $g(\omega) = 0$ for all $\omega \in \Omega \backslash A$.

Then g is $\mathcal{G}_t$-measurable (since $A \in \mathcal{G}_t$) and

$$\begin{array}{ll} E[gf] & = \int_\Omega gfdP \\ & = \int_A MfdP \\ & = MK \to -\infty \end{array}$$

by letting $M \to +/-\infty$ (depending on the sign of K). Hence,

$$\inf_{\{g \in \mathcal{X}_t\}} E[gf] > -\infty$$

implies that $\int_A fdP = 0$ for all $A \in \mathcal{G}_t$.

- Conversely, assume that $\int_A fdP = 0$ for all $A \in \mathcal{G}_t$. Let g be a simple, $\mathcal{G}_t$-measurable function, so $g = \sum_{A \in \mathcal{G}_t} g_A \mathbf{1}_A$, where g_A is a constant for all $A \in \mathcal{G}_t$ (and $\mathbf{1}_A$ denotes the indicator function of the set A). Then,

$$\begin{array}{ll} E[gf] & = \int_\Omega gfdP \\ & = \sum_{A \in \mathcal{G}_t} \int_A g_A fdP \\ & = \sum_{A \in \mathcal{G}_t} g_A \int_A fdP \\ & = 0. \end{array}$$

  Since all $\mathcal{G}_t$-measurable functions can be approximated arbitrarily well by a sequence of simple $\mathcal{G}_t$-measurable functions (see for example Shilling [41], Theorem 8.8), and $g \mapsto E[gf]$ is linear (hence continuous), $E[gf] = 0$ for any $\mathcal{G}_t$-measurable g. Therefore, $\int_A fdP = 0$ for all $A \in \mathcal{G}_t$ implies that $\inf_{\{g \in \mathcal{X}_t\}} E[gf] = 0 > -\infty$.

$\square$

The derivation of this section will also use the next lemma (where the notation is the same as in Lemma 6.6):

Lemma 6.7 $\inf_{\{g \in \mathcal{X}_t\}} E[fg] > -\infty$ *implies that* $\inf_{\{g \in \mathcal{X}_t\}} E[fg] = 0$.

Proof: Assume that $\inf_{\{g \in \mathcal{X}_t\}} E[gf] > -\infty$. It is known that $\inf_{\{g \in \mathcal{X}_t\}} E[gf] \leq 0$, since $g = 0$ (the zero-function) is a feasible solution. For contradiction, assume that

$$\inf_{\{g \in \mathcal{X}_t\}} E[gf] = K > -\infty, K \neq 0.$$

Then there exists a g which is $\mathcal{G}_t$-measurable such that $-\infty < E[gf] < 0$ (since there exists a sequence of g's such that $E[gf]$ is arbitrarily close to K, from the definition of the infimum). Then, Mg is also $\mathcal{G}_t$-measurable, so by letting $M \to \infty$, $E[Mgf] = ME[gf] \to -\infty$, hence $\inf_{\{g \in \mathcal{X}_t\}} E[gf] = -\infty$, which is a contradiction. Therefore, $\inf_{\{g \in \mathcal{X}_t\}} E[gf] > -\infty$ implies that

$$\inf_{\{g \in \mathcal{X}_t\}} E[gf] = 0.$$

$\square$

By combining Lemma 6.7 with Lemma 6.6, it follows that $\inf_{\{g \in \mathcal{X}_t\}} E[gf] = 0$ if and only if $\int_A f dP = 0$ for all $A \in \mathcal{G}_t$.

Consider a seller of a contingent claim B in the market with information corresponding to a filtration $(\mathcal{G}_t)_t$. Assume there is no arbitrage in the market w.r.t. the filtration $(\mathcal{G}_t)_t$. The pricing problem of the seller is

$$
\begin{array}{lrcll}
\inf_{v,H} & & v & & \\
\text{subject to} & & & & \\
& S(0) \cdot H(0) & \leq & v & \\
& B & \leq & S(T) \cdot H(T-1) & \text{for all } \omega \in \Omega, \\
& S(t) \cdot \Delta H(t) & = & 0 & \text{for } t \in \{1, \dots, T-1\}, \\
& & & & \text{and for all } \omega \in \Omega,
\end{array}
\tag{6.4}
$$

where the minimization is done over $v \in \mathbb{R}$ and all processes H that are adapted (using adapted or predictable here is just a matter of translation) to the filtration $(\mathcal{G}_t)_t$, since the seller must choose her trading strategy given what she knows at a particular time. Recall that the price process S is assumed to be adapted to $(\mathcal{G}_t)_t$, that $\mathcal{G}_0 = \{\emptyset, \Omega\}$ and $\mathcal{G}_T = \mathcal{P}(\Omega)$. Here $S(0), H(0)$ denotes the composed price process and trading strategy at time 0 (respectively) and $\Delta H(t) := H(t) - H(t-1)$.

Note also that the first constraint must hold with equality in an optimal solution (v, H), because if it held with strict inequality, the seller could fix the trading strategy H and decrease the price v without violating any of the constraints. Since this would give a better optimal value, (v, H) could not have been an optimal solution. Therefore, problem (6.4) is equivalent to

$$
\begin{array}{lrcll}
\inf_{v,H} & & v & & \\
\text{subject to} & & & & \\
& S(0) \cdot H(0) & = & v & \\
& B & \leq & S(T) \cdot H(T-1) & \text{for all } \omega \in \Omega, \\
& S(t) \cdot \Delta H(t) & = & 0 & \text{for } t \in \{1, \dots, T-1\}, \\
& & & & \text{and for all } \omega \in \Omega,
\end{array}
$$

which is again equivalent to

$$
\begin{array}{lrcll}
\inf_{H} & & S(0) \cdot H(0) & & \\
\text{subject to} & & & & \\
& B - S(T) \cdot H(T-1) & \leq & 0 & \text{for all } \omega \in \Omega, \\
& S(t) \cdot \Delta H(t) & = & 0 & \text{for } t \in \{1, \dots, T-1\}, \\
& & & & \text{and for all } \omega \in \Omega.
\end{array}
\tag{6.5}
$$

It is useful to transform the problem even more. Actually, problem (6.5) is equivalent to the following problem (6.6):

$$
\begin{array}{lrcll}
\inf_H & S(0)\cdot H(0) & & & \\
\text{subject to} & & & & \\
 & B - S(T)\cdot H(T-1) & \leq & 0 & \text{for all } \omega \in \Omega, \\
 & S(t)\cdot \Delta H(t) & = & 0 & \text{for } t \in \{1,\dots,T-1\}, \\
 & & & & \text{and for all } \omega \in \Omega, \\
 & S(0)\cdot H(0) & \geq & 0. &
\end{array} \tag{6.6}
$$

Why are problems (6.5) and (6.6) equivalent? Since there is no arbitrage w.r.t. $(\mathcal{G}_t)_t$ in the market, there is no self-financing trading strategy H such that $S(0)\cdot H(0) < 0$, but $S(T)\cdot H(T-1) \geq B$ since $B \geq 0$ (for all $\omega \in \Omega$). Hence, there is no feasible solution to problem (6.5) with $S(0)\cdot H(0) < 0$, and therefore the two problems are equivalent.

Problem (6.6) is (by basic algebra) equivalent to

$$
\begin{array}{lrcll}
\inf_H & S(0)\cdot H(0) & & & \\
\text{subject to} & & & & \\
 & B - S(T)\cdot H(T-1) & \leq & 0 & \text{for all } \omega \in \Omega, \\
 & S(t)\cdot \Delta H(t) & \leq & 0 & \text{for } t \in \{1,\dots,T-1\}, \\
 & & & & \text{and for all } \omega \in \Omega, \\
 & -S(t)\cdot \Delta H(t) & \leq & 0 & \text{for } t \in \{1,\dots,T-1\}, \\
 & & & & \text{and for all } \omega \in \Omega, \\
 & -S(0)\cdot H(0) & \leq & 0 & (\text{for all } \omega \in \Omega)).
\end{array} \tag{6.7}
$$

(Note that the final constraint may be assumed to hold for all $\omega \in \Omega$, since $S(0)$ and $H(0)$ are also random variables, we just happen to know they are deterministic.)

Now, the conjugate duality method of Rockafellar [34] will be applied. Therefore, define the perturbation space U by

$$U := \{\bar{u} := (u, (v_1^t)_t, (v_2^t)_t, w) : \bar{u} \in \mathcal{L}^p(\Omega, \mathcal{F}, P; \mathbb{R}^{2T})\}$$

where $1 \leq p < \infty$ and the $\mathbb{R}^{2T}$ originates from that there is one component of the perturbation function for each constraint in the primal problem. By counting the constraints one sees that there are $2T$ constraints that each hold for every $\omega \in \Omega$ (the constraint $-S(0)\cdot H(0) \leq 0$ also holds for each $\omega \in \Omega$, but since $H(0)$ and $S(0)$ are $\mathcal{G}_0$-measurable, $S(0), H(0)$ are constants).

Let Y be the dual space of U, so $Y := U^* = \mathcal{L}^q(\Omega, \mathcal{F}, P; \mathbb{R}^{2T})$ where $\frac{1}{p}+\frac{1}{q} = 1$ (see for example Pedersen [28] for more on dual spaces). Consider the pairing (see Definition 2.31) between U and V

$$\langle (u, (v_1^t)_t, (v_2^t)_t, w), y\rangle := E[y_1 u] + \textstyle\sum_{t=1}^{T-1}(E[y_2^t v_1^t] + E[y_3^t v_2^t]) + E[y_4 w].$$

Then U and Y are paired spaces from Example 2.35 (see also Pennanen and Perkkiö [30] for more on this). Actually, it is not necessary to assume that U and

Y are $\mathcal{L}^p$-spaces, it is sufficient that U is a decomposable space, see Pennanen and Perkkiö [30].

Define the perturbation function F as in Example 2.29, so for $(u, v_1, v_2, w) \in U$ where $v_i := (v_i^t)_{t\in\{1,\dots,T-1\}}$ for $i = 1, 2$, $F(H, (u, v_1, v_2, w)) := S(0)\cdot H(0)$ if $B(\omega) - S(T,\omega)\cdot H(T-1,\omega) \leq u(\omega)$ for all $\omega \in \Omega$ and $S(t,\omega)\cdot \Delta H(t,\omega) \leq v_1^t(\omega)$ for all $t \in \{1, ..., T-1\}$, $\omega \in \Omega$, $-S(t,\omega)\cdot\Delta H(t,\omega) \leq v_2^t(\omega)$ for all $t \in \{1, ..., T-1\}$, $\omega \in \Omega$, $-S(0)\cdot H(0) \leq w$ for all $\omega \in \Omega$, and $F(H, (u, v_1, v_2, w)) := \infty$ otherwise.

Hence,

$$
\begin{array}{rl}
K(H,y) & := \inf_{(u,v,z)\in U}\{F(H,(u,v,z)) + \langle (u,v,z), y\rangle\} \\
& = \begin{cases} S(0)\cdot H(0) \quad + & E[y_1(B - S(T)\cdot H(T-1))] - E[y_4 S(0)\cdot H(0)] \\ & +\sum_{t=1}^{T-1} E[(y_2^t - y_3^t)S(t)\cdot\Delta H(t)] \\ & \text{if } y_1, y_2^t, y_3^t, y_4 \geq 0 \text{ a.e. for all } t \in \{0,\dots,T-1\}, \\ -\infty & \text{otherwise.} \end{cases} \\
& = \begin{cases} S(0)\cdot H(0) \quad + & E[y_1(B - S(T)\cdot H(T-1))] \\ & +\sum_{t=1}^{T-1} E[\bar{y}_t S(t)\cdot\Delta H(t)] - E[y_4]S(0)\cdot H(0) \\ & \text{if } y_1, y_4 \geq 0 \text{ a.e. for all } t \in \{0,\dots,T-1\}, \\ -\infty & \text{otherwise.} \end{cases}
\end{array}
$$

where $\bar{y}_t := y_2^t - y_3^t$ is a free variable (i.e. the sign of $\bar{y}$ is not clear a priori).

Note that

$$
\begin{array}{rl}
S(0)\cdot H(0) & +E[y_1(B - S(T)\cdot H(T-1))] + \sum_{t=1}^{T-1} E[\bar{y}_t S(t)\cdot\Delta H(t)] \\
& -E[y_4]S(0)\cdot H(0) \\
= & E[S(0)\cdot H(0)] + E[y_1(B - S(T)\cdot H(T-1))] + \sum_{t=1}^{T-1} E[\bar{y}_t S(t)\cdot\Delta H(t)] \\
& -E[y_4 S(0)\cdot H(0)] \\
= & E[S(0)\cdot H(0)] + E[y_1(B - S(T)\cdot H(T-1))] + E[\bar{y}_1 S(1)\cdot(H(1) - H(0))] \\
& +E[\bar{y}_2 S(2)\cdot(H(2) - H(1))] + E[\bar{y}_3 S(3)\cdot(H(3) - H(2))] + ...+ \\
& +E[\bar{y}_{T-2} S(T-2)\cdot(H(T-2) - H(T-3))] \\
& +E[\bar{y}_{T-1} S(T-1)\cdot(H(T-1) - H(T-2))] - E[y_4 S(0)\cdot H(0)] \\
= & E[H(0)(S(0)(1-y_4) - S(1)\bar{y}_1)] + E[H(1)(\bar{y}_1 S(1) - \bar{y}_2 S(2))] \\
& +E[H(2)(\bar{y}_2 S(2) - \bar{y}_3 S(3))] \\
& +... + E[H(T-2)(\bar{y}_{T-2} S(T-2) - \bar{y}_{T-1} S(T-1))] \\
& +E[H(T-1)(\bar{y}_{T-1} S(T-1) - y_1 S(T))] + E[y_1 B] \\
= & E[H(0)(S(0)[1-y_4] - S(1)\bar{y}_1)] + \sum_{t=1}^{T-2} E[H(t)(\bar{y}_t S(t) - \bar{y}_{t+1} S(t+1))] \\
& +E[H(T-1)(\bar{y}_{T-1} S(T-1) - y_1 S(T))] + E[y_1 B]
\end{array}
$$

From Lemma 6.6 dual objective function g is given by

$$
\begin{aligned}
g(y) &:= \inf_{\{H\,:\,(\mathcal{G}_t)-\text{adapted}\}} K(H,y) \\
&= \begin{cases}
E[y_1 B] + & \inf_{H(0)}\{E[H(0)(S(0)[1-y_4] - S(1)\bar{y}_1)]\} + \\
& \sum_{t=1}^{T-2} \inf_{H(t)}\{E[H(t)(\bar{y}_t S(t) - \bar{y}_{t+1}S(t+1))]\} + \\
& \inf_{H(T-1)}\{E[H(T-1)(\bar{y}_{T-1}S(T-1) - y_1 S(T))]\} \\
& \text{ if } y_1 \geq 0, y_4 \geq 0 \text{ a.e.} \\
-\infty & \text{otherwise.}
\end{cases} \\
&= \begin{cases}
E[y_1 B] & \text{if } y_1, y_4 \geq 0 \text{ a.e.}, \int_A S(0)[1-y_4]dP = \int_A \bar{y}_1 S(1)dP \\
& \text{for all } A \in \mathcal{G}_0, \int_A \bar{y}_t S(t)dP = \int_A \bar{y}_{t+1}S(t+1)dP \\
& \text{for all } A \in \mathcal{G}_t, \text{ for } t \in \{1, ..., T-2\} \\
& \text{and } \int_A \bar{y}_{T-1}S(T-1)dP = \int_A y_1 S(T)dP \\
& \text{for all } A \in \mathcal{G}_{T-1}, \\
-\infty & \text{otherwise.}
\end{cases}
\end{aligned}
$$

where the final equality uses the comment after Lemmas 6.6 and 6.7.

Hence, the dual problem is

$$
\begin{array}{lrcl}
\sup_{y \in Y, y_1, y_4 \geq 0 \text{ a.e.}} & & E[y_1 B] & \\
\text{subject to} & & & \\
& \int_A \bar{y}_1 S(1)dP & = & \int_A [1-y_4]S(0)dP \\
& & & \text{for all } A \in \mathcal{G}_0, \\
& \int_A \bar{y}_{t+1}S(t+1)dP & = & \int_A \bar{y}_t S(t)dP \\
& & & \text{for all } A \in \mathcal{G}_t, \text{ for } t = 1, ..., T-2, \\
& \int_A y_1 S(T)dP & = & \int_A \bar{y}_{T-1}S(T-1)dP \\
& & & \text{for all } A \in \mathcal{G}_{T-1}.
\end{array}
\tag{6.8}
$$

Actually, problem (6.8) is equivalent to a simpler problem:

$$
\begin{array}{lrcl}
\sup_{y \in Y, y_1 \geq 0} & & E[y_1 B] & \\
\text{subject to} & & & \\
& \int_A \bar{y}_1 S(1)dP & = & \int_A S(0)dP \\
& & & \text{for all } A \in \mathcal{G}_0, \\
& \int_A \bar{y}_{t+1}S(t+1)dP & = & \int_A \bar{y}_t S(t)dP \\
& & & \text{for all } A \in \mathcal{G}_t, \text{ for } t = 1, ..., T-2, \\
& \int_A y_1 S(T)dP & = & \int_A \bar{y}_{T-1}S(T-1)dP \\
& & & \text{for all } A \in \mathcal{G}_{T-1}.
\end{array}
\tag{6.9}
$$

Why are problems (6.8) and (6.9) equivalent?

Clearly, if $\bar{y}$ is a feasible solution for problem (6.9), then $(\bar{y}, 0)$ is a feasible solution for problem (6.8), and the corresponding value functions coincide.

Hence, the optimal value of problem (6.9) is less than or equal the optimal value of problem (6.8).

Conversely, assume y is a feasible solution of problem 6.8, where $y_4 > 0$ on a set of P-measure greater than 0. What happens to the value function if y_4 is reduced to zero almost everywhere? From the dual feasibility conditions, the reduction of y_4 implies that one must increase $\int_A S(1)\bar{y}_1 dP$ for $A := \Omega \in \mathcal{G}_0$. This again implies an increase in $\int_A S(2)\bar{y}_2 dP$ for $A = \Omega$ (since $\mathcal{G}_0 \subseteq \mathcal{G}_1$), and by continuing to use the dual feasibility conditions systematically, the reduction of y_4 finally leads to an increase in $\int_A y_1 S(T) dP$ for $A = \Omega$. Hence, from the reduction of y_4, one has the freedom to choose a y_1' such that

$$\int_\Omega S(T) y_1' dP > \int_\Omega S(T) y_1 dP,$$

$S(T)$ is unaltered, so this means that one can increase y_1' (compared to y_1) on a set of measure greater than zero, without altering y_1' (compared to y_1) anywhere else, so $y_1'(\omega) = y_1(\omega)$ for all $\omega \in \Omega \setminus A$. Hence, one can choose $A \in \mathcal{G}_T$ such that $P(A) > 0$, $B(A) > 0$ and let $y_1'(\omega) > y_1(\omega)$ for all $\omega \in A$. But then

$$\begin{array}{rcl} E[(y_1' - y_1)B] & = & \int_\Omega (y_1' - y_1) B dP \\ & = & \int_A (y_1' - y_1) B dP + \int_{\Omega \setminus A} (y_1' - y_1) B dP \\ & = & \int_A (y_1' - y_1) B dP + 0 \\ & > & 0. \end{array}$$

This implies that $E[y_1' B] > E[y_1 B]$, hence, for any feasible solution y of problem (6.8) with $y_4 > 0$, one can construct another feasible solution y' with $y_4' = 0$ (almost everywhere), which corresponds to a greater value function. Hence, it is sufficient to only consider solutions where $y_4 = 0$ (P-a.e.) in problem (6.8). Therefore, problems (6.8) and (6.9) are equivalent.

