NO FREE LUNCH AND RISK MEASURES ON ORLICZ SPACES

by

Theresa Maria Offwood

under the supervision of

Coenraad C.A. Labuschagne

Programme in Advanced Mathematics of Finance

School of Computational and Applied Mathematics

University of the Witwatersrand,

Private Bag 3, Wits, 2050

South Africa

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy.

February 2012

Abstract

The importance of Orlicz spaces in the study of mathematics of finance came to the for in the 2000's when Frittelli and his collaborators connected the theory of utility functions to Orlicz spaces. In this thesis, we look at how Orlicz spaces play a role in financial mathematics. After giving an overview of scalar-valued Orlicz spaces, we look at the first fundamental theorem of asset pricing in an Orlicz space setting. We then give a brief summary of scalar risk measures, followed by the representation result for convex risk measures on Orlicz hearts. As an example of a risk measure, we take a detailed look at the Wang transform both as a pricing mechanism and as a risk measure. As the theory of financial mathematics is moving towards the set-valued setting, we give a description of vector-valued Orlicz hearts and their duals using tensor products. Lastly, we look at set-valued risk measures on Orlicz hearts, proving a robust representation theorem via a tensor product approach.

Acknowledgements

I would like to thank the following people, who made it possible for me to complete this thesis:

- My supervisor, Coenraad Labuschagne, for sharing his ideas and his knowledge and helping me make this a thesis I am proud of.
- My mother, Beni Offwood, for her unconditional support, love and encouragement.
- My husband, Jacques le Roux, for his endless patience, love and support.
- My sister, Sonja Offwood, for helpful discussions and for reading my thesis.
- Olivier Menokou Pamen, Stuart Cullender, Koos Grobler and Marco Frittelli for their helpful comments and suggestions.

Declaration

I hereby declare that this thesis was independently written by me. No material was used other than that referred to. Sources directly quoted and ideas used, including figures, tables and sketches, have been correctly denoted. Those not otherwise indicated belong to the author. This thesis is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

Signed: ____________________ Date: __________

Contents

Table of Contents **v**

1 Introduction **1**

2 Scalar-valued Orlicz spaces **6**
2.1 Definition . . . 7
2.2 Comparing Orlicz spaces with L_p-spaces . . . 10
2.3 The Δ_2-condition . . . 12
2.4 The Orlicz norm . . . 12
2.5 Duality for Orlicz spaces . . . 14
2.6 Optimisation in Orlicz spaces . . . 17
2.7 Conditional expectations on Orlicz spaces . . . 18

3 No free lunch in Orlicz spaces **20**
3.1 No free lunch . . . 21
3.2 Equivalent martingale measures on Orlicz hearts . . . 25
3.2.1 Yan's theorem in Banach lattices . . . 25
3.2.2 Equivalent martingale measures . . . 28
3.3 Utility functions . . . 31
3.4 No market free lunch . . . 36
3.5 No free lunch in Orlicz spaces . . . 40
3.5.1 The optimisation problem . . . 40
3.5.2 The Orlicz space associated with u . . . 45
3.5.3 The minimax theorem . . . 49
3.5.4 No market free lunch* . . . 57

4 Scalar-valued risk measures on Orlicz hearts **60**
4.1 Definition . . . 61
4.2 Acceptance sets . . . 64
4.3 Risk measures on Orlicz hearts . . . 69
4.3.1 Monotone functionals on Banach lattices . . . 69
4.3.2 Dual Representation . . . 71
4.4 Examples of risk measures . . . 80

5 Wang transform as a risk measure **84**
5.1 Comonotonicity and the Choquet integral . . . 85
5.1.1 Comonotonicity . . . 85
5.1.2 The Choquet integral . . . 89
5.1.3 Integral representation theorems . . . 91
5.2 Choquet pricing . . . 92
5.3 Stochastic differential equations . . . 95
5.4 The Wang transform . . . 97
5.4.1 Pricing financial risk . . . 103
5.4.2 Pricing exotic options . . . 107
5.5 The Esscher-Girsanov transform vs. the Wang transform . . . 113
5.5.1 The Esscher transform . . . 114
5.5.2 The Esscher-Girsanov transform . . . 116
5.5.3 The Esscher-Girsanov transform vs. the Wang transform . . . 117
5.5.4 Discussion . . . 120
5.6 Comonotonic convex risk measures . . . 120
5.7 The Wang transform as a risk measure . . . 123

6 Vector-valued Orlicz spaces **125**
6.1 Connecting $H_\Phi(P,Y)$ to $H_\Phi(P)\widetilde{\otimes}_l Y$ . . . 126
6.2 $(H_\Phi(P,Y))^*$ . . . 129
6.3 Associativity of the l-norm . . . 132
6.4 The Radon-Nikodým property in $H_\Phi(P,Y)$ . . . 137
6.5 Martingale convergence in $H_\Phi(P,Y)$ . . . 138

7 Set-valued risk measures on Orlicz hearts **141**
7.1 Set-valued setting . . . 142
7.2 Duality in tensor products . . . 154
7.3 Risk measures on the l-tensor product . . . 159

A Appendix: Functional analysis background **169**
A.1 Riesz spaces . . . 169
A.2 Order duality . . . 172
A.3 Banach lattices . . . 174
A.4 Norm duality . . . 175
A.5 Semi-M-spaces . . . 176
A.6 Measure Theory . . . 178
A.7 L_p-spaces . . . 181
A.8 Martingales . . . 183

B Appendix: Tensor products on Banach spaces **186**
B.1 Definition . . . 186
B.2 Norms on tensor products . . . 188
B.3 The Bochner space $L_p(P,Y)$ as a tensor product . . . 189

C Appendix: Convex analysis **191**

Chapter 1

Introduction

Mathematical finance is a subset of applied mathematics, that attempts to understand and model the financial markets. Financial mathematics uses highly technical and abstract branches of mathematics to explain very practical applications that affect peoples' everyday lives. The first to enter into this field was Louis Bachelier [9], with a discussion of the use of Brownian motion to evaluate stock options. However, at the time, this contribution was only of interest in academia. One of the first influential works of mathematical finance was by Harry Markowitz, who looked at the theory of portfolio optimisation. Simultaneously, William Sharpe developed the theory