From now on, we only consider problem (6.9). This problem will be transformed into a more familiar form. Actually, it will be shown that there is a one-to-one correspondence between feasible dual solutions (i.e. solutions to problem (6.9)) and absolutely continuous martingale measures w.r.t. $(\mathcal{G}_t)_t$.

The following analysis will use a rule for change of measure under conditional expectation, which is formulated as a lemma (see Davis [6]):

Lemma 6.8 *(Change of measure under conditional expectation) Let $(\Omega, \mathcal{F}, P)$ be a given probability space, and define a new probability measure Q by $\frac{dQ}{dP} = Z$, where Z is some $\mathcal{F}$-measurable random variable such that $Z \geq 0$ P-a.s. and $E[Z] = 1$. Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra of $\mathcal{F}$, and let X be any $\mathcal{F}$-measurable random variable. Then*

$$E[Z|\mathcal{G}] E_Q[X|\mathcal{G}] = E[ZX|\mathcal{G}].$$

Now, we will show the correspondence between feasible dual solutions and absolutely continuous martingale measures w.r.t. $(\mathcal{G}_t)_t$.

First, *assume that* $Q \in \mathcal{M}^a(S, \mathcal{G})$, i.e. that Q is an absolutely continuous martingale measure w.r.t. the filtration $(\mathcal{G}_t)_t$. We want to show that

Q corresponds to a feasible dual solution. The idea is to define $y_1 := \frac{dQ}{dP}$ (this is $\mathcal{F}$-measurable since any Radon-Nikodym derivative is measurable), and $\bar{y}_t := E[y_1|\mathcal{G}_t]$ (note that this definition is OK since $E[y_1|\mathcal{G}_t]$ is $\mathcal{G}_t$-measurable), and then check that the dual feasibility conditions of problem (6.9) hold.

- First, consider the final dual feasibility condition:

$$\int_A y_1 S(T)dP = \int_A \bar{y}_{T-1} S(T-1)dP \quad \text{for all } A \in \mathcal{G}_{T-1}. \tag{6.10}$$

  By using the definition of conditional expectation, this condition is equivalent to $\bar{y}_{T-1}S(T-1) = E[y_1 S(T)|\mathcal{G}_{T-1}]$. From the rule for change of measure under conditional expectation, Lemma 6.8,

$$E[y_1|\mathcal{G}_{T-1}]E_Q[S(T)|\mathcal{G}_{T-1}] = E[y_1 S(T)|\mathcal{G}_{T-1}].$$

  Hence, condition (6.10) is equivalent to

$$\bar{y}_{T-1}S(T-1) = E[y_1|\mathcal{G}_{T-1}]E_Q[S(T)|\mathcal{G}_{T-1}]. \tag{6.11}$$

  So it is sufficient to check that $y_1, \bar{y}_{T-1}$ satisfies equation (6.11). But this is clear, since $\bar{y}_{T-1} := E[y_1|\mathcal{G}_{T-1}]$, and $E_Q[S(T)|\mathcal{G}_{T-1}] = S(T-1)$ since Q is a martingale measure.

- To show that $y_1, (\bar{y}_t)_t$ satisfy the second dual feasibility equation, consider first this equation for $T-2$, i.e.

$$\int_A \bar{y}_{T-1}S(T-1)dP = \int_A \bar{y}_{T-2}S(T-2)dP \text{ for all } A \in \mathcal{G}_{T-2}.$$

  $\mathcal{G}_{T-2} \subseteq \mathcal{G}_{T-1}$, so every $A \in \mathcal{G}_{T-2}$ is also in $\mathcal{G}_{T-1}$. From the final dual feasibility equation, which was proven to hold for $y_1, \bar{y}_{T-1}$ in the previous item

$$\int_A \bar{y}_{T-1}S(T-1)dP = \int_A y_1 S(T)dP \text{ for all } A \in \mathcal{G}_{T-1}.$$

  In particular

$$\int_A \bar{y}_{T-1}S(T-1)dP = \int_A y_1 S(T)dP \text{ for all } A \in \mathcal{G}_{T-2}.$$

  Hence, it is sufficient to prove that

$$\int_A \bar{y}_{T-2}S(T-2)dP = \int_A y_1 S(T)dP \text{ for all } A \in \mathcal{G}_{T-2},$$

  which, from the definition of conditional expectation, is equivalent to

$$\bar{y}_{T-2}S(T-2) = E[y_1 S(T)|\mathcal{G}_{T-2}] = E[y_1|\mathcal{G}_{T-2}]E_Q[S(T)|\mathcal{G}_{T-2}], \tag{6.12}$$

  where the final equality follows from Lemma 6.8. But equation (6.12) holds since $\bar{y}_{T-2} := E[y_1|\mathcal{G}_{T-2}]$ and $E_Q[S(T)|\mathcal{G}_{T-2}] = S(T-2)$ since Q is a martingale measure. By the same kind of argument for $t = T-3, T-4, ..., 1$ (or by backwards induction), the second dual feasibility equation holds for all t.

- It remains to prove that the first dual feasibility condition holds for $\bar{y}_1$, i.e. that

$$\int_A \bar{y}_1 S(1) dP = \int_A S(0) dP \text{ for all } A \in \mathcal{G}_0$$

Since $\mathcal{G}_0 := \{\emptyset, \Omega\}$ and $\int_\emptyset \bar{y}_1 S(1) dP = \int_\emptyset S(0) dP$ trivially, it only remains to show that $E[\bar{y}_1 S(1)] = E[S(0)] = S(0)$. From the previous item it follows that

$$\begin{aligned} E[\bar{y}_1 S(1)] &= E[\bar{y}_2 S(2)] \\ &= \ldots(\text{induction}) \\ &= E[\bar{y}_{T-1} S(T-1)] \\ &= E[y_1 S(T)] \\ &= E_Q[S(T)] \\ &= E_Q[S(T)|\mathcal{G}_0] \\ &= S(0) \end{aligned}$$

where the final equality follows because Q is a martingale measure w.r.t. $(\mathcal{G}_t)$. But this implies that the first dual feasibility equation holds as well.

Hence, for each absolutely continuous martingale measure, there is a feasible dual solution.

Conversely, *assume there exists a feasible dual solution* $y_1, \bar{y}_t$ for $t = 1, \ldots, T-1$. We want to show that this dual solution corresponds to an equivalent martingale measure. Define $Q(F) := \int_F y_1(\omega) dP(\omega)$ for all $F \in \mathcal{F}$ (note that $y_1 \geq 0$, since it is feasible in the dual problem, and that one may assume $E[y_1] = 1$, since the dual problem is invariant under translation). The problem is to prove that for any $\bar{y}_t$, the dual feasibility conditions can be interpreted as martingale conditions.

The dual feasibility condition

$$\int_A y_1 S(T) dP = \int_A \bar{y}_{T-1} S(T-1) dP \text{ for all } A \in \mathcal{G}_{T-1}$$

is, from the definition of conditional expectation, equivalent to

$$\bar{y}_{T-1} S(T-1) = E[y_1 S(T)|\mathcal{G}_{T-1}] = E[y_1|\mathcal{G}_{T-1}] E_Q[S(T)|\mathcal{G}_{T-1}]$$

where the last equality follows from Lemma 6.8.

From this, it follows that $E_Q[S(T)|\mathcal{G}_{T-1}] = S(T-1)$ (which is a martingale condition) if $\bar{y}_{T-1} = E[y_1|\mathcal{G}_{T-1}]$.

Therefore, one must prove that $\bar{y}_{T-1} = E[y_1|\mathcal{G}_{T-1}]$: The third dual feasibility condition implies that $\int_A \bar{y}_{T-1} S(T-1) dP = \int_A y_1 S(T) dP$. This is a vector equation, so by considering the component corresponding to S_0 (the bond), and using that the market is assumed to be normalized, so $S_0(t, \omega) = 1$ for all $t \in \{0, 1, \ldots, T\}$ and for all $\omega \in \Omega$, it follows that

$$\int_A y_1 dP = \int_A \bar{y}_{T-1} dP \text{ for all } A \in \mathcal{G}_{T-1}.$$

From the definition of conditional expectation, this is equivalent to

$$\bar{y}_{T-1} = E[y_1|\mathcal{G}_{T-1}].$$

Hence, $E_Q[S(T)|\mathcal{G}_{T-1}] = S(T-1)$. We would like to prove this for a general time t, i.e. that $E_Q[S(T)|\mathcal{G}_t] = S(t)$. Take $t \in \{1, ..., T-2\}$. The second dual feasibility equation states that

$$\int_A \bar{y}_t S(t) dP = \int_A \bar{y}_{t+1} S(t+1) dP \text{ for all } A \in \mathcal{G}_t \text{ for } t \in \{1, ..., T-2\}.$$

Since $\mathcal{G}_t \subseteq \mathcal{G}_{t+1}$, this implies that $\int_A \bar{y}_t S(t) dP = \int_A \bar{y}_{t+2} S(t+2) dP$ for all $A \in \mathcal{G}_t$, so by induction, $\int_A \bar{y}_t S(t) dP = \int_A \bar{y}_{T-1} S(T-1) dP$ for all $A \in \mathcal{G}_t$. From the final dual feasibility condition, $\int_A \bar{y}_{T-1} S(T-1) dP = \int_A y_1 S(T) dP$ for all $A \in \mathcal{G}_{T-1}$, in particular, this holds for all $A \in \mathcal{G}_t$ (since $\mathcal{G}_t \subseteq \mathcal{G}_{T-1}$ because $t \leq T-1$). Hence

$$\int_A \bar{y}_t S(t) dP = \int_A y_1 S(T) dP \text{ for all } A \in \mathcal{G}_t. \tag{6.13}$$

From the definition of conditional expectation, (6.13) is equivalent to

$$\begin{aligned} \bar{y}_t S(t) &= E[y_1 S(T)|\mathcal{G}_t] \\ &= E[y_1|\mathcal{G}_t] E_Q[S(T)|\mathcal{G}_t] \end{aligned}$$

where the last equality uses Lemma 6.8. Hence, it suffices to show that $\bar{y}_t = E[y_1|\mathcal{G}_t]$.

By considering equation (6.13) for S_0 (the bond)

$$\int_A \bar{y}_t dP = \int_A y_1 dP \text{ for all } A \in \mathcal{G}_t$$

which, from the definition of conditional expectation, implies that $\bar{y}_t = E[y_1|\mathcal{G}_t]$. Hence, $E[S(T)|\mathcal{G}_t] = S(t)$ for $t = 1, ..., T-1$.

From the first dual feasibility equation, it follows that $E[S(T)|\mathcal{G}_0] = S(0)$: The first dual feasibility condition states that

$$\int_A \bar{y}_1 S(1) dP = \int_A S(0) dP \text{ for all } A \in \mathcal{G}_0.$$

From the same argument as before, using the second and third dual feasibility equations, and that $\mathcal{G}_0 \subseteq \mathcal{G}_t$ for all $t \geq 0$, it follows that $\int_A \bar{y}_1 S(1) dP = \int_A \bar{y}_{T-1} S(T-1)$ for all $A \in \mathcal{G}_0$. From the final dual feasibility condition, $\int_A \bar{y}_{T-1} S(T-1) dP = \int_A y_1 S(T) dP$ for all (in particular) $A \in \mathcal{G}_0$. Hence

$$\int_A S(0) dP = \int_A y_1 S(T) dP \text{ for all } A \in \mathcal{G}_0.$$

Since $\Omega \in \mathcal{G}_0$, this means in particular that

$$E[S(0)] = E[y_1 S(T)] = E_Q[S(T)] = E_Q[S(T)|\mathcal{G}_0]$$

so the martingale condition is OK also for $t = 0$. Note that $E_Q[S(T)|\mathcal{G}_T] = S(T)$, since $S(T)$ is $\mathcal{G}_T$-measurable.

Finally, to generalize this to the regular martingale condition $E[S(t)|\mathcal{G}_s] = S(s)$ for $s \leq t$, $s, t \in \{0, 1, ..., T\}$. Choose s and t such that $s \leq t$, and $s, t \in \{0, 1, ..., T\}$. Then

$$\begin{aligned} E_Q[S(t)|\mathcal{G}_s] &= E_Q[E_Q[S(T)|\mathcal{G}_t]|\mathcal{G}_s] \\ &= E_Q[S(T)|\mathcal{G}_s] \\ &= S(s). \end{aligned}$$

Here, the first equation follows from the martingale condition that has already been proved, and the second equality comes from the rule of double expectation, and that $\mathcal{G}_s \subseteq \mathcal{G}_t$ for $s \leq t$.

Hence, Q turns the price process S into a martingale, and each feasible dual solution corresponds to an absolutely continuous martingale measure w.r.t. $(\mathcal{G}_t)$, i.e. $Q \in \mathcal{M}^a(S, \mathcal{G})$.

To summarize, it has been proven that the dual problem of the seller's pricing problem (in a normalized market) for a seller with information corresponding to the filtration $(\mathcal{G}_t)_t$ can be rewritten

$$\sup_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B], \tag{6.14}$$

where $\mathcal{M}^a(S, \mathcal{G})$ denotes the family of all absolutely continuous martingale measures with respect to the filtration $(\mathcal{G}_t)_t$.

As for proving that there is no duality gap, i.e. that the value of problem (6.14) is equal to the value of problem (6.4), this can be done via Theorem 9 in Pennanen and Perkkiö [30], which will be called Theorem 6.9 in this thesis. This theorem is based on a conjugate duality setting similar to that of Rockafellar, and gives conditions for the value function φ (see Section 2.3) to be lower semi-continuous. Hence, from Theorem 2.44 (which is Theorem 7 in Rockafellar [34]), if these conditions hold, there is no duality gap (since $\varphi(\cdot)$ is convex, because the perturbation function F was chosen to be convex).

The theorem is rewritten to suit the notation of this thesis.

Theorem 6.9 *Assume there exists a $y \in Y$ and an $m \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ such that for P-a.e. $\omega \in \Omega$*

$$F(H, u, \omega) \geq u \cdot y(\omega) + m(\omega) \text{ for all } (H, u) \in \mathbb{R}^{T(N+1)} \times \mathbb{R}^{2T}, \tag{6.15}$$

where $\cdot$ denotes the standard Euclidean inner product. Assume also that $A := \{H \in \mathcal{H}_{\mathcal{G}} : F^{\infty}(H(\omega), 0, \omega) \leq 0 \ P - a.s.\}$ is a linear space. Then, the value function $\varphi(u)$ is lower semi-continuous on U and the infimum of the primal problem is attained for all $u \in U$.

In Theorem 6.9, $H \in \mathbb{R}^{T(N+1)}$ is a vector representing a stochastic process H with $N+1$ components at each time $t \in \{0, 1, \ldots, T-1\}$ and $\mathcal{H}_{\mathcal{G}}$ denotes the family of all stochastic processes that are adapted to the filtration $(\mathcal{G}_t)_t$.

F^∞ is the *recession function* of F. In general, if $h(x,\omega)$ is a function and $\mathrm{dom}(h(\cdot,\omega)) \neq \emptyset$, h^∞ is given by the formula (see Pennanen and Perkkiö [30])

$$h^\infty(x,\omega) = \sup_{\lambda>0} \frac{h(\lambda x + \bar{x},\omega) - h(\bar{x},\omega)}{\lambda} \tag{6.16}$$

(which is independent of $\bar{x}$).

For the proof of Theorem 6.9, see Pennanen and Perkkiö [30].

Actually, there is one difference between the frameworks of Rockafellar [34], and Pennanen and Perkkiö [30]. In [30] it is assumed that the perturbation function F is a so-called convex normal integrand (in order to be able to change the order of minimization and integration). Example 1 in Pennanen [29] considers the same choice of perturbation function F as we have chosen above, and this examples states that F is a convex normal integrand if the objective function and the constraint functions of the primal problem are convex normal integrands. Clearly, for our primal problem (6.7), all of these functions are convex. To check that they are normal integrands, an example from Rockafellar and Wets [36] will be applied. Example 14.29 in [36] states that all *Caratheodory integrands* are normal integrands, i.e. that all functions $f : \mathbb{R}^n \times \Omega \to \bar{\mathbb{R}}$ such that $f(x,\omega)$ is measurable in ω for each x, and continuous in x for each $\omega \in \Omega$, is a normal integrand (where x may depend measurably on ω). Hence, we check that this holds for one of the constraint functions of the primal problem (6.7). Let $f_t(H,\omega) := S(t,\omega) \cdot \Delta H(t,\omega)$. Choose $H \in \mathbb{R}^{(N+1)T}$. Then, $f_t(H,\omega) = S(t,\omega) \cdot \Delta H(t)$ (where $H(t)$ denotes the part of the vector $H \in \mathbb{R}^{(N+1)T}$ corresponding to time t) is $\mathcal{F}$-measurable, since $S(t)$ is $\mathcal{F}_t$-measurable for all t and $\mathcal{F}_t \subseteq \mathcal{F}$. Now, choose $\omega \in \Omega$, then $f_t(H,\omega) = S(t,\omega) \cdot \Delta H(t)$, which is continuous in H (since it is linear). Hence, f_t is a convex, normal integrand. Since the objective function and constraint functions of the primal problem (6.7) are similar (and B is $\mathcal{F}_T$-measurable), the same type of arguments prove that (from Example 14.29 in Rockafellar and Wets [36]) all of these functions are convex normal integrands. Hence (from Example 1 in Pennanen [29]), F is a convex normal integrand, and therefore, the framework of Pennanen and Perkkiö, in particular Theorem 6.9, can be applied.

Now, assume that the set A in Theorem 6.9 is a linear space. We would like to prove that the inequality (6.15) holds for the perturbation function corresponding to the primal problem (6.4).

As shown previously, the perturbation function takes the form

$$F(H,(u,v_1,v_2,w)) := S(0) \cdot H(0)$$

if $B(\omega) - S(T,\omega) \cdot H(T-1,\omega) \leq u(\omega)$ for all $\omega \in \Omega$ and $S(t,\omega) \cdot \Delta H(t,\omega) \leq v_1^t(\omega)$ for all $t \in \{1,...,T-1\}$, $\omega \in \Omega$, $-S(t,\omega) \cdot \Delta H(t,\omega) \leq v_2^t(\omega)$ for all $t \in \{1,...,T-1\}$, $\omega \in \Omega$, $-S(0) \cdot H(0) \leq w$ and $F(H,(u,v_1,v_2,w)) := \infty$ otherwise.

Now, choose $y(\omega) := (0,(0)_t,(0)_t,-1)$ for all $\omega \in \Omega$. Then, $y \in \mathcal{L}^q$ (since P is a finite measure). Also, choose $m(\omega) = -1$ for all $\omega \in \Omega$. $m \in \mathcal{L}^1$ since P is a finite measure.

Then:

$$\begin{array}{rcll} F(H,\bar{u}) & \geq & S(0)\cdot H(0) & \text{(from the definition of } F) \\ & \geq & -w & \text{(from the definition of } F) \\ & \geq & \bar{u}\cdot y(\omega)+m(\omega) & \text{(from the choice of } y \text{ and } m) \end{array}$$

for all $(H,\bar{u}) \in \mathbb{R}^{(N+1)T} \times \mathbb{R}^{2T}$.

Hence, the conditions of Theorem 6.9 are satisfied, and the theorem implies that there is no duality gap. This means that the seller's price of the contingent claim B (i.e. the optimal value of problem (6.4)) is equal to the optimal value of problem (6.14).

Hence, the following theorem has been proved.

Theorem 6.10 $\diamondsuit$

Consider a normalized market and a seller of the claim B with information corresponding to the filtration $(\mathcal{G}_t)_t$. Assume that there is no arbitrage in the market w.r.t. $(\mathcal{G}_t)_t$ and that $A := \{H \in \mathcal{H}_{\mathcal{G}} : F^{\infty}(H(\omega),0,\omega) \leq 0 \; P\text{- a.s.}\}$ is a linear space. Then, the seller's price of a contingent claim B is

$$\sup_{Q\in\mathcal{M}^a(S,\mathcal{G})} E_Q[B]$$

where $\mathcal{M}^a(S,\mathcal{G})$ denotes the family of all absolutely continuous martingale measures with respect to the filtration $\mathcal{G}_t$.

From general pricing theory, it is known that the seller's price, when the seller has full information corresponding to the filtration $(\mathcal{F}_t)_t$, will be

$$\sup_{Q\in\mathcal{M}^e(S,\mathcal{F})} E_Q[B]$$

(if the market is assumed to be normalized) where $\mathcal{M}^e(S,\mathcal{F})$ is the set of equivalent martingale measures w.r.t. the filtration $(\mathcal{F}_t)_t$. We will prove that this is consistent with Theorem 6.10. In order to do this, the following lemma is useful.

Lemma 6.11 *Assume there is no arbitrage in the market. For any $g \in \mathcal{L}^{\infty}(\Omega,\mathcal{F},P)$,*

$$E_Q[g] \leq 0 \textit{ for all } Q \in \mathcal{M}^a(S,\mathcal{F})$$

if and only if

$$E_Q[g] \leq 0 \textit{ for all } Q \in \mathcal{M}^e(S,\mathcal{F}).$$

Proof: Since there is no arbitrage in the market (by assumption), there exists at least one $Q^* \in \mathcal{M}^e(S,\mathcal{F})$ (and hence also a $Q \in \mathcal{M}^a(S,\mathcal{F})$, since $\mathcal{M}^e(S,\mathcal{F}) \subseteq \mathcal{M}^a(S,\mathcal{F})$) from the fundamental theorem of mathematical finance (since the time is assumed to be discrete, see Delbaen [7]). For any $Q \in \mathcal{M}^a(S,\mathcal{F})$ and $\lambda \in (0,1)$, $\lambda Q^* + (1-\lambda)Q \in \mathcal{M}^e(S,\mathcal{F})$. Hence, $\mathcal{M}^e(S,\mathcal{F})$ is dense in $\mathcal{M}^a(S,\mathcal{F})$, and the lemma follows. $\square$

Lemma 6.12 $\sup_{Q \in \mathcal{M}^a(S,\mathcal{F})} E_Q[B] = \sup_{Q \in \mathcal{M}^e(S,\mathcal{F})} E_Q[B]$.

Proof:

- To prove that $\sup_{Q \in \mathcal{M}^a(S,\mathcal{F})} E_Q[B] \leq \sup_{Q \in \mathcal{M}^e(S,\mathcal{F})} E_Q[B]$: Define $x := \sup_{Q \in \mathcal{M}^a(S,\mathcal{F})} E_Q[B]$. Then $E_Q[B] \leq x$ for all $Q \in \mathcal{M}^a(S,\mathcal{F})$. But from Lemma 6.11 (with $g = B - x$), this implies that $E_Q[B] \leq x$ for all $Q \in \mathcal{M}^e(S,\mathcal{F})$, so $\sup_{Q \in \mathcal{M}^e(S,\mathcal{F})} E_Q[B] \leq x$, and hence the inequality follows.
- The opposite inequality is shown similarly, or by using that $\mathcal{M}^e(S,\mathcal{F}) \subseteq \mathcal{M}^a(S,\mathcal{F})$.

$\square$

Hence, the seller's price which has been derived via conjugate duality is the same as the seller's price which is familiar from pricing theory.

We will now consider pricing with different levels of information. As in Section 6.1, consider two sellers of the same claim B. Seller A_1 has information corresponding to the filtration $(\mathcal{H}_t)_t$ where $\mathcal{H}_0 = \{\emptyset, \Omega\}$, and $\mathcal{H}_T = \mathcal{P}(\Omega)$ (the σ-algebra consisting of all measurable subsets of Ω) and $\mathcal{F}_t \subseteq \mathcal{H}_t$ for all t (so the seller A_1 knows the prices at all times). Similarly, seller A_2 has filtration $(\mathcal{J}_t)_t$ such that $\mathcal{J}_0 = \{\emptyset, \Omega\}$, and $\mathcal{J}_T = \mathcal{P}(\Omega)$. A_1 is assumed to have less information than A_2. Hence, $\mathcal{H}_t \subseteq \mathcal{J}_t$ for all $t = 0, 1, \ldots, T$, i.e. $(\mathcal{H}_t)_t$ is nested in $(\mathcal{J}_t)_t$. Since $(\mathcal{F}_t)_t$ is nested in $(\mathcal{H}_t)_t$, it is nested in $(\mathcal{J}_t)_t$ as well. Note that Lemma 6.4 goes through even though Ω is arbitrary, so $\mathcal{M}^a(S,\mathcal{J}) \subseteq \mathcal{M}^a(S,\mathcal{H})$. Therefore, Theorem 6.5 holds also for an arbitrary Ω, and the following theorem holds.

Theorem 6.13 $\diamondsuit$

If $(\mathcal{H}_t)_t$ and $(\mathcal{J}_t)_t$ are filtrations such that $(\mathcal{H}_t)_t$ is nested in $(\mathcal{J}_t)_t$, then

$$\sup_{Q \in \mathcal{M}^a(S,\mathcal{J})} E_Q[B] \leq \sup_{Q \in \mathcal{M}^a(S,\mathcal{H})} E_Q[B].$$

Hence, a well informed seller can offer the contingent claim B at a price which is less than or equal to the price offered by a less informed seller (if we assume that the set A of Theorem 6.10 is a linear space).

As previously, the buyer's problem is analogous to the seller's problem, so the price a buyer with filtration $(\mathcal{G}_t)_t$ is willing to pay for the claim B is $\inf_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B]$ (if the set A of Theorem 6.10 is a linear space). Hence, the same kind of arguments that lead to Theorem 6.13 imply that a well informed buyer will be willing to pay more for the claim B than a less informed buyer. This, together with Theorem 6.13, implies that the probability of a seller and buyer agreeing on a price is smaller in a market where the agents have little information than in a market with more information.

Finally, note that Lemma 6.11 also holds for a general filtration $(\mathcal{G}_t)_t$, so the price offered by a seller with filtration $(\mathcal{G}_t)_t$ will actually be $\sup_{Q \in \mathcal{M}^e(S,\mathcal{G})} E_Q[B]$, where $\mathcal{M}^e(S,\mathcal{G})$ denotes the family of all probability measures equivalent to P that turn S into a martingale with respect to the filtration $\mathcal{G}$ (if $\mathcal{M}^e(S,\mathcal{G}) \neq \emptyset$).

6.3 The pricing problem with short-selling constraints: Finite Ω $\diamondsuit$

Consider a financial market where the time t is discrete and the scenario space Ω is finite, so the scenario tree model of Section 4.3 applies. Given a contingent claim B in the market, consider a seller of this claim with filtration $(\mathcal{G}_t)_{t \in \{0,1,\ldots,T\}}$ such that $\mathcal{G}_0 = \{\emptyset, \Omega\}$, $\mathcal{G}_T$ is the σ-algebra that corresponds to the partition $\{\{\omega_1\}, ..., \{\omega_M\}\}$, and $\mathcal{F}_t \subseteq \mathcal{G}_t$ for $t = 0, 1, \ldots, T$. Hence the seller knows the prices of the assets at any time. We will consider the pricing problem of this seller, as in Section 6.1, but with a twist: the seller is not allowed to short-sell in one specific asset, asset $n^* \neq 0$ (i.e. we assume that the seller can short-sell in the bond). In the following, assume that the market is normalized, so $S_k^0 = 1$ for all k.

Hence, the pricing problem of this seller is

$$
\begin{array}{lrcll}
\min & & v & & \\
\text{subject to} & & & & \\
& S_0 \cdot H_0 & \leq & v, & \\
& B_k & \leq & S_k \cdot H_{a(k)} & \text{for all } k \in \mathcal{N}_T^{\mathcal{G}}, \\
& S_k \cdot H_k & = & S_k \cdot H_{a(k)} & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}},\ t \in \{1, 2, \ldots, T-1\}, \\
& H_k^{n^*} & \geq & 0 & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}}, t \in \{0, 1, \ldots, T-1\},
\end{array}
\tag{6.17}
$$

where the minimization is done with respect to $v \in \mathbb{R}$ and $H_k \in \mathbb{R}^{N+1}$ for $k \in \mathcal{N}_t^{\mathcal{G}}$ for $t = 0, ..., T-1$. Here, $\mathcal{N}_t^{\mathcal{G}}$ denotes the set of time t-vertices (nodes) in the scenario tree representing the filtration $\mathcal{G}$, and B_k denotes the value of the claim B in the vertex $k \in \mathcal{N}_T$. Recall that $a(k)$ denotes the ancestor of vertex k, see Section 4.3. Hence, the seller's problem is: Minimize the price v of the claim B such that the seller is able to pay B at time T from investments in a self-financing, predictable portfolio that costs less than or equal to v at time 0 and which does not sell short in the asset n^*. As in Section 6.1, note that the portfolio process H has been translated, so that H is adapted to $(\mathcal{G}_t)_t$, not predictable. This is done without loss of generality.

Problem (6.17) is a linear programming (LP) problem, and one can find the dual of this problem using standard LP-duality techniques. However, it turns out to be easier to find the dual problem via Lagrange duality (see Section 5.4). Since the linear programming dual problem is a special case of the Lagrange dual problem, the linear programming duality theorem implies that there is no duality gap (this can also be shown via the Slater condition, since $v > \sup_{\omega \in \Omega} B(\omega)$ and putting everything in the bank is a strictly feasible solution of problem (6.17)). Also, since problem (6.17) is an LP problem, the simplex algorithm is an efficient computational method which can be used to find optimal solutions in specific examples.

Problem (6.17) can be rewritten

$$
\begin{array}{lrcll}
\min & v & & & \\
\text{subject to} & & & & \\
& S_0 \cdot H_0 - v & \leq & 0 & \\
& B_k - S_k \cdot H_{a(k)} & \leq & 0 & \text{for all } k \in \mathcal{N}_T^{\mathcal{G}}, \\
& S_k \cdot (H_k - H_{a(k)}) & \leq & 0 & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}},\ t \in \{1, 2, \ldots, T-1\}, \\
& -S_k \cdot (H_k - H_{a(k)}) & \leq & 0 & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}},\ t \in \{1, 2, \ldots, T-1\}, \\
& H_k^{n^*} & \geq & 0 & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}}, t \in \{0, 1, \ldots, T-1\}.
\end{array} \tag{6.18}
$$

Problem (6.18) is suitable for the Lagrange duality method.

Let $y_0 \geq 0$, $z_k \geq 0$ for all $k \in \mathcal{N}_T^{\mathcal{G}}$ and $y_k^1, y_k^2 \geq 0$ for all $k \in \mathcal{N}_t^{\mathcal{G}}$, $t \in \{1, 2, \ldots, T-1\}$ be the Lagrange multipliers. Then, the Lagrange dual problem is

$$
\begin{array}{ll}
\sup_{y_0, z, y^1, y^2 \geq 0} & \inf_{v, H : H_k^{n^*} \geq 0\ \forall k} \{v + y_0(S_0 \cdot H_0 - v) + \sum_{k \in \mathcal{N}_T^{\mathcal{G}}} z_k(B_k - S_k \cdot H_{a(k)}) \\
& + \sum_{t=1}^{T-1} \sum_{k \in \mathcal{N}_t^{\mathcal{G}}} y_k S_k (H_k - H_{a(k)})\} \\
= \sup_{y_0, z \geq 0, y} \{ & \inf_v \{v(1 - y_0)\} \\
& + \inf_{H_0 : H_0^{n^*} \geq 0} \{y_0 S_0 \cdot H_0 - \sum_{m \in \mathcal{C}_{\mathcal{G}}(0)} y_m S_m \cdot H_0\} \\
& + \sum_{t=1}^{T-2} \sum_{k \in \mathcal{N}_t^{\mathcal{G}}} \inf_{H_k : H_k^{n^*} \geq 0} \{y_k S_k \cdot H_k - \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m S_m \cdot H_k\} \\
& + \sum_{k \in \mathcal{N}_{T-1}^{\mathcal{G}}} \inf_{H_k : H_k^{n^*} \geq 0} \{y_k S_k \cdot H_k - \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} z_m S_m \cdot H_k\} \\
& + \sum_{k \in \mathcal{N}_T^{\mathcal{G}}} z_k B_k\}
\end{array}
$$

where z, y^1, y^2 denotes the vectors of z_k's, y_k^1's and y_k^2's respectively and $y_k := y_k^1 - y_k^2$ is a free variable (i.e. the sign of y_k is not clear a priori).

Consider each of the minimization problems separately. In order to have a feasible dual solution, all of these minimization problems must have optimal value greater than $-\infty$.

- $\inf_v \{v(1 - y_0)\} > -\infty$ (that is, there is a feasible dual solution) if and only if $y_0 = 1$.
- $\inf_{\{H_0 \ :\ H_0^{n^*} \geq 0\}} \{H_0 \cdot (y_0 S_0 - \sum_{m \in \mathcal{C}_{\mathcal{G}}(0)} y_m S_m)\} > -\infty$ if and only if $y_0 S_0^n = \sum_{m \in \mathcal{C}_{\mathcal{G}}(0)} y_m S_m^n$ for all $n \neq n^*$, and $y_0 S_0^{n^*} \geq \sum_{m \in \mathcal{C}_{\mathcal{G}}(0)} y_m S_m^{n^*}$. In this case the infimum is equal to 0.
- $\inf_{\{H_k \ :\ H_k^{n^*} \geq 0\}} \{H_k \cdot (y_k S_k - \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m S_m)\} > -\infty$ if and only if $y_k S_k^n = \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m S_m^n$ for all $n \neq n^*$ and $y_k S_k^{n^*} \geq \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m S_m^{n^*}$. Note that in this case, the infimum is 0. Therefore, in order to get a dual solution, this must hold for all $k \in \mathcal{N}_t^{\mathcal{G}}$ for $t = 1, ..., T-2$.
- Finally, $\inf_{\{H_k \ :\ H_k^{n^*} \geq 0\}} \{H_k \cdot (y_k S_k - \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} z_m S_m)\} > -\infty$ if and only if $y_k S_k^n = \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} z_m S_m^n$ for all $n \neq n^*$ and $y_k S_k^{n^*} \geq \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} z_m S_m^{n^*}$.

In order to get a feasible dual solution this must hold for all $k \in \mathcal{N}_{T-1}^{\mathcal{G}}$. Note that in this case the infimum is 0.

Hence, the dual problem is

$$
\begin{array}{ll}
\sup_{y_0, z \geq 0, y} & \sum_{k \in \mathcal{N}_T^{\mathcal{G}}} z_k B_k \\
\text{subject to} &
\end{array}
$$

$$
\begin{array}{rcll}
y_0 & = & 1, & \\
y_k S_k^n & = & \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m S_m^n & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}}, \\
& & & t = 0, 1, ..., T-2, n \neq n^*, \\
y_k S_k^{n^*} & \geq & \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m S_m^{n^*} & \text{for all } k \in \mathcal{N}_t^{\mathcal{G}}, \\
& & & t = 0, 1, ..., T-2, \\
y_k S_k^n & = & \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} z_m S_m^n & \text{for all } k \in \mathcal{N}_{T-1}^{\mathcal{G}}, n \neq n^*, \\
y_k S_k^{n^*} & \geq & \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} z_m S_m^{n^*} & \text{for all } k \in \mathcal{N}_{T-1}^{\mathcal{G}}.
\end{array}
$$

By considering the second equation above for the non-risky asset (the bond), i.e. for S_k^0, one sees that in order to have a feasible dual solution

$$
y_k S_k^0 = \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m S_m^0
$$

must hold. Here, we have used that the assumption that the seller is allowed to short-sell in the bond, i.e. $n^* \neq 0$. Since the market is normalized, $S_k^0 = 1$ for all k, therefore

$$
y_k = \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m
$$

for all $k \in \mathcal{N}_t^{\mathcal{G}}$, $t = 1, ..., T-2$. Hence, in particular $\sum_{m \in \mathcal{C}_{\mathcal{G}}(0)} y_m = y_0 = 1$. Therefore $\sum_{k \in \mathcal{N}_{T-1}^{\mathcal{G}}} y_k = y_0 = 1$ (by induction). From the final dual feasibility condition (considered for the bond, i.e. $n = 0$), $\sum_{k \in \mathcal{N}_T^{\mathcal{G}}} z_k = \sum_{k \in \mathcal{N}_{T-1}^{\mathcal{G}}} y_k = 1$ and $z_k \geq 0$ for all k (since z is a Lagrange multiplier). Therefore, $\{z_k\}_{k \in \mathcal{N}_T^{\mathcal{G}}}$ can be identified with a probability measure Q (on the terminal vertices of the scenario tree) such that the Q-probability of ending up in terminal vertex k is z_k. Then, as in Section 6.1, the conditions $y_k S_k^n = \sum_{m \in \mathcal{C}_{\mathcal{G}}(k)} y_m S_m^n$ are martingale conditions (w.r.t. the filtration $(\mathcal{G}_t)_t$) for each asset $n \neq n^*$, of the form $S_{t-1}^n = E[S_t^n | \mathcal{G}_{t-1}]$, which can be shown to imply the general martingale condition in this discrete time case: Let $s \leq t$,

$$
\begin{array}{rcll}
E[S_t^n|\mathcal{G}_s] & = & E[E[S_t^n|\mathcal{G}_{t-1}]|\mathcal{G}_s] & \text{(from double expectation)} \\
& = & E[S_{t-1}^n|\mathcal{G}_s] & \text{(from the one step martingale condition)} \\
& = & E[E[S_{t-1}^n|\mathcal{G}_{t-2}]|\mathcal{G}_s] & \text{(double expectation)} \\
& = & E[S_{t-2}^n|\mathcal{G}_s] & \text{(one step martingale condition)} \\
& = & \dots & \\
& = & E[S_{s+1}^n|\mathcal{G}_s] & \\
& = & S_s^n &
\end{array}
$$

hence the martingale condition holds.

Similarly, the condition $y_k S_k^{n^*} \geq \sum_{m \in \mathcal{C}_\mathcal{G}(k)} y_m S_m^{n^*}$ is a super-martingale condition (w.r.t. the filtration $(\mathcal{G}_t)_t$):

From the definition of conditional expectation

$$\sum_{m \in \mathcal{C}_\mathcal{G}(k)} y_m S_m^{n^*} = y_k E_Q[S_{t+1}^{n^*}|\mathcal{G}_t]_k \text{ for all } k \in \mathcal{N}_t^\mathcal{G}$$

where $E_Q[S_{t+1}^{n^*}|\mathcal{F}_t]_k$ denotes the conditional expectation of $S_{t+1}^{n^*}$ given $\mathcal{G}_t$ w.r.t. the probability measure Q corresponding to the feasible dual solution y, evaluated in node $k \in \mathcal{N}_t^\mathcal{G}$. Hence

$$y_k S_k^{n^*} \geq y_k E_Q[S_{t+1}^{n^*}|\mathcal{F}_t]_k \text{ for all } k \in \mathcal{N}_t^\mathcal{G}$$

So, if $y_k > 0$, the super-martingale condition holds. If $y_k = 0$ (note that $y_k \geq 0$ since $z_k \geq 0$), then the Q-probability of node k happening is 0. Therefore, the conditional expectation $E_Q[S_{t+1}^{n^*}|\mathcal{F}_t]_k$ is defined (by convention) to be 0. Since the price process is non-negative, the super-martingale condition holds in this case as well.

This proves that any feasible dual solution can be identified with a probability measure Q which is absolutely continuous w.r.t. P such that assets $1, \dots, n^* - 1, n^* + 1, \dots, N$ are martingales w.r.t. Q and asset n^* is a super-martingale w.r.t. Q. Denote the set of such probability measures by $\bar{\mathcal{M}}_{n^*}^a(S, \mathcal{G})$.

The converse also holds: Take $Q \in \bar{\mathcal{M}}_{n^*}^a(S, \mathcal{G})$, and define $z_m := Q(\omega_m)$ for $m = 1, ..., M$, $y_k := \sum_{m \in \mathcal{C}_\mathcal{G}(k)} z_m$ for $k \in \mathcal{N}_{T-1}^\mathcal{G}$ and $y_k := \sum_{m \in \mathcal{C}_\mathcal{G}(k)} y_m$ for $k \in \mathcal{N}_t^\mathcal{G}$, $0 \leq t \leq T-2$. It can be checked (from these definitions) that this is a feasible dual solution.

Hence, the Lagrange dual problem can be rewritten

$$\beta := \sup_{Q \in \bar{\mathcal{M}}_{n^*}^a(S,\mathcal{G})} E_Q[B]$$

where the maximization is done over the set of probability measures on $(\Omega, \mathcal{G}_T)$ such that assets $1, \dots, n^* - 1, n^* + 1, \dots, N$ are martingales w.r.t. Q and asset n^* is a super-martingale w.r.t. Q (w.r.t. the filtration $(\mathcal{G}_t)_t$).

As explained previously, LP-duality implies that there is no duality gap. Hence, the optimal primal value, i.e. the seller's price of the contingent claim B, is equal to the optimal dual value, that is β.

The following theorem summarizes what we have shown in this section.

Theorem 6.14 $\diamondsuit$

Consider a normalized financial market with discrete time and finite scenario space Ω. Consider also a seller of a contingent claim B, who has information represented by the filtration $(\mathcal{G}_t)_t$, such that $(\mathcal{F}_t)_t$ is nested in $(\mathcal{G}_t)_t$, and who is not allowed to short-sell risky asset n^*. This seller will sell B at the price

$$\sup_{Q \in \bar{\mathcal{M}}^a_{n^*}(S,\mathcal{G})} E_Q[B]$$

where the maximization is done over the set of probability measures on $(\Omega, \mathcal{G}_T)$ such that assets $1, \ldots, n^* - 1, n^* + 1, \ldots, N$ are martingales w.r.t. Q and asset n^* is a super-martingale w.r.t. Q (w.r.t. the filtration $(\mathcal{G}_t)_t$).

Note that the previous arguments go through in the same way if there are short-selling constraints on several of the risky assets (as long as the seller is allowed to sell short in the bond), the only difference in the result will be that the final maximization in the dual problem will be over the set of probability measures that turn all the shorting-prohibited assets into super-martingales, and the rest of the assets into martingales.

Note that since all martingales are super-martingales

$$\sup_{Q \in \mathcal{M}^a(S,\mathcal{G})} E_Q[B] \leq \beta.$$

Hence, a seller who has short-selling prohibitions on one, or several, of the risky assets will be forced to demand a higher price for the claim B than a seller without prohibitions (for seller's with the same, general, level of inside information).

6.4 Pricing with short-selling constraints: Arbitrary Ω $\diamondsuit$

The goal of this section is to generalize the results of Section 6.3.

Consider the usual financial market, but with an arbitrary scenario space Ω. That is: Consider a financial market based on a general probability space $(\Omega, \mathcal{F}, P)$. There are N risky assets with price processes $S_1(t), \ldots, S_N(t)$ and one bond with price process $S_0(t)$ and the time $t \in \{0, 1, \ldots, T\}$. Let $(\mathcal{F}_t)_{t=0}^T$ be the filtration generated by the price process. Assume that the market is normalized, so $S_0(t, \omega) = 1$ for all $\omega \in \Omega$, $t \in \{0, 1, \ldots, T\}$.

The problem facing the seller of a contingent claim B, who is not allowed to short-sell in risky asset n^*, and who has filtration $(\mathcal{F}_t)_t$, is

$$
\begin{array}{llcll}
\inf_{\{v\in\mathbb{R}, H \; (\mathcal{F}_t)_t\text{-adapted}\}} & v & & & \\
\text{subject to} & & & & \\
& S(0)\cdot H(0) & \leq & v, & \\
& S(T)\cdot H(T-1) & \geq & B & \text{for all } \omega\in\Omega, \\
& S(t)\cdot \Delta H(t) & = & 0 & \text{for all } t\in\{1,\dots,T-1\}, \\
& & & & \text{and for all } \omega\in\Omega, \\
& H_{n^*}(t) & \geq & 0 & \text{for all } t\in\{0,\dots,T-1\}, \\
& & & & \text{and for all } \omega\in\Omega.
\end{array}
\tag{6.19}
$$

Assume that there is no arbitrage in the market with respect to $(\mathcal{F}_t)_t$, i.e. that $\mathcal{M}^e(S,\mathcal{F})\neq\emptyset$ (from the fundamental theorem of mathematical finance, see Delbaen and Schachermayer [9]). This implies that problem (6.19) is equivalent to

$$
\begin{array}{llcll}
\inf_{\{v\in\mathbb{R}, H \; (\mathcal{F}_t)_t\text{-adapted}\}} & v & & & \\
\text{subject to} & & & & \\
& S(0)\cdot H(0) & \leq & v, & \\
& S(T)\cdot H(T-1) & \geq & B & \text{for all } \omega\in\Omega, \\
& S(t)\cdot \Delta H(t) & = & 0 & \text{for all } t\in\{1,\dots,T-1\}, \\
& & & & \text{and for all } \omega\in\Omega, \\
& H_{n^*}(t) & \geq & 0 & \text{for all } t\in\{0,\dots,T-1\}, \\
& & & & \text{and for all } \omega\in\Omega, \\
& H(0)\cdot S(0) & \geq & 0. &
\end{array}
\tag{6.20}
$$

Why are these two problems equivalent? Since there is no arbitrage in the market, there does not exist any self-financing trading strategy H such that $S(0)\cdot H(0)<0$, but $S(T)\cdot H(T-1)\geq 0$. Hence, there does not exist any self-financing trading strategy H such that $S(0)\cdot H(0)<0$ and $S(T)\cdot H(T-1)\geq B\geq 0$, and therefore problems (6.19) and (6.20) are equivalent.

From the same kind of argument as in Section 6.3, this problem can be rewritten

$$
\begin{array}{llcll}
\inf_{\{H \; (\mathcal{F}_t)_t\text{-adapted}\}} & S(0)\cdot H(0) & & & \\
\text{subject to} & & & & \\
& B - S(T)\cdot H(T-1) & \leq & 0 & \text{for all } \omega\in\Omega, \\
& S(t)\cdot \Delta H(t) & = & 0 & \text{for all } t\in\{1,\dots,T-1\}, \\
& & & & \text{and for all } \omega\in\Omega, \\
& -H_{n^*}(t) & \leq & 0 & \text{for all } t\in\{0,\dots,T-1\}, \\
& & & & \text{and for all } \omega\in\Omega, \\
& -H(0)\cdot S(0) & \leq & 0 & (\text{for all } \omega\in\Omega).
\end{array}
\tag{6.21}
$$

(Note that the final feasibility condition may be assumed to hold for all $\omega\in\Omega$, since $S(0)$ and $H(0)$ are random variables.)

Equation (6.21) fits into the conjugate duality framework of Rockafellar [34]. It turns out that the Lagrange function is the same whether one rewrites the equality constraint in problem (6.21) to two inequality constraints or not. We work with problem (6.21) to keep notation as simple as possible. Hence, let

$$U := \{u = (\bar{u}, (v_t)_{t=1}^{T-1}, (z_t)_{t=0}^{T-1}, x) : u \in \mathcal{L}^p(\Omega, \mathcal{F}, P : \mathbb{R}^{2T+1})\}$$

Let $Y := U^* = \mathcal{L}^q(\Omega, \mathcal{F}, P : \mathbb{R}^{2T+1})$, the dual space of U, where $\frac{1}{p} + \frac{1}{q} = 1$. Consider the following pairing of U and Y

$$\begin{array}{l} \langle(\bar{u}, (v_t)_{t=1}^{T-1}, (z_t)_{t=0}^{T-1}, (x_t)_{t=0}^{T-1}), y\rangle \\ = E[y_1 \bar{u}] + \sum_{t=1}^{T-1} E[y_2^t v_t] + \sum_{t=0}^{T-1} E[y_3^t z_t] + E[x y_4] \end{array}$$

Define (for notational convenience) $v := (v_t)_{t=1}^{T-1}$ and $z := (z_t)_{t=0}^{T-1}$. Choose the perturbation function F to be $F(H, (\bar{u}, v, z, x)) := S(0) \cdot H(0)$ if $B(\omega) - S(T, \omega) \cdot H(T-1, \omega) \leq \bar{u}(\omega)$ for all $\omega \in \Omega$, $S(t, \omega) \cdot \Delta H(t, \omega) = v_t(\omega)$ for all $t \in \{1, \ldots, T-1\}$, for all $\omega \in \Omega$, and $-H_{n^*}(t, \omega) \leq z_t(\omega)$ for all $t \in \{0, \ldots, T-1\}$, for all $\omega \in \Omega$, $-H(0) \cdot S(0) \leq x$ for all $\omega \in \Omega$, and $F(H, (\bar{u}, v, z, x)) := \infty$ otherwise.

Note that $v_t = S(t) \cdot \Delta H(t)$ is feasible since $S(t) \cdot \Delta H(t)$ is $\mathcal{F}_t$-measurable.

Note also that we, contrary to Section 6.3, introduce dual variables for the constraints $H_{n^*}(t, \omega) \geq 0$ (for all t, ω). This is done because the separable problems coming from the Lagrange function are more difficult to solve when Ω is arbitrary.

The Lagrange function is

$$K(H, y) := \inf_{(\bar{u}, v, z, x) \in U} \{F(H, (\bar{u}, v, z, x)) + \langle(\bar{u}, v, z, x), y\rangle\},$$

so

$$\begin{array}{rcl} K(H, y) & = & S(0) \cdot H(0) + E[y_1(B - S(T) \cdot H(T-1))] \\ & & + \sum_{t=1}^{T-1} E[y_2^t S(t) \cdot \Delta H(t)] - \sum_{t=0}^{T-1} E[y_3^t H_{n^*}(t)] \\ & & - E[y_4 S(0) \cdot H(0)] \end{array}$$

if $y_1 \geq 0, y_3^t, y_4 \geq 0$ a.e. for all $t \in \{0, \ldots, T-1\}$ and $K(H, y) = -\infty$ otherwise.

Consider

$$
\begin{array}{rl}
S(0)\cdot H(0) & +E[y_1 B] - E[y_4 S(0)\cdot H(0)] - E[y_1 S(T)\cdot H(T-1)] \\
& +(E[y_2^1 S(1)\cdot)(H(1)-H(0))] + E[y_2^2 S(2)\cdot)(H(2)-H(1))] + \ldots + \\
& +E[y_2^{T-1} S(T-1)\cdot(H(T-1)-H(T-2))]) \\
& -(E[y_3^0 H_{n^*}(0)] + E[y_3^1 H_{n^*}] + \ldots - E[y_3^{T-1} H_n(T-1)]) \\
= & E[\sum_{i\neq n^*} H_i(0)\{S_i(0)[1-y_4] - y_2^1 S_i(1)\}] \\
& +E[H_{n^*}(0)\{S_{n^*}(0)[1-y_4] - y_2^1 S_{n^*}(1) - y_3^0\}] \\
& +\sum_{t=1}^{T-2}\{E[\sum_{i\neq n^*} H_i(t)\{y_2^t S_i(t) - y_2^{t+1} S_i(t+1)\}] \\
& +E[H_{n^*}(t)\{y_2^t S_{n^*}(t) - y_2^{t+1} S_{n^*}(t+1) - y_3^t\}]\} \\
& +E[\sum_{i\neq n^*} H_i(T-1)\{-y_1 S_i(T) + y_2^{T-1} S_i(T-1)\} \\
& +H_{n^*}(T-1)\{-y_1 S_{n^*}(T) + y_2^{T-1} S_{n^*}(T-1) - y_3^{T-1}\}] \\
& +E[y_1 B]
\end{array}
$$

(recall that $S(0)\cdot H(0)$ is deterministic, so $S(0)\cdot H(0) = E[S(0)\cdot H(0)]$.)

Hence, the dual objective function is

$$
\begin{array}{rcl}
g(y) & := & \inf_{H:(\mathcal{F}_t)_t-\text{adapted}} K(H,y) \\
& = & E[y_1 B] + \sum_{i\neq n^*} \inf_{H_i(0)}\{E[H_i(0)\{S_i(0)(1-y_4) - y_2^1 S_i(1)\}]\} \\
& & +\inf_{H_{n^*}(0)}\{E[H_{n^*}\{S_{n^*}(0)(1-y_4) - y_2^1 S_{n^*}(1) - y_3^0\}]\} \\
& & +\sum_{t=1}^{T-2}(\sum_{i\neq n^*} \inf_{H_i(t)}\{E[H_i(t)(y_2^t S_i(t) - y_2^{t+1} S_i(t+1))]\} \\
& & +\inf_{H_{n^*}(t)}\{E[H_{n^*}(t)(y_2^t S_{n^*}(t) - y_2^{t+1} S_{n^*}(t+1) - y_3^t)]\}) \\
& & +\sum_{i\neq n^*} \inf_{H_i(T-1)}\{E[H_i(T-1)(-y_1 S_i(T) + y_2^{T-1} S_i(T-1))]\} \\
& & +\inf_{H_{n^*}(T-1)}\{E[H_{n^*}(T-1)(-y_1 S_{n^*}(T) \\
& & +y_2^{T-1} S_{n^*}(T-1) - y_3^{T-1})]\}
\end{array}
$$

There exists a feasible dual solution if and only if all the minimization problems above have an optimal value greater than $-\infty$. Hence, by considering each of the minimization problems separately and using the comments after Lemma 6.6 and Lemma 6.7, the dual problem is

$$
\begin{array}{lll}
\sup_{\{y \in Y: y_1 \geq 0, y_3^t, y_4 \geq 0 \text{ a.e. } \forall\ t\}} & & E[y_1 B] \\
\text{subject to} & & \\
\int_A S_i(0)(1-y_4)dP & = & \int_A y_2^1 S_i(1)dP \text{ for all } A \in \mathcal{F}_0, i \neq n^*, \\
\int_A S_{n^*}(0)(1-y_4)dP & = & \int_A y_2^1 S_{n^*}(1)dP + \int_A y_3^0 dP \text{ for all } A \in \mathcal{F}_0, \\
\int_A y_2^t S_i(t)dP & = & \int_A y_2^{t+1} S_i(t+1)dP \text{ for all } A \in \mathcal{G}_t, \\
& & t = 1, \ldots, T-1, i \neq n^*, \\
\int_A S_{n^*}(t) y_2^t dP & = & \int_A y_2^{t+1} S_{n^*}(t+1)dP + \int_A y_3^t dP + \int_A y_3^t dP \\
& & \text{for all } A \in \mathcal{F}_t, t = 1, \ldots, T-2, \\
\int_A y_1 S_i(T)dP & = & \int_A y_2^{T-1} S_i(T-1)dP \text{ for all } A \in \mathcal{G}_{T-1}, i \neq n^*, \\
\int_A y_2^{T-1} S_{n^*}(T-1)dP & = & \int_A y_1 S_{n^*}(T)dP + \int_A y_3^{T-1} dP \text{ for all } A \in \mathcal{F}_{T-1}.
\end{array}
\tag{6.22}
$$

Note that since $y_3^t \geq 0$ a.e. for all $t \in \{0, 1, \ldots, T-1\}$ and P is a probability measure (hence non-negative), $\int_A y_3^t dP \geq 0$ for all $A \in \mathcal{F}_t$ for all t. Therefore, all the conditions regarding y_3^t can be replaced by altering the conditions of the form $\int_a y_2^t S_{n^*}(t)dP = \int_A y_2^{t+1} S_{n^*}(t+1)dP + \int_A y_3^t dP$ for all $A \in \mathcal{F}_t$ by $\int_A y_2^t S_{n^*}(t)dP \geq \int_A y_2^{t+1} S_{n^*}(t+1)dP$ (similarly, $\int_A S_{n^*}(0)(1-y_4)dP = \int_A y_2^1 S_{n^*}(1)dP + \int_A y_3^0 dP$ for all $A \in \mathcal{F}_0$ is replaced by $\int_A S_{n^*}(0)(1-y_4)dP \geq \int_A y_2^1 S_{n^*}(1)dP$ for all $A \in \mathcal{F}_0$). Hence, problem (6.22) is equivalent to

$$
\begin{array}{llll}
\sup_{y \in Y: y_1 \geq 0, y_4 \geq 0 \text{ a.e. } \forall\ t} & & E[y_1 B] & \\
\text{subject to} & & & \\
\int_A S_i(0)(1-y_4)dP & = & \int_A y_2^1 S_i(1)dP & \text{for all } A \in \mathcal{F}_0, i \neq n^*, \\
\int_A S_{n^*}(0)(1-y_4)dP & \geq & \int_A y_2^1 S_{n^*}(1)dP & \text{for all } A \in \mathcal{F}_0, \\
\int_A y_2^t S_i(t)dP & = & \int_A y_2^{t+1} S_i(t+1)dP & \text{for all } A \in \mathcal{F}_t, \\
& & & t = 1, \ldots, T-2, i \neq n^*, \\
\int_A S_{n^*}(t) y_2^t dP & \geq & \int_A y_2^{t+1} S_{n^*}(t+1)dP & \text{for all } A \in \mathcal{F}_t, \\
& & & t = 1, \ldots, T-2, \\
\int_A y_1 S_i(T)dP & = & \int_A y_2^{T-1} S_i(T-1)dP & \text{for all } A \in \mathcal{F}_{T-1}, i \neq n^*, \\
\int_A y_2^{T-1} S_{n^*}(T-1)dP & \geq & \int_A y_1 S_{n^*}(T)dP & \text{for all } A \in \mathcal{F}_{T-1}.
\end{array}
\tag{6.23}
$$

From the same kind of arguments as in Section 6.2, one can show that it is sufficient to only maximize over feasible solutions y of problem (6.23) where $y_4 = 0$ P-almost everywhere (see the arguments after problem (6.9), and apply similar reasoning to the dual feasibility constraints with equality of problem (6.23)) if such a solution exists.

Hence, (if there exists a feasible dual solution where $y_4 = 0$ P-a.e.) the dual problem can be written

$$
\begin{array}{llll}
\sup_{y \in Y: y_1 \geq 0} & E[y_1 B] & & \\
\text{subject to} & & & \\
\int_A S_i(0)dP & = & \int_A y_2^1 S_i(1)dP & \text{for all } A \in \mathcal{F}_0, i \neq n^*, \\
\int_A S_{n^*}(0)dP & \geq & \int_A y_2^1 S_{n^*}(1)dP & \text{for all } A \in \mathcal{F}_0, \\
\int_A y_2^t S_i(t)dP & = & \int_A y_2^{t+1} S_i(t+1)dP & \text{for all } A \in \mathcal{F}_t, \\
 & & & t = 1, \dots, T-1, i \neq n^*, \\
\int_A S_{n^*}(t) y_2^t dP & \geq & \int_A y_2^{t+1} S_{n^*}(t+1)dP & \text{for all } A \in \mathcal{F}_t, \\
 & & & t = 1, \dots, T-2, \\
\int_A y_1 S_i(T)dP & = & \int_A y_2^{T-1} S_i(T-1)dP & \text{for all } A \in \mathcal{F}_{T-1}, i \neq n^*, \\
\int_A y_2^{T-1} S_{n^*}(T-1)dP & \geq & \int_A y_1 S_{n^*}(T)dP & \text{for all } A \in \mathcal{F}_{T-1}.
\end{array}
\tag{6.24}
$$

Actually, it will be shown that there is a one-to-one correspondence between feasible dual solutions where $y_4 = 0$ and a certain kind of measures $Q \in \bar{\mathcal{M}}^a_{n^*}(S, \mathcal{F})$ (defined similarly as in Section 6.3). Hence, since we have assumed that the set $\mathcal{M}^e(S, \mathcal{F}) \neq \emptyset$ and $\mathcal{M}^e(S, \mathcal{F}) \subseteq \bar{\mathcal{M}}^a_{n^*}(S, \mathcal{F})$ (since all martingales are super-martingales) there exists a feasible dual solution where $y_4 = 0$.

To prove this one-to-one correspondence, techniques similar to those of Section 6.2 will be used.

Recall that a stochastic process $(M_t)_t$ is a *super-martingale* on a filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_t)$ if $M_t \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ for all t, $(M_t)_t$ is $(\mathcal{F}_t)_t$-adapted and $E[M_t|\mathcal{F}_s] \leq M_s$ for all $s \leq t$.

Define, as in Section 6.3, $\bar{\mathcal{M}}^a_{n^*}(S, \mathcal{F})$ as the set of probability measures Q on $(\Omega, \mathcal{F})$ that are absolutely continuous w.r.t. P and are such that the price processes $S_1, \dots, S_{n^*-1}$, $S_{n^*+1}, \dots, S_N$ are Q-martingales (w.r.t. the filtration $(\mathcal{F}_t)_t$) and S_{n^*} is a Q-super-martingale (w.r.t. $(\mathcal{F}_t)_t$).

We will prove that the dual problem (6.24) is equivalent to another problem, namely

$$\sup_{Q \in \bar{\mathcal{M}}^a_{n^*}(S, \mathcal{F})} E_Q[B]. \tag{6.25}$$

The approach for doing this is similar to that of Section 6.2.

Therefore, *assume there exists a* $Q \in \bar{\mathcal{M}}^a_{n^*}(S, \mathcal{F})$. Define $y_1 := \frac{dQ}{dP}$ (the Radon-Nikodym derivative of Q w.r.t. P), and $y_2^t := E[y_1|\mathcal{F}_t]$ (this is OK since $E[y_1|\mathcal{F}_t]$ is $\mathcal{F}_t$-measurable) for $t = 1, \dots, T-1$. As in Section 6.2, we prove that y_1, y_2^t satisfy the dual feasibility conditions of problem (6.24). Note that the conditions for assets $1, \dots, n^*-1, n^*+1, \dots, N$ are precisely as in Section 6.2, and these price processes are Q-martingales (from the definition of $\bar{\mathcal{M}}^a_{n^*}(S, \mathcal{F})$), hence these dual feasibility conditions hold by the same arguments as in Section 6.2. The conditions for asset n^* must be checked separately.

- $\int_A y_2^{T-1} S_{n^*}(T-1)dP \geq \int_A y_1 S_{n^*}(T)dP$ for all $A \in \mathcal{F}_{T-1}$: Recall that $E[y_1 S_{n^*}(T)|\mathcal{F}_{T-1}]$ is the unique $\mathcal{F}_{T-1}$-measurable random variable such that

$$\int_A E[y_1 S_{n^*}(T)|\mathcal{F}_{T-1}]dP = \int_A y_1 S_{n^*}(T)dP \text{ for all } A \in \mathcal{F}_{T-1}.$$

Hence, it is sufficient to prove that

$$\int_A E[y_1 S_{n^*}(T)|\mathcal{F}_{T-1}]dP \leq \int_A y_2^{T-1} S_{n^*}(T-1)dP \text{ for all } A \in \mathcal{F}_{T-1}.$$

In particular, it suffices to prove that

$$E[y_1 S_{n^*}(T)|\mathcal{F}_{T-1}] \leq y_2^{T-1} S_{n^*}(T-1) \text{ for all } \omega \in \Omega.$$

From the rule for change of measure under conditional expectation, Lemma 6.8

$$E[y_1 S_{n^*}(T)|\mathcal{F}_{T-1}] = E[y_1|\mathcal{F}_{T-1}] E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}].$$

Hence, it is enough to show that

$$E[y_1|\mathcal{F}_{T-1}] E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}] \leq y_2^{T-1} S_{n^*}(T-1).$$

But this is OK, since $y_2^{T-1} := E[y_1|\mathcal{F}_{T-1}]$ and $E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}] \leq S_{n^*}(T-1)$, since S_{n^*} is a Q-super-martingale.

- $\int_A S_{n^*}(t) y_2^t dP \geq \int_A y_2^{t+1} S_{n^*}(t+1)dP$ for all $A \in \mathcal{F}_t, t = 1, \ldots, T-2$: We begin by proving this inequality for $T-2$, i.e. we want to show that

$$\int_A y_2^{T-2} S_{n^*}(T-2)dP \geq \int_A y_2^{T-1} S_{n^*}(T-1)dP \text{ for all } A \in \mathcal{F}_{T-2}.$$

Note that the following holds for all $A \in \mathcal{F}_{T-2}$

$$\begin{array}{ll} \int_A y_2^{T-1} S_{n^*}(T-1)dP & = \int_A E[y_2^{T-1} S_{n^*}(T-1)|\mathcal{F}_{T-2}]dP \\ & = \int_A E[E[y_1|\mathcal{F}_{T-1}] S_{n^*}(T-1)|\mathcal{F}_{T-2}]dP \\ & = \int_A E[E[y_1 S_{n^*}(T-1)|\mathcal{F}_{T-1}]|\mathcal{F}_{T-2}]dP \\ & = \int_A E[y_1 S_{n^*}(T-1)|\mathcal{F}_{T-2}]dP \end{array}$$

where the first equality uses the definition of conditional expectation, the second uses the definition of y_2^{T-1}, the third uses that $S_{n^*}(T-1)$ is $\mathcal{F}_{T-1}$-measurable and the final equality uses the rule of double expectation and that $\mathcal{F}_{T-2} \subseteq \mathcal{F}_{T-1}$. Hence, from Lemma 6.8, it is sufficient to prove that

$$\begin{array}{ll} y_2^{T-2} S_{n^*}(T-2) & \geq E[y_1 S_{n^*}(T-1)|\mathcal{F}_{T-2}] \\ & = E[y_1|\mathcal{F}_{T-2}] E_Q[S_{n^*}(T-1)|\mathcal{F}_{T-2}]. \end{array}$$

But this is clearly true, since $y_2^{T-2} := E[y_1|\mathcal{F}_{T-2}]$ and $S_{n^*}(T-2) \geq E_Q[S_{n^*}(T-1)|\mathcal{F}_{T-2}]$ since asset n^* is a Q-super-martingale. The same kind of argument goes through for all the other dual feasibility conditions of this type.

- $\int_A S_{n^*}(0)dP \geq \int_A y_2^1 S_{n^*}(1)dP$ for all $A \in \mathcal{F}_0$: Recall that $\mathcal{F}_0 = \{\emptyset, \Omega\}$. The inequality is trivially true for $A = \emptyset$. Hence, it only remains to check that $E[y_2^1 S_{n^*}(1)] \leq E[S_{n^*}(0)] = S_{n^*}(0)$. Note that

$$\begin{aligned} E[y_2^1 S_{n^*}(1)] &= E[y_2^1 S_{n^*}(1)|\mathcal{F}_0] \\ &= E[E[y_1|\mathcal{F}_1]S_{n^*}(1)|\mathcal{F}_0] \\ &= E[E[y_1 S_{n^*}(1)|\mathcal{F}_1]|\mathcal{F}_0] \\ &= E[y_1 S_{n^*}(1)|\mathcal{F}_0] \\ &= E[y_1 S_{n^*}(1)] \\ &= E_Q[S_{n^*}(1)] \\ &= E_Q[S_{n^*}(1)|\mathcal{F}(0)] \\ &\leq S_{n^*}(0) \end{aligned}$$

  where the second equality follows from the definition of y_2^1, the fourth equality from the rule of double expectation and the inequality from that S_{n^*} is a Q-martingale. Hence, the final dual feasibility condition is OK as well.

This proves that any $Q \in \bar{\mathcal{M}}^a_{n^*}(S, \mathcal{F})$ corresponds to a feasible dual solution.

Conversely, *assume there exists a feasible dual solution $y_1 \geq 0, (y_2^t)_{t=1}^{T-1}$*. Define $Q(F) := \int_F y_1 dP$ for all $F \in \mathcal{F}$. This defines a probability measure since $y_1 \geq 0$, and one can assume that $E[y_1] = 1$ since the dual problem (6.24) is invariant under translation.

The goal is to show that $Q \in \bar{\mathcal{M}}^a_{n^*}(S, \mathcal{F})$, i.e. that the dual feasibility conditions can be interpreted as martingale and super-martingale conditions.

- First, consider the final dual feasibility condition for $i \neq n^*$:

$$\int_A y_1 S_i(T)dP = \int_A y_2^{T-1} S_i(T-1)dP \text{ for all } A \in \mathcal{F}_{T-1} \tag{6.26}$$

  and for n^*:

$$\int_A y_1 S_{n^*}(T)dP \leq \int_A y_2^{T-1} S_{n^*}(T-1)dP \text{ for all } A \in \mathcal{F}_{T-1}. \tag{6.27}$$

  The definition of conditional expectation implies that equation (6.26) is equivalent to $E[y_1 S_i(T)|\mathcal{F}_{T-1}] = y_2^{T-1} S_i(T-1)$. But

$$E[y_1 S_i(T)|\mathcal{F}_{T-1}] = E[y_1|\mathcal{F}_{T-1}]E_Q[S_i(T)|\mathcal{F}_{T-1}]$$

  from change of measure under conditional expectation (see Lemma 6.8). Hence, $E_Q[S_i(T)|\mathcal{F}_{T-1}] = S_i(T-1)$ if $y_2^{T-1} = E[y_1|\mathcal{F}_{T-1}]$, so this is what we want to prove.

  By considering equation (6.26) for the bond and using that the market is normalized (by assumption), so $S_0(t, \omega) = 1$ for all t, ω

$$\int_A y_1 dP = \int_A y_2^{T-1} dP \text{ for all } A \in \mathcal{F}_{T-1}.$$

From the definition of conditional expectation, this means that $y_2^{T-1} = E[y_1|\mathcal{F}_{T-1}]$. Hence, $E_Q[S_i(T)|\mathcal{F}_{T-1}] = S_i(T-1)$.

- The argument is similar for equation (6.27). From the definition of conditional expectation, equation (6.27) implies that

$$\int_A E[y_1 S_{n^*}(T)|\mathcal{F}_{T-1}]dP \leq \int_A y_2^{T-1} S_{n^*}(T-1)dP \text{ for all } A \in \mathcal{F}_{T-1}.$$

Note that from Lemma 6.8

$$\begin{aligned} \textstyle\int_A E[y_1 S_{n^*}(T)|\mathcal{F}_{T-1}]dP &= \textstyle\int_A E[y_1|\mathcal{F}_{T-1}] E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}]dP \\ &= \textstyle\int_A y_2^{T-1} E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}]dP \end{aligned}$$

where the last equality uses the result of the previous item. Hence

$$\int_A y_2^{T-1} E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}]dP \leq \int_A y_2^{T-1} S_n(T-1)dP \text{ for all } A \in \mathcal{F}_{T-1}.$$

So

$$0 \leq \int_A y_2^{T-1}(S_{n^*}(T-1) - E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}])dP \text{ for all } A \in \mathcal{F}_{T-1}.$$

Note that $y_2^{T-1} \geq 0$ since $y_2^{T-1} = E[y_1|\mathcal{F}_{T-1}]$ from the previous item, and $y_1 \geq 0$. Note also that for all $A \in \mathcal{F}_{T-1}$ such that $y_2^{T-1}(A) \geq 0$, but not identically equal 0 a.e., the super-martingale condition $S_{n^*}(T-1) \geq E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}]$ holds. If $y_2^{T-1}(A) = 0$ a.e., then $Q(A) = 0$, and therefore $E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}](A) := 0$ by convention, so also in this case, the super-martingale condition $S_{n^*}(T-1) \geq E_Q[S_{n^*}(T)|\mathcal{F}_{T-1}]$ holds. (Here $y_2^t(A)$ denotes the value y_2^t takes on A, the notation is similar for the other measurable random variables.)

We would like to show the same result for a general time t, i.e. we want to show that for $i \neq n^*$

$$E_Q[S_i(T)|\mathcal{F}_t] = S_i(T) \text{ for all } t \leq T.$$

This is OK by the same kind of argument as in Section 6.2. We also need to prove that

$$E_Q[S_{n^*}(T)|\mathcal{F}_t] \leq S_{n^*}(t) \text{ for all } t \leq T.$$

The second dual feasibility condition for asset n^* states that $\int_A y_2^t S_{n^*}(t)dP \geq \int_A y_2^{t+1} S_n(t+1)dP$ for all $A \in \mathcal{F}_t$. Note that by using this equation for $t+1, t+2, \ldots, T-2$, it follows that

$$\begin{aligned} \textstyle\int_A y_2^t S_{n^*}(t)dP &\geq \textstyle\int_A y_2^{t+1} S_{n^*}(t+1)dP && \forall A \in \mathcal{F}_t \\ &\geq \textstyle\int_A y_2^{t+2} S_{n^*}(t+2)dP && \forall A \in \mathcal{F}_{t+1}, \\ & && \text{in particular } \forall A \in \mathcal{F}_t \\ &\geq \ldots \\ &\geq \textstyle\int_A y_2^{T-1} S_{n^*}(T-1)dP && \forall A \in \mathcal{F}_t \\ &\geq \textstyle\int_A y_1 S_{n^*}(T)dP && \forall A \in \mathcal{F}_t \end{aligned}$$

where the final inequality uses the third dual feasibility condition. From the change of measure under conditional expectation, Lemma 6.8

$$\int_A y_1 S_{n^*}(T)dP = \int_A E[y_1|\mathcal{F}_t]E_Q[S_{n^*}(T)|\mathcal{F}_t]dP.$$

Hence

$$\int_A y_2^t S_{n^*}(t)dP \geq \int_A E[y_1|\mathcal{F}_t]E_Q[S_{n^*}(T)|\mathcal{F}_t]dP \;\forall\; A \in \mathcal{F}_t.$$

Recall that $y_2^t = E[y_1|\mathcal{F}_t]$, so

$$\int_A y_2^t(S_{n^*}(t) - E_Q[S_{n^*}(T)|\mathcal{F}_t])dP \geq 0 \;\forall\; A \in \mathcal{F}_t.$$

If $y_2^t(A) \geq 0$, but not identically equal 0 a.e., this implies the super-martingale condition

$$S_{n^*}(t, A) \geq E_Q[S_{n^*}(T)|\mathcal{F}_t](A) \text{ for all } A \in \mathcal{F}_t.$$

If $y_2^t(A) = 0$ a.e., then $Q(A) = 0$, so $E_Q[S_{n^*}(T)|\mathcal{F}_t](A) = 0$ by convention, and hence, since the price processes are non-negative, the super-martingale condition $S_{n^*}(t) \geq E_Q[S_{n^*}(T)|\mathcal{F}_t]$ holds.

From the same kind of arguments as in Section 6.2, for $i \neq n^*$, $E_Q[S_i(T)|\mathcal{F}_t] = S_i(t)$ (for all $t \leq T$) can be generalized to $E_Q[S_i(t)|\mathcal{F}_s] = S_i(s)$ for all $s \leq t$.

Similarly, $E_Q[S_{n^*}(t)|\mathcal{F}_s] \leq S_{n^*}(s)$ for all $s \leq t$: We know that

$$E_Q[S_{n^*}(T)|\mathcal{F}_s] \leq S_n(s) \;\forall\; s \leq T.$$

Consider $t \geq s$. Then, for all $A \in \mathcal{F}_s$,

$$\begin{aligned}
\int_A y_2^s S_{n^*}(s)dP &\geq \int_A y_2^t S_{n^*}(t)dP \\
&= \int_A E[y_2^t S_{n^*}(t)|\mathcal{F}_s]dP \\
&= \int_A E[E[y_1|\mathcal{F}_t]S_{n^*}(t)|\mathcal{F}_s]dP \\
&= \int_A E[E[y_1 S_{n^*}(t)|\mathcal{F}_t]|\mathcal{F}_s]dP \\
&= \int_A E[y_1 S_{n^*}(t)|\mathcal{F}_s]dP \\
&= \int_A E[y_1|\mathcal{F}_s]E_Q[S_{n^*}(t)|\mathcal{F}_s]dP
\end{aligned}$$

where the first inequality follows from second dual feasibility condition iterated, the first equality from the definition of conditional expectation, the second equality from that $y_2^t = E[y_1|\mathcal{F}_t]$ (from the same kind of arguments as in Section 6.2). The fourth equality uses the rule of double expectation and the final equality uses Lemma 6.8, regarding change of measure under conditional expectation.

Hence, since $y_2^s = E[y_1|\mathcal{F}_s] \geq 0$ (because $y_1 \geq 0$)

$$\int_A y_2^s(S_{n^*}(s) - E_Q[S_{n^*}(t)|\mathcal{F}_s])dP \geq 0 \text{ for all } A \in \mathcal{F}_s.$$

Again, if $y_2^s(A) \geq 0$, but not identically equal 0 a.e., the super-martingale condition $S_{n^*}(s, A) \geq E_Q[S_{n^*}(t)|\mathcal{F}_s](A)$ for all $A \in \mathcal{F}_s$ holds. If $y_2^s(A) = 0$ a.e., then $Q(A) = 0$, so by convention $E_Q[S_{n^*}(t)|\mathcal{F}_s](A) = 0$. Therefore, since the price processes are non-negative, the super-martingale condition $S_{n^*}(s, A) \geq E_Q[S_{n^*}(t)|\mathcal{F}_s](A)$ holds.

Hence

$$E_Q[S_{n^*}(t)|\mathcal{F}_s] \leq S_{n^*}(s) \text{ for all } s \leq t.$$

This implies that $Q \in \bar{\mathcal{M}}^a_{n^*}(S, \mathcal{F})$, and therefore, the dual problem of the seller's pricing problem with short-selling constraints (in a normalized market) is

$$\sup_{Q \in \bar{\mathcal{M}}^a_{n^*}(S,\mathcal{F})} E_Q[B].$$

Finally, to show that there is no duality gap, Theorem 6.9 will be applied. In order to apply this theorem, it is necessary to rewrite the perturbation function F so that the equality constraint is represented by two inequality constraints instead (in order to be consistent with the setting of Pennanen and Perkkiö). The alternative perturbation function (which, again, would result in the same Lagrange function as previously, it just requires extra notation) is $F(H, (\bar{u}, v^{(1)}, v^{(2)}, z, x)) = S(0) \cdot H(0)$ if $B - S(t) \cdot H(T-1) \leq \bar{u}$ for all ω, $S(t) \cdot \Delta H(t) \leq v^{(1)}_t$ for all t, ω, $-S(t) \cdot \Delta H(t) \leq v^{(2)}_t$ for all t, ω, $-H_{n^*}(t) \leq z_t$ for all t, ω, $-H(0) \cdot S(0) \leq x$, and $F(H, (\bar{u}, v^{(1)}, v^{(2)}, z, x)) = \infty$ otherwise. One can prove that F is a convex normal integrand (which is necessary for the framework of Pennanen and Perkkiö), from arguments similar to those of Section 6.2.

As in Section 6.2, assume that the set A of Theorem 6.9 is a linear space. To check the other assumption of the theorem, choose

$$y = (0, (0)_t, (0)_t, (0)_t, -1) \in \mathcal{L}^q(\Omega, \mathcal{F}, P : \mathbb{R}^{3T})$$

(since P is a finite measure), where 0 represents the 0-function. Also, choose $m(w) = -1$ for all $\omega \in \Omega$. Then $m \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ (again, since P is a finite measure). Then, given $(H, u) \in \mathbb{R}^{(N+1)T} \times \mathbb{R}^{2T+1}$:

$$\begin{array}{rcll} F(H,u) & \geq & S(0) \cdot H_0 & \\ & \geq & -x & \text{(from the definition of } F) \\ & = & u \cdot y(\omega) & \text{(from the choice of } y) \\ & \geq & m(\omega) + u \cdot y(\omega) & \text{(from the choice of } m) \end{array}$$

where the notation H_0 is used to emphasize that this is a deterministic vector. Hence, the the conditions of Theorem 6.9 are satisfied, and therefore, there is no duality gap, so the seller's price of the contingent claim is

$$\sup_{Q \in \bar{\mathcal{M}}^a_{n^*}(S,\mathcal{F})} E_Q[B].$$

Note that the previous derivation also goes through in the same way if there are short-selling constraints on several of the risky assets (the notation is just a little more complicated). Hence we have proved the following theorem.

Theorem 6.15 $\diamondsuit$

Consider a financial market with N risky assets and one bond, based on a filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t=0}^T)$ (where $(\mathcal{F}_t)_t$ is generated by the price process), where the time is discrete, $t \in \{0, 1, \dots, T\}$, and the scenario space Ω is arbitrary. Let $\bar{\mathcal{M}}^a_{n_1,\dots,n_K}(S, \mathcal{F})$ denote the set of probability measures Q which are absolutely continuous w.r.t. P, such that S_i for $i \neq n_1, \dots, n_K$ is a Q-martingale, and S_{n_k} is a Q-super-martingale for $k = 1, \dots, K$ (w.r.t. the filtration $(\mathcal{F}_t)_t$). Assume $\mathcal{M}^e(S, \mathcal{F}) \neq \emptyset$ (so the market has no arbitrage), and that the set $A := \{H \in \mathcal{H}_{\mathcal{G}} : F^{\infty}(H(\omega), 0, \omega) \leq 0$ P-a.s.$\}$ is a linear space (where F is defined as above). The seller of a contingent claim B who has short selling constraints on risky assets $n_1, \dots, n_K$ will offer the claim at a price

$$\sup_{Q \in \bar{\mathcal{M}}^a_{n_1,\dots,n_K}(S,\mathcal{F})} E_Q[B].$$

Note that the arguments of this section would go through in the same way if the seller's filtration was not $(\mathcal{F}_t)_t$, but some $(\mathcal{G}_t)_t$ such that $(\mathcal{F}_t)_t$ is nested in $(\mathcal{G}_t)_t$, i.e. some general level of inside information.

Also, since all martingales are super-martingales, the price offered by a seller without short-selling constraints will be less than or equal the price offered by the seller facing constraints (given that the two sellers have the same level of information).

6.5 The constrained pricing problem: Finite Ω $\diamondsuit$

Consider a normalized financial market based on a probability space $(\Omega, \mathcal{F}, P)$ where Ω is finite and the time $t \in \{0, 1, \dots, T\}$ is discrete. Recall that this means that the financial market can be modeled by a scenario tree as in Section 4.3. The financial market consists of N risky assets and one non-risky asset (bond). The price processes of the assets are $S_1, \dots, S_N$ and S_0 respectively. Let $(\mathcal{F}_t)_{t=0}^T$ denote the filtration generated by the price processes.

This section considers the pricing problem of a seller facing a constraint on how much she is allowed to purchase of the risky asset $n^* \in \{1, 2, \dots, N\}$. The constraint facing the seller is $H_{n^*}(t, \omega) \in [A, C]$ for all times $t \in [0, T]$ and all $\omega \in \Omega$, where $0 \in [A, C]$. Hence, the seller's problem for a claim B is:

$$\begin{array}{llll}
\min & v & & \\
\text{subject to} & & & \\
S_0 \cdot H_0 & \leq & v, & \\
B_k & \leq & S_k \cdot H_{a(k)} & \text{for all } k \in \mathcal{N}_T, \\
S_k \cdot H_k & = & S_k \cdot H_{a(k)} & \text{for all } k \in \mathcal{N}_t, t \in \{1, \dots, T-1\}, \\
H_k^{n^*} & \in & [A, C] & \text{for all } k \in \mathcal{N}_t, t \in \{0, 1, \dots, T-1\}
\end{array} \tag{6.28}$$

where $S_k := (S_k^0, S_k^1, \dots, S_k^N)$ and $H_k := (H_k^0, H_k^1, \dots, H_k^N)$.

Note that problem (6.28) is a linear programming problem, and hence the linear programming duality theorem implies that there is no duality gap. As in previous sections, the dual problem is found using Lagrange duality (which gives the linear programming dual problem, since linear programming is a special case of Lagrange duality). Recall that because problem (6.28) is a linear programming problem, the simplex algorithm is an efficient method for computing optimal solutions in specific examples.

We rewrite the equality constraint of problem (6.28) to two inequality constraints, in order to make problem (6.28) suitable for the Lagrange duality method:

$$
\begin{array}{lrcll}
\min & & v & & \\
\text{subject to} & & & & \\
& S_0 \cdot H_0 - v & \leq & 0, & \\
& B_k - S_k \cdot H_{a(k)} & \leq & 0 & \text{for all } k \in \mathcal{N}_T, \\
& S_k \cdot (H_k - H_{a(k)}) & \leq & 0 & \text{for all } k \in \mathcal{N}_t, t \in \{1, \dots, T-1\}, \\
& -S_k \cdot (H_k - H_{a(k)}) & \leq & 0 & \text{for all } k \in \mathcal{N}_t, t \in \{1, \dots, T-1\}, \\
& -H_k^{n^*} & \leq & -A & \text{for all } k \in \mathcal{N}_t, t \in \{0, 1, \dots, T-1\} \\
& H_k^{n^*} & \leq & C & \text{for all } k \in \mathcal{N}_t, t \in \{0, 1, \dots, T-1\}.
\end{array}
\tag{6.29}
$$

Let $y_0, (z_k)_k, (y_k^1)_k, (y_k^2)_k, (w_1^k)_k, (w_2^k)_k \geq 0$ (componentwise) be Lagrange multipliers. The Lagrange dual problem is

$$
\begin{array}{ll}
\sup \inf_{v,H} & \{v + y_0(S_0 \cdot H_0 - v) + \sum_{k \in \mathcal{N}_T} z_k (B_k - S_k \cdot H_{a(k)}) \\
& + \sum_{t=1}^{T} \sum_{k \in \mathcal{N}_t} (y_k^1 - y_k^2) S_k (H_k - H_{a(k)}) + \sum_{t=0}^{T-1} \sum_{k \in \mathcal{N}_t} w_1^k (A - H_k^{n^*}) \\
& + \sum_{t=0}^{T-1} \sum_{k \in \mathcal{N}_t} w_2^k (H_k^{n^*} - C)\} \\
= \sup\{ & \inf_v \{v(1 - y_0)\} + \sum_{n \neq n^*} \inf_{H_0^n} (y_0 S_0^n - \sum_{m \in \mathcal{C}(0)} y_m S_m^n) H_0^n \\
& + \inf_{H_0^{n^*}} (y_0 S_0^{n^*} - \sum_{m \in \mathcal{C}(0)} y_m S_m^{n^*} - w_1^0 + w_2^0) H_0^{n^*} \\
& + \sum_{t=1}^{T-2} \sum_{k \in \mathcal{N}_t} \sum_{n \neq n^*} \inf_{H_k^n} (y_k S_k^n - \sum_{m \in \mathcal{C}(k)} y_m S_m^n) H_k^n \\
& + \inf_{H_k^{n^*}} (y_k S_k^{n^*} - \sum_{m \in \mathcal{C}(k)} y_m S_m^{n^*} - w_1^k + w_2^k) H_k^{n^*} \\
& + \sum_{k \in \mathcal{N}_{T-1}} \sum_{n \neq n^*} \inf_{H_k} (y_k S_k^n - \sum_{m \in \mathcal{C}(k)} z_m S_m^n) H_k^n \\
& + \inf_{H_k^{n^*}} (y_k S_k^{n^*} - \sum_{m \in \mathcal{C}(k)} z_m S_m^{n^*} - w_1^k + w_2^k) H_k^{n^*} \\
& + \sum_{t=0}^{T-1} \sum_{k \in \mathcal{N}_t} A w_1^k - C w_2^k + \sum_{k \in \mathcal{N}_T} z_k B_k\}
\end{array}
\tag{6.30}
$$

where $y_k := y_k^1 - y_k^2$ for all $k \in \mathcal{N}_t, t \in \{1, \dots, T-1\}$, and the supremum is taken over all $y_0, (z_k)_k, (w_1^k)_k, (w_2^k)_k \geq 0, (y_k)$ free (componentwise).

Consider each of the minimization problems of equation (6.30) separately. In order for there to be a feasible dual solution, each of these problems must have optimal value greater than $-\infty$. This leads to the Lagrange dual problem

$$
\begin{array}{l}
\sup \sum_{k \in \mathcal{N}_T} z_k B_k + \sum_{t=0}^{T-1} \sum_{k \in \mathcal{N}_t} A w_1^k - C w_2^k \\
\text{subject to}
\end{array}
$$

$$
\begin{array}{lcll}
y_0 & = & 1 & \\
y_0 S_0^n & = & \sum_{m \in \mathcal{C}(0)} y_m S_m^n & \text{for all } n \neq n^*, \\
y_0 S_0^{n^*} & = & \sum_{m \in \mathcal{C}(0)} y_m S_m^{n^*} + w_1^0 - w_2^0, & \\
y_k S_k^n & = & \sum_{m \in \mathcal{C}(k)} y_m S_m^n & \text{for all } n \neq n^*, k \in \mathcal{N}_t, \\
 & & & t \in \{1, \ldots, T-2\}, \\
y_k S_k^{n^*} & = & \sum_{m \in \mathcal{C}(k)} y_m S_m^{n^*} + w_1^k - w_2^k & \text{for all } k \in \mathcal{N}_t, t \in \{1, \ldots, T-2\}, \\
y_k S_k^n & = & \sum_{m \in \mathcal{C}(k)} z_m S_m^n & \text{for all } n \neq n^*, k \in \mathcal{N}_{T-1}, \\
y_k S_k^{n^*} & = & \sum_{m \in \mathcal{C}(k)} z_m S_m^{n^*} + w_1^k - w_2^k & \text{for all } k \in \mathcal{N}_{T-1}
\end{array}
\tag{6.31}
$$

where the maximization is done over the set $\{y_0, (z_k), (w_1^k), (w_2^k) \geq 0\}$ (componentwise), (y_k) is a free variable (i.e. the sign is not clear a priori).

Note that since $w_1^k, w_2^k \geq 0$ for all k, $w_1^k - w_2^k$ is a free variable.

From the same kind of arguments as in Section 6.1 one can show that if there exists a feasible dual solution, i.e. a feasible solution to problem (6.31), then there exists a probability measure Q on $(\Omega, \mathcal{F})$ which is absolutely continuous w.r.t. P such that S_n is a Q-martingale for all $n \neq n^*$ (one cannot say anything about S_{n^*} being a martingale/super-martingale/sub-martingale etc.). Conversely, if there exists a probability measure Q which is absolutely continuous w.r.t. P such that S_n is a Q-martingale for all $n \neq n^*$, then there exists a feasible dual solution. The reasoning behind this is the same as in Section 6.1, except that one has to choose appropriate w_1^k, w_2^k for the feasibility equations involving S_{n^*} to hold. Hence, there exists a feasible dual solution if and only if there exists a probability measure Q which is absolutely continuous w.r.t. P such that S_n is a Q-martingale for all $n \neq n^*$. Note that these arguments use that the market is normalized, and that there are no constraints on the non-risky asset (bond).

Therefore, the dual problem can be rewritten

$$
\begin{array}{lll}
\sup & E_Q[B] + \sum_{t=0}^{T-1} \sum_{k \in \mathcal{N}_t} A w_1^k - C w_2^k & \\
\text{subject to} & & \\
 & w_1^k - w_2^k = Q(k)(S_k^{n^*} - E_Q[S^{n^*}(t+1)|\mathcal{F}_t]_k) & \text{for all } k \in \mathcal{N}_t \\
 & & t \in \{0, 1, \ldots, T-1\}.
\end{array}
\tag{6.32}
$$

where the maximization is done over the set $\{Q, (w_1^k), (w_2^k) \geq 0 : S_n \text{ is } Q-$martingale for all $n \neq n^*\}$, $E_Q[S^{n^*}(t+1)|\mathcal{F}_t]_k$ denotes the value of the random variable $E_Q[S^{n^*}(t+1)|\mathcal{F}_t]$ in the node $k \in \mathcal{N}_t$ and $Q(k)$ denotes the Q-probability of ending up in node k. As mentioned, the linear programming duality theorem implies that there is no duality gap, so the seller's price of the claim B is equal to the optimal value of problem (6.32).

Finally, note that the same kind of argument as above would go through if there was more than one constrained risky asset. This leads to a similar result as the one derived above.

This chapter has presented various applications of duality theory in mathematical finance. Sections 6.1, 6.3, and 6.5 use the scenario tree framework of Section 4.3 to convert the stochastic pricing problem into a linear programming problem. Lagrange duality is then applied to find the dual problem (since linear programming is a special case of Lagrange duality), and the linear programming duality theorem is used to prove that there is no duality gap. This is possible since the time is assumed to be finite and discrete, and the scenario space Ω is assumed to be finite. Sections 6.2 and 6.4 also consider a discrete time model, but with an arbitrary scenario space. This complicates things, and requires tools from stochastic analysis, functional analysis and measure theory. Also, the conjugate duality theory of Chapter 2 is used instead of Lagrange duality and linear programming. However, the results derived for finite Ω and arbitrary Ω are consistent. The next chapter will continue along these lines, but will consider arbitrage problems instead of pricing problems.

Chapter 7

Arbitrage, EMMs and duality

This chapter illustrates connections between duality methods and arbitrage-theory.

Section 7.1 considers a financial market model where the scenario space Ω is arbitrary and the time is finite and discrete. The main purpose of this section is to prove that for this model, there is no arbitrage in the market if and only if there is no free lunch with vanishing risk. This is done via the generalized Lagrange duality method of Section 5.4.

Finally, Section 7.2 proves a slightly weaker version of the fundamental theorem of mathematical finance via conjugate duality.

7.1 NFLVR and EMMs via Lagrange duality $\diamondsuit$

This section presents a proof of the equivalence between the existence of an equivalent martingale measure (EMM) and the no free lunch with vanishing risk condition (NFLVR), via generalized Lagrange duality (see Section 5.4), for finite, discrete time and arbitrary scenario space Ω. This is a slightly weaker result than the fundamental theorem of mathematical finance (also known as the Dalang-Morton-Willinger theorem, see Dalang et al. [5]).

The idea of this section is to begin with a version of the arbitrage problem (the NFLVR-problem), and show how one can quickly derive a dual problem which has a convenient form.

The setting is as follows. A probability space $(\Omega, \mathcal{F}, P)$ is given, where the scenario space Ω is arbitrary. As previously, there is one bond (non-risky asset) with price process $S_0(t)$, and N risky assets represented by the price processes $S_n(t)$, $n \in \{1, 2, \dots, N\}$ where the time $t \in \{0, 1, \dots, T\}$. Let $S = (S_0, S_1, ..., S_N)$ be the composed price process. Assume that the market is normalized, so $S_0(t) = 1$ for all $t \in \{0, 1, ..., T\}$. Also, let $(\mathcal{F}_t)_{t=0}^T$ be the filtration generated by the price process S, and consider the filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_t)$.

One says that there is *no free lunch with vanishing risk* if there does not exist any sequence of self-financing trading strategies with initial values 0 that approximate an arbitrage strategy, in the sense that the negative parts of the terminal value of the portfolio converge to 0 uniformly for all $\omega \in \Omega$ (see Delbaen and Schachermayer [9]).

Theorem 7.1 *Given the setting above, there is no free lunch with vanishing risk if and only if there exists an equivalent martingale measure.*

Proof: The idea of the proof is as follows

- Show that it is sufficient to show the theorem for $T = 1$. This proof is a version of the one in Delbaen and Schachermayer's *The Mathematics of Arbitrage* [9], section 6.7, adapted to this situation.
- Assume $T = 1$. Set up the NFLVR-problem, and call this the primal problem.
- Define the generalized Lagrange function as in Section 5.4.
- Derive the dual problem, which turns out to have a very convenient form.
- Show that there is no duality gap using the generalized Slater condition from Section 5.4, and deduce the theorem.

Now, for the actual proof: It is fairly simple to prove that the existence of an equivalent martingale measure implies NFLVR (see Delbaen and Schachermayer [8]), so this is omitted.

Instead, consider the opposite direction: NFLVR implies the existence of an equivalent martingale measure. Therefore, assume that NFLVR holds, i.e. that there is no free lunch with vanishing risk from time $t = 0$ to time T. This implies that there is no free lunch with vanishing risk from time $t = 0$ to time $t = 1$ (one time step), and that there is no free lunch with vanishing risk from time $t = 1$ to time T ($T-1$ time steps). Why is this? Assume, for contradiction, that there exists a free lunch with vanishing risk from time $t = 0$ to time $t = 1$. This implies that there exists a sequence of trading strategies $(\bar{H}_n^0)_n$ such that $(\bar{H}_n^0)_n$ is an arbitrage in the limit (the 0 in $\bar{H}_n^0$ symbolizes that the trading strategy is chosen at time $t = 0$). But then, $(H_n)_n$, such that $H_n^0 := \bar{H}_n^0$ and H_n^t, $t \geq 1$, simply places all values in the bond, is a free lunch with vanishing risk from time $t = 0$ to time T. This is a contradiction. A similar argument proves that there is no free lunch with vanishing risk from time $t = 1$ to time T.

Hence, NFLVR holds from time $t = 0$ to $t = 1$ and from time $t = 1$ to time T. To show that it is sufficient to consider $T = 1$, induction will be applied. We formulate what we would like to prove as a separate theorem:

Theorem 7.2 *Assume NFLVR holds for time $t = 0$ to time T. Then, there exists an equivalent martingale measure Q, i.e. a probability measure Q such that Q is equivalent to P and* *(i)* $S_t \in \mathcal{L}^1(\Omega, \mathcal{F}_T, Q)$ *for all* $t \in \{0, 1, \ldots, T\}$.

(ii) $(S_t)_{t=0}^T$ *is a* Q*-martingale, i.e.* $E_Q[S_t | \mathcal{F}_{t-1}] = S_{t-1}$ *for* $t = 1, 2, \ldots, T$.

(iii) The Radon-Nikodym derivative $\frac{dQ}{dP}$ *is bounded.*

Starting condition: Theorem 7.1 holds for $T = 1$ (this will be shown right after the induction).

Induction hypothesis: Assume Theorem 7.1 holds for time $T-1$ (i.e. for $T-1$ time steps).

Since NFLVR holds from time $t = 0$ to time T, it also holds from time $t = 1$ to time T (i.e. for $T-1$ time steps). Hence, the induction hypothesis (i.e. use of Theorem 7.2 for $T-1$ time- steps) implies that there exists a probability measure Q' on $(\Omega, \mathcal{F}_T)$ such that Q' is equivalent to P, $\frac{dQ'}{dP} \in \mathcal{L}^\infty$ and the price process $(S(t))_{t=1}^T$ is integrable w.r.t. Q' and is a martingale w.r.t. Q' (and the filtration $(\mathcal{F}_t)_t$). The martingale condition can be written

$$\int_A S_t dQ' = \int_A S_{t+1} dQ' \text{ for all } A \in \mathcal{F}_t, \; t \in \{1, 2, \ldots, T-1\}$$

(from the definition of conditional expectation).

By using the starting condition, i.e. the one time step version of Theorem 7.2 on the process $(S_t)_{t=0}^1$, the probability space $(\Omega, \mathcal{F}_1, Q')$ and the filtration $(\mathcal{F}_t)_{t=0}^1$, one finds that there exists a bounded function f_1 such that f_1 is $\mathcal{F}_1$-measurable, $f_1 > 0$, $E_{Q'}[f_1] = 1$ and such that

$$\int_A S(0) f_1 dQ' = \int_A S(1) f_1 dQ' \text{ for all } A \in \mathcal{F}_0.$$

(note that f_1 is actually the Radon-Nikodym derivative of the martingale measure $\bar{Q}$ coming from Theorem 7.2 w.r.t. the measure Q', so $f_1 := \frac{d\bar{Q}}{dQ'}$).

Then, define the measure Q on $(\Omega, \mathcal{F}_T)$ by

$$Q(A) = \int_A f_1 dQ' \text{ for all } A \in \mathcal{F}_T$$

(so $\frac{dQ}{dP} = f_1 \frac{dQ'}{dP}$).

Q is bounded (since Q' is bounded and f_1 is bounded), hence it can be transformed into a probability measure, also $\frac{dQ}{dP} > 0$ (since $f_1 > 0$), and hence Q is equivalent to P.

To check the integrability condition of Theorem 7.2, note that for $t = 1, 2, \ldots, T$,

$$\int_\Omega |S(t)| dQ = \int_\Omega |S(t)| f_1 dQ' < \infty,$$

since f_1 is bounded and $S_t \in \mathcal{L}^1(\Omega, Q', \mathcal{F}_T)$.

To check the martingale condition, note that for all $A \in \mathcal{F}_0$

$$\int_A S(0)dQ = \int_A S(0)f_1 dQ' = \int_A S(1)f_1 dQ' = \int_A S(1)dQ,$$

from the rule for change of measure and the martingale condition.

For $t \geq 1$, f_1 is $\mathcal{F}_t$-measurable (since it is $\mathcal{F}_1$-measurable, and $\mathcal{F}_1 \subseteq \mathcal{F}_t$ for all $t \geq 1$) and bounded, hence

$$\int_A S(t)dQ = \int_A S(t)f_1 dQ' = \int_A S(t+1)f_1 dQ' = \int_A S(t+1)dQ,$$

where the first and last equality follow from a change of measure, and the middle inequality can be proved by approximating the measurable function f_1 by a sequence of simple functions (this is always possible, see Shilling [41]) and using the martingale condition for the measure Q'.

Hence, the induction is OK, and it is sufficient to prove the theorem for $T = 1$.

For $T = 1$, the NFLVR-problem, which will be called the primal problem, takes the form

$$\begin{array}{lrcll} \inf_{H,k} & k & & & \\ \text{subject to} & & & & \\ & E[\Delta S \cdot H] & \geq & 1 & \\ & \Delta S \cdot H + k & \geq & 0 & \text{for all } \omega \in \Omega, \\ & k & \geq & 0, & \end{array} \tag{7.1}$$

where $H \in \mathbb{R}^{N+1}$ is the trading strategy vector chosen at time 0, S is the price process and k is a slack variable.

Why is this the NFLVR-problem? Assume that the optimal value of problem (7.1) is 0. Then there exists a free lunch with vanishing risk since there exists (from the definition of infimum) a sequence $(H_n, k_n) \in \mathbb{R}^{N+2}$ such that $k_n \to 0$, $E[\Delta S \cdot H_n] \geq 1 > 0$ and $\Delta S \cdot H_n = \Delta S \cdot H_n + k_n \geq 0$. Conversely, assume there exists a free lunch with vanishing risk, i.e. that there is a sequence $(H_n, k_n) \in \mathbb{R}^{N+2}$ such that $k_n \to 0$, $\Delta S \cdot H_n + k_n \geq 0$ for all $\omega \in \Omega$ and $E[\Delta S \cdot H_n] > 0$. Then, clearly, the optimal value of problem (7.1) is 0.

Rewriting problem (7.1) slightly, to suit the Lagrange duality setting of Luenberger [24], gives

$$\begin{array}{lrcll} \inf_{H,k} & k & & & \\ \text{subject to} & & & & \\ & 1 - E[\Delta S \cdot H] & \leq & 0 & \\ & -\Delta S \cdot H - k & \leq & 0 & \text{for all } \omega \in \Omega \\ & -k & \leq & 0. & \end{array}$$

Note that the composed constraints may be considered as a function G of (k, H) from $\mathbb{R} \times \mathcal{L}^p(\Omega, \mathcal{F}, P; \mathbb{R}^{N+1})$ into the normed space $R := \mathbb{R} \times \mathcal{L}^p(\omega, \mathcal{F}, P) \times \mathbb{R}$ (where we assume that the price process and trading strategies must be in $\mathcal{L}^p$ for each time t). This is a normed space with the norm $||(x,y,z)|| = |x| + ||y||_p + |z|$ (it is straightforward to check that this is a norm), where $(x,y,z) \in R$.

The dual space of R is $R^* = \mathbb{R} \times \mathcal{L}^q(\omega, \mathcal{F}, P) \times \mathbb{R}$, where $\frac{1}{p} + \frac{1}{q} = 1$. The Lagrange function L takes the form

$$L((H,k),(y_1,\bar{y},y_2)) = k + y_1(1 - E[\Delta S \cdot H]) + E[\bar{y}(-\Delta S \cdot H - k)] + y_2(-k),$$

where $(y_1, \bar{y}, y_2) \in R_+^* := \{(y_1, \bar{y}, y_2 \in R^* : y_1 \geq 0, \bar{y} \geq 0 \ P\text{-a.e.}, y_2 \geq 0)\}$.

Consider

$$\inf_{k,H} L((H,k),(y_1,\bar{y},y_2)) = \inf_k \{k(1 - y_2 - E[\bar{y}])\} + \inf_H \{-H(\Delta S(y_1 + \bar{y}))\} + y_1.$$

$\inf_k \{k(1 - y_2 - E[\bar{y}])\} > -\infty$ if and only if $1 - y_2 - E[\bar{y}] = 0$, and in this case the infimum is zero. Similarly, $\inf_H \{-H(\Delta S(y_1 + \bar{y}))\} > -\infty$ if and only if $E[\Delta S(y_1 + \bar{y})] = 0$, and in this case the infimum is zero (from the comments after Lemma 6.7).

Hence, the dual problem is

$$\begin{array}{lrcll} \sup_{y_1,y_2,\bar{y}} & & y_1 & & \\ \text{subject to} & & & & \\ & E[\bar{y}] & = & 1 - y_2 & \\ & E[(\bar{y} + y_1)\Delta S_n] & = & 0 & \text{for all } n \leq N, \\ & \bar{y} & \geq & 0 & P\text{-a.s.}, \\ & y_1, y_2 & \geq & 0. & \end{array}$$

The y_2-variable can be eliminated by requiring $E[\bar{y}] \leq 1$ (because then one can choose $y_2 := 1 - E[\bar{y}] \geq 0$ as a feasible solution). The final dual problem is

$$\begin{array}{lrcll} \sup_{\bar{y},y_1} & & y_1 & & \\ \text{subject to} & & & & \\ & E[\bar{y}] & \leq & 1 & \\ & E[(\bar{y} + y_1)\Delta S_n] & = & 0 & \text{for all } n \leq N, \\ & \bar{y} & \geq & 0 & P\text{-a.s.}, \\ & y_1 & \geq & 0. & \end{array} \tag{7.2}$$

So, problems (7.1) and (7.2) are Lagrange dual problems.

To prove that there is no duality gap, the generalized Slater condition of Section 5.4 (Theorem 5 in Luenberger [24]) will be applied. In the primal problem (7.1), the objective function is convex and real-valued, hence, we only need to find a strictly feasible solution to the primal problem to prove that there is no duality gap. Two cases will be considered. First, assume there exists a trading strategy H such that $E[\Delta S \cdot H] > 0$. Then (by multiplying H by a positive constant), there exists a trading strategy $\bar{H}$ such that $E[\Delta S \cdot \bar{H}] > 1$. Hence, $1 - E[\Delta S \cdot \bar{H}] < 0$. Let $k := |\inf_{\omega \in \Omega} \Delta S(\omega) \cdot H(\omega)| + 1$. Then, $k + \Delta S \cdot H > 0$ for all $\omega \in \Omega$ and $k > 0$. Therefore, this is a strictly feasible solution to the primal problem, and hence there is no duality gap from Theorem 5 in Luenberger [24].

Conversely, assume there does not exists a trading strategy H such that $E[\Delta S \cdot H] > 0$. In this case, we will not use Lagrange duality to prove the equivalence between EMM's and NFLVR, but do a direct argument instead. Note that under this assumption, there is is no trading strategy H such that $E[\Delta S \cdot H] < 0$ either, because if there was such a strategy H, $E[\Delta S(-H)] > 0$, which is a contradiction. Hence, every measurable random variable H is such that $E[\Delta S{\cdot}H] = 0$. In particular, choose a probability measure Q which is equivalent to P, and consider the Radon-Nikodym derivative $\frac{dQ}{dP}$. Then, $H := \frac{dQ}{dP}e_n$ (where $e_n \in \mathbb{R}^{N+1}$ denotes the n'th unit vector, so $e_n := (0, \dots, 0, 1, 0, \dots, 0)$), must be such that $E[\Delta S \cdot H] = 0$. But this implies that $E[\Delta S_n \frac{dQ}{dP}] = 0$ for all n. So, $E_Q[\Delta S_n] = 0$ for all n, so the price process S is a Q-martingale. Hence, from the choice of Q, Q is an equivalent martingale measure. Therefore, in particular, if there is no free lunch with vanishing risk (in this case), then there exists an equivalent martingale measure (since there always exists such a measure). Conversely, note that the assumption directly implies that there is no free lunch with vanishing risk (from the definition of NFLVR). Hence, Theorem 7.1 always holds under this assumption on the market.

Therefore, consider only the case where there exists a trading strategy H such that $E[\Delta S \cdot H] > 0$, and hence there is no duality gap, for the remainder of the proof.

To conclude the proof:

- From Delbaen and Schachermayer [8], the existence of an equivalent martingale measure implies that there is no free lunch with vanishing risk.

- Conversely, assume that there is no free lunch with vanishing risk, i.e. that the optimal primal value is greater than 0. Since there is no duality gap, the optimal dual value is greater than zero. Hence, there exists $\bar{y} \geq 0$, $y_1 > 0$ such that $E[\bar{y}] \leq 1$, $E[(\bar{y} + y_1)\Delta S_n] = 0$ for all n. Define

  $$Q'(F) := \int_F (\bar{y}(\omega) + y_1) dP(\omega) \text{ for all } F \in \mathcal{F}.$$

  ($\bar{y}$ is integrable since $E[\bar{y}] \leq 1$). Q' is a measure on $(\Omega, \mathcal{F})$ and Q' is equivalent to P (since $\bar{y} \geq 0$ P-a.e., so $\bar{y}(\omega) + y_1 > 0$ P-a.e.). Define

  $$Q(F) = \frac{Q'(F)}{Q'(\Omega)},$$

  then Q is a probability measure which is equivalent to P (note that $Q'(\Omega) < \infty$ since P is a probability measure, $E[\bar{y}] \leq 1$ and y_1 is bounded since it is the optimal primal value, which is less than ∞ since there always exists a feasible primal solution).

  Note that for all $n \leq N$,

$$\begin{array}{rcl} 0 & = & E_P[\Delta S_n(\bar{y}+y_1)] \\ & = & \int_\Omega \Delta S_n(\bar{y}+y_1)dP \\ & = & \int_\Omega \Delta S_n dQ' \\ & = & \int_\Omega \Delta S_n Q'(\Omega)dQ \\ & = & Q'(\Omega)\int_\Omega \Delta S_n dQ, \end{array}$$

so $E_Q[\Delta S_n] = \int_\Omega \Delta S_n dQ = 0$. Therefore, Q is an equivalent martingale measure.

Hence, it is proved that there is no free lunch with vanishing risk in the market if and only if there exists an equivalent martingale measure, and this completes the proof of Theorem 7.1. $\square$

7.2 Proof of a slightly weaker version of the FTMF $\diamondsuit$

This section presents an alternative proof of a version of the fundamental theorem of mathematical finance (abbreviated FTMF) for finite, discrete time and arbitrary scenario space. The proof is based on the conjugate duality theory introduced by Rockafellar [34], which was presented in Chapter 2.

The idea of the proof is to begin with the arbitrage problem, and derive a dual problem using conjugate duality. Then, Theorem 6.9 will be used to show that there is no duality gap. However, as in previous sections (for example in Section 6.2) we will assume that the set A of Theorem 6.9 is a linear space. Therefore, the result derived here is slightly weaker than the fundamental theorem of mathematical finance.

The setting is as follows. A probability space $(\Omega, \mathcal{F}, P)$ is given, where the scenario space Ω is arbitrary. As previously, there is one bond (non-risky asset) with price process $S_0(t)$, and N risky assets represented by the price processes $S_n(t)$, $n \in \{1, 2, \ldots, N\}$. The time $t \in \{0, 1, \ldots, T\}$, where $T < \infty$. $S = (S_0, S_1, ..., S_N)$ is the composed price process. Assume that the market is normalized, so $S_0(t) = 1$ for all $t \in \{0, 1, ..., T\}$. Also, let $(\mathcal{F}_t)_t$ be the filtration generated by the price process S.

The following theorem is due to Dalang, Morton and Willinger [5] and is called the fundamental theorem of mathematical finance.

Theorem 7.3 *(Fundamental theorem of mathematical finance) Given the setting above, there is no arbitrage if and only if there exists an equivalent martingale measure.*

To prove that the existence of an equivalent martingale measure implies that there is no arbitrage is quite simple, see Delbaen and Schachermayer [9]. To prove the other direction of Theorem 7.3, i.e. that no arbitrage implies the existence of an equivalent martingale measure, conjugate duality (see Chapter 2) will be applied. In order to prove this, it is sufficient to prove it for $T = 1$. This is proved in Delbaen and Schachermayer's *The Mathematics of Arbitrage* [9], section 6.7, but is omitted here. The proof is nearly identical to the argument in Section 7.1, and it is quite short. Therefore, assume $T = 1$. Now, the idea is as follows: set up the arbitrage problem, and call this the primal problem. Then, rewrite this problem to suit the setting of Chapter 2, and choose appropriate paired spaces. Define a perturbation function F as in Example 2.29, and define the Lagrange function as in Section 2.6. Then we will derive the dual problem, which turns out to have a very convenient form.

Proceeding with this plan: From Delbaen and Schachermayer [9] it is sufficient to consider $T = 1$. In this case, the arbitrage problem, which will be called the primal problem, takes the form

$$\alpha := \sup_{H\in\mathbb{R}^{N+1}}\{E_P[\Delta S \cdot H] : \Delta S \cdot H \geq 0\}, \tag{7.3}$$

where $\Delta S := S(1) - S(0) = (S_1(1) - S_1(0), ..., S_N(1) - S_N(0))$ is the price change vector and $H \in \mathbb{R}^{N+1}$ is the (deterministic) vector representing the trading strategy chosen at time 0.

This problem is equivalent to the following problem

$$\sup_{H\in\mathbb{R}^{N+1}}\{E_P[S(1) \cdot H] : S(0) \cdot H = 0, S(1) \cdot H \geq 0\},$$

which is more closely related to the definition of an arbitrage.

Note that the optimal value of problem (7.3) is always greater than or equal zero, since $H = 0 \in \mathbb{R}^{N+1}$ is a feasible solution. If the optimal value of problem (7.3) is zero, i.e. $\alpha = 0$, then there is no arbitrage. If the optimal value is greater than zero, then the problem is unbounded, because if H generates a positive optimal value, so will KH for $K > 0$. Therefore, by letting $K \to \infty$, the problem is unbounded, and there exists an arbitrage in the market.

Note also that reducing the problem to $T = 1$ removes the self-financing condition by taking it implicitly into the terminal portfolio value.

Rewriting problem (7.3) as a minimization problem (in order to get into the setting of Chapter 2) gives

$$\begin{array}{ll} \inf_{H\in\mathbb{R}^N} & -E_P[\sum_{n=1}^N \Delta S_n H_n] \\ \text{subject to} & \\ & -\sum_{n=1}^N \Delta S_n(\omega) H_n \leq 0 \text{ for all } \omega \in \Omega. \end{array} \tag{7.4}$$

(where we have omitted a minus in front of the infimum to simplify notation. This will be included later.)

Let $U := \mathcal{L}^p(\Omega, \mathcal{F}, P)$ (where $1 \leq p < \infty$) be the Banach space paired with $Y := (\mathcal{L}^p(\Omega, \mathcal{F}, P))^* = \mathcal{L}^q(\Omega, \mathcal{F}, P)$ where $\frac{1}{p} + \frac{1}{q} = 1$ (its dual, see for example Pedersen [28]) via the pairing $\langle u, y\rangle = \int_\Omega u(\omega)y(\omega)dP(\omega) = E_P[uy]$, $u \in U$, $y \in Y$. Note that, from Rockafellar [34], there is some ambiguity in the choice of U and Y. The choice of these spaces may make it more or less difficult to show the absence of a duality gap (without having to make any additional assumptions).

Define the perturbation function F by

$$F(H,u) = \begin{cases} -E_P[\sum_{n=1}^N \Delta S_n H_n] & \text{if } -\sum_{n=1}^N \Delta S_n(\omega) H_n \leq u(\omega) \text{ for all } \omega \in \Omega, \\ \infty & \text{otherwise.} \end{cases}$$

Then, F is a convex function (can be checked, and is also stated in Rockafellar [34]).

Now, define the corresponding Lagrange function

$$\begin{array}{rcl} K(H,y) & := & \inf_{u \in U}\{F(H,u) + \langle y,u \rangle\} \\ & = & \begin{cases} -E_P[\sum_{n=1}^N \Delta S_n H_n] + E_P[y(-\sum_{n=1}^N \Delta S_n H_n)] & \text{if } y \geq 0 \ P\text{-a.s.,} \\ -\infty & \text{otherwise.} \end{cases} \end{array}$$

Then, the dual objective function is defined by $g(y) := \inf_{H \in X} K(H,y)$, hence

$$\begin{array}{rcl} g(y) & = & \begin{cases} 0 & \text{if } y \geq 0 \ P\text{-a.e. and } E_P[\Delta S_n] = -E_P[y \Delta S_n] \text{ for all } n \leq N, \\ -\infty & \text{otherwise,} \end{cases} \\ & = & \begin{cases} 0 & \text{if } y \geq 0 \ P\text{-a.e. and } E_P[\Delta S_n(y+1)] = 0 \text{ for all } n \leq N, \\ -\infty & \text{otherwise.} \end{cases} \end{array}$$

Hence, the dual problem is

$$\sup_{y \in Y} g(y) = \sup_y 0, \tag{7.5}$$

where the final maximization is done over the set $\{y \in Y : y \geq 0 \ P\text{-a.e. and } E_P[\Delta S_n(y+1)] = 0 \text{ for all } n \leq N\}$. Note that the dual objective function is a constant, equal to 0, so the dual problem is really a feasibility problem.

Assume that there exists a feasible dual solution, i.e. a $y \in Y$ such that $y \geq 0$ P-a.e. and $E_P[\Delta S_n(y+1)] = 0$ for all $n \leq N$. Define

$$Q'(F) := \int_F (y(\omega)+1) dP(\omega) \text{ for all } F \in \mathcal{F}.$$

Q' is a measure on $(\Omega, \mathcal{F})$ and Q' is equivalent to P (since $y \geq 0$ P-a.e., so $y+1 > 0$ P-a.e.). Since $y \in Y = \mathcal{L}^q(\Omega, \mathcal{F}, P) \subseteq \mathcal{L}^1(\Omega, \mathcal{F}, P)$ (where the set inclusion follows since P is a probability measure, so all measurable sets have finite measure), y is integrable. Therefore, $y+1$ is integrable, so $Q'(\Omega) < \infty$. Define

$$Q(F) = \frac{Q'(F)}{Q'(\Omega)}.$$

Then, Q is a probability measure which is equivalent to P, and for all $n \leq N$,

$$\begin{aligned} 0 &= E_P[\Delta S_n(y+1)] \\ &= \int_\Omega \Delta S_n(y+1)dP \\ &= \int_\Omega \Delta S_n dQ' \\ &= \int_\Omega \Delta S_n Q'(\Omega)dQ \\ &= Q'(\Omega)\int_\Omega \Delta S_n dQ, \end{aligned}$$

hence $E_Q[\Delta S_n] = \int_\Omega \Delta S_n dQ = 0$. This implies that Q is an equivalent martingale measure. The same kind of argument can be done in the opposite direction to get a one-to-one correspondence between feasible dual solutions and equivalent martingale measures.

Recall that the original primal arbitrage problem (7.3) was transformed to suit the setting of Rockafellar [34]. In this transformation a minus-sign was skipped to simplify notation. This minus should be included into the dual problem. Therefore, the problem which is dual to (7.3) is

$$\inf_{\{Q: Q \text{ is EMM}\}} 0. \tag{7.6}$$

where EMM is short for equivalent martingale measure.

This derivation shows how equivalent martingale measures are naturally connected to the existence of arbitrage in the market.

Now, to show that there is no duality gap. Let

$$f(H,u) := \begin{cases} -\sum_{n=1}^N \Delta S_n H_n & \text{if } -\sum_{n=1}^N \Delta S_n(\omega)H_n \leq u(\omega) \text{ for all } \omega \in \Omega, \\ \infty & \text{otherwise.} \end{cases}$$

Assume that the set $A := \{H \in \mathcal{H}_\mathcal{G} : f^\infty(H(\omega), 0, \omega) \leq 0 \ P\text{-a.s.}\}$ of Theorem 6.9 is a linear space (see equation (6.16) for a definition of f^∞). The slight alteration of the perturbation function in the set A is due to a difference in notation between Rockafellar [34] and Pennanen and Perkkiö [30]. Note that the alternative perturbation function f is a convex normal integrand from arguments similar to those of Section 6.2, hence the framework of Pennanen and Perkkiö [30] can be applied. To show that the first condition of Theorem 6.9 holds, choose $y(\omega) = -1$ and $m(\omega) = -1$ for all $\omega \in \Omega$. Then $y \in \mathcal{L}^q(\Omega, \mathcal{F}, P)$ and $m \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ since P is a finite measure. Note that

$$\begin{aligned} f(H,u) &\geq \sum_{n=1}^N \Delta S_n H_n && \text{(from the definition of } f) \\ &\geq -u && \text{(from the definition of } f) \\ &= u \cdot y(\omega) && \text{(from the choice of } y) \\ &\geq u \cdot y(\omega) + m(\omega) && \text{(from the choice of } m), \end{aligned}$$

for any $(H,u) \in \mathbb{R}^{N+1} \times \mathbb{R}$.

Hence, the first condition of Theorem 6.9 holds, and therefore, the theorem implies that there is no duality gap (by using the comments after Theorem 2.44, since the perturbation function is convex).

To conclude, assume there is no arbitrage in the market. Then, the optimal primal value is 0 (from the definition of the primal problem (7.3)). Since there is no duality gap, the optimal value of the dual problem (7.6) is 0 as well. Hence, there must exist a feasible dual solution (if not, the dual optimal value would be ∞), i.e. there exists an equivalent martingale measure.

Conversely, assume there exists an equivalent martingale measure. Then, it follows from Delbaen and Schachermayer [9] that there is no arbitrage in the market.

To summarize, the following theorem is true:

Theorem 7.4 *Given the setting above, assume that the set* $A := \{H \in \mathcal{H}_{\mathcal{G}} : f^{\infty}(H(\omega), 0, \omega) \leq 0 \ P\text{-a.s.}\}$ *is a linear space. Then, there is no arbitrage in the market if and only if there exists an equivalent martingale measure.*

This chapter completes our study of how duality methods can be used in mathematical finance. The chapter has shown how generalized Lagrange duality (see Section 5.4) and conjugate duality (see Chapter 2) can be used to solve arbitrage problems in mathematical finance. The next, and final, chapter consists of some concluding remarks.

Chapter 8

Final comments

The purpose of this thesis has been to combine stochastic analysis, functional analysis, measure- and integration theory, real analysis and convex analysis to investigate, and prove, some results in mathematical finance. Background theory, such as convexity, convex analysis, the conjugate duality framework of Rockafellar [34], Lagrange duality and the general stochastic analysis framework for modeling a financial market, has been covered. Then, these results and frameworks have been applied to various central problems in mathematical finance, for instance convex risk measures, utility maximization, pricing, and arbitrage problems. I believe the results are interesting, and that they indicate the potential of exploiting duality in mathematical finance, and more generally, in stochastic control problems.

A central technique of this thesis can be roughly summarized as follows:

- Given a problem in mathematical finance, formulate the problem as an optimization problem in a suitable space. Typically, it will be a constrained stochastic optimization problem. Call this the primal problem. Try to solve the primal problem by standard optimization techniques.
- If one is unable to solve this primal problem as it is: Find a dual to the problem. The optimal value of the dual problem will provide bounds on the optimal value of the primal problem. Try to solve this dual problem. If one is unable to solve this dual problem as well, it may help to transform either the primal or the dual problem (by basic algebra, exploiting that some inequalities must hold with equality in the optimum etc.).
- Try to show that there is no duality gap, i.e. that the optimal primal and dual values coincide. There are hopes of this if one is working with a convex objective function and convex constraints.

However, as this thesis has shown, there are many types of duality and some are simpler to handle than others. For example, linear programming is less complicated than the conjugate duality framework. So how can one know which duality method to try?

- If the time is finite and discrete, the scenario space is finite, and the objective function as well as the constraints are linear: Linear programming (LP) can be applied. There will not be a duality gap from the linear programming duality theorem. Note that sometimes it may be easier to find the Lagrange dual problem than finding the LP dual (LP is a special case of Lagrange duality).

- If there are finitely many constraints (even though the scenario space is arbitrary and time is continuous) Lagrange duality can be applied. If the objective function and constraint functions are convex, the Slater condition can be used to show that there is no duality gap. If the constraints are linear, this boils down to checking whether there exists a feasible solution to the primal problem.

- If there are arbitrarily many constraints (for instance one constraint for each $\omega \in \Omega$), the conjugate duality framework may work. There are several ways to define the perturbation space and the perturbation function, and hence there are several different dual problems. It may be necessary to try different perturbation spaces.

  If the objective function and constraints are convex, there is hope that there is no duality gap. In order to show this, one must essentially show that the optimal value function φ is lower semi-continuous. However, there are several other theorems guaranteeing this, for instance the generalized Slater condition of Example 4" in Rockafellar and Theorem 9 (Theorem 6.9) in Pennanen and Perkkiö [30].

Though closing the duality gap is a desirable result in many examples, it is not always possible. However, all is not lost. Even if one cannot close the duality gap, the dual problem gives bounds on the optimal primal value, and an iterative method where one computes primal and dual solutions every other time will give an interval where the optimal primal (and dual) value(s) must be. Also, note the elegance and step-by-step approach that duality methods bring to many well-known problems in mathematical finance. Another advantage regarding most of the duality approaches of this thesis is that one does not have to make a lot of assumptions on the market structure.

However, even though duality methods provide a step-by-step approach, many difficulties can occur. A lot of stochastic and functional analysis is necessary in order to study the functions that arise naturally from duality. When working with conjugate duality, it may be difficult to choose appropriate paired spaces, and an appropriate perturbation function, in order to attain useful properties assuring, for instance, no duality gap. Sometimes, the primal and dual problems have to be transformed in order to apply duality methods, or perhaps in order to get results that can be interpreted. These transformations may require a lot of stochastic analysis, measure theory, and clever observations.

Many results in this thesis, in particular the results in the final chapters, have been proved in discrete time. The reason for this is that this makes it simpler to apply the conjugate duality framework of Rockafellar [34]. However, it may be possible to generalize the results presented here to continuous time, possibly using a discrete time approximation. However, this is beyond the grasp of this thesis, and open for further research.

I hope this thesis has enlightened, and structured, the vast and (at least to me in the beginning) rather confusing topic of duality methods in mathematical finance. A topic which I feel has a lot of potential still to be exploited.

Appendix A

Matlab programs

This appendix consists of two Matlab programs connected to Chapter 5. The first program checks whether a financial market is complete, and the second program uses Theorem 5.8 to compute the optimal trading strategy for an investor with ln-utility in a given complete financial market. See Chapter 5 for further comments.

```
clear all

%PROGRAM FOR DECIDING WHETHER A FINANCIAL MARKET IS COMPLETE.

%Assume the market is normalized. All asset prices have been
%divided through by the bank.

%General:
N=4; %Number of scenarios \omega in \Omega
T=2; %Number of time steps
M=1; %Number of risky assets

%Information structure:
%List all scenarios and give them same numeral value as order
omega=zeros(N,1);
for n=1:N
    omega(n)=n;
end

%Write partition at different times, represent \omega_n by n
%Don't need to write first and last partition since assumed
%to be trivial and biggest sigma alg.

%Notation: Fi=time i partition

%t=1:
```

```
%NB:these are cell arrays
F={{[1 2 3 4]},{[1 2],[3 4]},{[1],[2],[3],[4]}}; %F represents
%partitions. Add more partitions with comma and
%another curly bracket.

%Prices:
Price={4,[5;2], [7;3;4;1]}; %Prices risky asset: time 0,
%time 1 and time 2. If you have more than one risky asset:
%Add more rows with semi colon to separate.

%Want to check if there exists measure such that the price
%process is a martingale. Need to construct correct matrix:

Antallrader = 0;

for t=1:T
    Antallrader=Antallrader + size(F{t},2);
end

Antallrader=Antallrader + 1; %Need one extra row for
%probability measure condition.

A=zeros(Antallrader,N);


if T>1
    for t = 2:T %for each time
        for m = 1:size(F{t},2) %check how many rows you need
            for k = F{t}{m}(1):F{t}{m}(size(F{t}{m},2))
            %and which columns to use
                A(m,k) = Price{t+1}(k) - Price{t}(m);
            end
        end
    end
else
    T=1; %Don't need to do anything if T=1.
end


%Time 1 condition:
for i=1:size(F{2},2)
    for j=F{2}{i}(1):F{2}{i}(size(F{2}{i},2))
        A(Antallrader - 1,j)= Price{2}(i);
    end
end


A(Antallrader, :) = ones(1,N); %Probability measure condition
```

```
%Make right hand side vector:

b=zeros(Antallrader,1);
b(Antallrader)= 1;
b(Antallrader - 1) =Price{1};

%Solve equation for Q:
Total=[A b];

%Have to read from the row reduced form whether market is
%complete or not.
RowReducedQMatrix=rref(Total)
```

```
clear all

%PROGRAM FOR MAXIMIZING EXPECTED UTILITY OF TERMINAL WEALTH
%AND COMPUTING CORR. OPTIMAL TRADING STRATEGIES IN A
%COMPLETE MARKET.


%Assume marked is  complete and normalized.
%Assume F_0 is trivial, F_T is the biggest sigma-alg.
%possible. Assume investor has utility function
%U(x)=ln(x), and wants to maximize expected utility
%of terminal wealth.


%INPUT

%General:
N=4; %Number of scenarios \omega in \Omega
T=2; %Number of time steps
M=1; %Number of risky assets

%Information structure:
%List all scenarios and give them same numeral value as order
omega=zeros(N,1);
for n=1:N
    omega(n)=n;
end

%Write partition at different times, represent \omega_n by n
%Don't need to write first and last partition since assumed
%to be trivial and biggest sigma alg.

%Notation: Fi=time i partition

%t=1:
```

```
%NB:these are cell arrays
F={{[1 2],[3 4]}}; %F represents partitions. Add more
%partitions with comma and another curly bracket.

%Prices:
Price={4,[5;2], [7;3;4;1]}; %Prices risky asset: time 0,
%time 1 and time 2. If you have more than one risky asset:
%Add more rows with semi colon to separate.

x=10; %Initial endowment of investor

%Probabilities:
P=[1/5; 1/5; 2/5; 1/5]; %Given probability measure
Q=[1/3;1/3;1/9;2/9]; %Risk neutral probability measure
%NOTE: Can compute this in program. Do check if complete.

%I(y)=1/y; Define previously computed func. from ln-utility

%COMPUTATIONS

dQdP=Q./P; %Find Radon Nikodym derivative by componentwise
%division of verctors.

yStar=1/x; %From previous computation for ln-utility.
%Change this if agent has another utility function.

%Use Thm. in Schachermayer's "Utility maximization in an
%incomplete market" to compute final optimal value.

FinalOptValue=ones(N,1)./(yStar*dQdP); %Compute final optimal
%value from formula in Thm.

%Print output
for n=1:N
    fprintf('The terminal value of the optimal portfolio is
     %g.\n in state omega %g.\n',FinalOptValue(n),n)
end


%Compute things needed for finding optimal trading strategy

if T==1
    PriceTime1=[Price{2} ones(size(Price{2},2),1)];
    TradeTime1=PriceTime1\FinalOptValue;
    fprintf(' In the following trading strategy vector the
    first component is the bank, the second risky asset 1
    and so on.')
    fprintf('The optimal trading strategy at time 0 is:')
    disp(TradeTime1)
```

```
else

    %Make list of price matrices

    for t=1:T
        A{t} = [Price{1,t+1}];
        for m=2:M
             A{t}=[A{t} Price{m,t+1}];
        end
        A{t}=[A{t} ones(size(Price{1,t+1}))];
    end
    %A is now the list consisting of all the price matrises
    %for the different times.

    %For all times we need to consider the optimal trading
    %strategy given the information at that time. Hence,
    %we must divide up our price matrices, in order to get
    %the correct computations done.

    %t gives the time you're in. k gives the element of the
    %partition at time t you're in.
    B{1,1}=A{1};
    for t=2:T %Don't need to divide up for first since first
    %sigma alg. is trivial.
        for k=1:size(F{t-1},2)
            B{t,k}= A{t}((F{t-1}{k}(1)):F{t-1}{k}...
            (size(F{t-1}{k},2)),:);
        end
    end
    %B{t,k} is the matrix corresponding to the trading
    %strategies you should choose at time t-1 given the
    %information at that time.


    %Need to split up final value vector as well, depending
    %on \Omega:
    for k=1:size(F{T-1},2)
        c{T,k}=FinalOptValue(F{T-1}{k}(1):F{T-1}{k}...
        (size(F{T-1}{k},2)));
        %The c's give the valueprocess
    end


    %Compute optimal trading strategy:Compute backwards
    for l=1:T
        t=(T+1)-l; %The indices become a little messy because
        %of backwards
        if t==1
            c1=[c{1,1};c{1,2}];
            b{t,1}=B{t,1}\c1;
```

```
            break
        else
            for k=1:size(F{t-1},2)
                b{t,k}=B{t,k}\c{t,k}; %b{t,k} is optimal
                %trading strategy to choose from time t-1
                %to t when you are in part k of the partition.
                %Need to compute optimal value of process one
                %step back.
                L{t-1,k}=A{t-1}(k,:);
                c{t-1,k}=L{t-1,k}*b{t,k};
            end
        end
    end

    %Note that in the calculations above we compute the
    %self-financing trading strategy that gives us the
    %optimal portfolio.


    %Make nice print

    fprintf(' \n \n In the following trading strategy
    vectors the first component is the bank, \n the
    second risky asset 1 and so on. \n')

    for t=1:T
        if t==1
            fprintf('The optimal trading strategy at
            time 0 is \n \n')
            disp(b{1,1})
        else
            for k=1:size(F{t-1},2)
                fprintf('The optimal trading strategy at
                time %g if you are in states omega \n \n',...
                t-1)
                disp(F{t-1}{k}')
                fprintf('is \n')
                disp(b{t,k})
            end
        end
    end

end
```

Bibliography

[1] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9.

[2] Dimitri Bertsekas, Angelia Nedic, and Asu Ozdaglar. *Convex analysis and optimization*. Athena scientific, Belmont MA, 2003.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, Cambrigde, 2004.

[4] Geir Dahl. An introduction to convexity. 2000.

[5] Robert Dalang, Andrew Morton, and Walter Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics and stochastic reports*, 29:185–201, 1990.

[6] Mark Davis. Mathematical option pricing. 2006.

[7] Freddy Delbaen. The dalang-morton-willinger theorem. 1999.

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

[9] Freddy Delbaen and Walter Schachermayer. *The mathematics of arbitrage*. Springer, Berlin, 2006.

[10] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems*. SIAM, 1999.

[11] Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and stochastics*, 6:429–447, 2002.

[12] Hans Föllmer and Alexander Schied. Robust preferences and convex measures of risk. In *Advances in finance and stochastics*, pages 39–56. Springer, Berlin, 2002.

[13] Hans Föllmer and Alexander Schied. Convex and coherent risk measures. 2008.

[14] Marco Frittelli and Emanuela Gianin. Putting order in risk measures. *Journal of banking and finance*, 26:1473–1486, 2002.

[15] Alena Henclova. Notes on free lunch in the limit and pricing by conjugate duality. *Kybernetika*, 42, 2006.

[16] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex analysis and minimization algoritms I.* Springer, Berlin, 1993.

[17] Ioannis Karatzas and Steven Shreve. *Methods of mathematical finance.* Springer, New York, 1998.

[18] Alan King. Duality and martingales: a stochastic programming perspective on contingent claims. *Mathematical programming*, 91:543–562, 2002.

[19] Alan King and Lisa Korf. Martingale pricing measure in incomplete markets via stochastic programming duality in the dual of $\mathcal{L}^{\infty}$. 2001.

[20] Hidetoshi Komiya. Elementary proof for sion's minimax theorem. *Kodai mathematical journal*, 11:5–7, 1988.

[21] Dmitry Kramkov and Walter Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annuals of applied probability*, 9:904–950, 1999.

[22] Dmitry Kramkov and Walter Schachermayer. Necessary and sufficient conditions in the problem of optimal investment in incomplete markets. *Annuals of applied probability*, 13:1504–1516, 2003.

[23] Tom Lindstrøm. *Kalkulus.* Universitetsforlaget, Oslo, 2006.

[24] David Luenberger. Convex programming and duality in normed space. *IEEE transactions and systems science and cybernetics*, 1968.

[25] Hans-Jakob Luthi and Jörg Doege. Convex risk measures for portfolio optimization and concepts of flexibility. *Mathematical programming*, 104:541–559, 2005.

[26] James Munkres. *Topology.* Prentice hall, Upper Saddle River, 2000.

[27] Bernt Øksendal. *Stochastic differential equations.* Springer, Berlin, 2003.

[28] Gert Pedersen. *Analysis now.* Springer, New York, 2002.

[29] Teemu Pennanen. Convex duality and stochastic programming in mathematical finance. *Mathematics of operations research*, 36:340–362, 2011.

[30] Teemu Pennanen and Ari-Pekka Perkkiö. Stochastic programs without duality gaps (to appear). 2012.

[31] Huyên Pham. *Continuous time stochastic control and optimization with financial applications*. Springer, Berlin, 2009.

[32] Stanley Pliska. *Introduction to mathematical finance*. Blackwell publishing, Malden, 2006.

[33] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, Princeton, 1972.

[34] Ralph Tyrell Rockafellar. *Conjugate duality and optimization*. SIAM, Philadelphia, 1974.

[35] Ralph Tyrell Rockafellar. Coherent approaches to risk in optimization under uncertainty. *INFORMS*, 2007.

[36] Ralph Tyrell Rockafellar and Roger Wets. *Variational analysis*. Springer, Berlin, 2004.

[37] Leonard Christopher Gorden Rogers. Duality in constrained optimal investment and consupmtion problems: a synthesis. 2001.

[38] Andrzej Ruszczynski and Alexander Shapiro. Optimization of convex risk functions. *Mathematics of operations research*, 31:433–452, 2006.

[39] Bryan Rynne and Martin Youngston. *Linear functional analysis*. Springer, London, 2008.

[40] Walter Schachermayer. Utility maximization in incomplete markets. 2004.

[41] Renè Shilling. *Measures, integrals and martingales*. Cambridge university press, Cambridge, 2005.

[42] Robert Vanderbei. *Linear programming: foundations and extensions*. Springer, Berlin, 2008.

[43] Günter Ziegler. *Lectures on polytopes*. Springer, Heidelberg, 1995.

Index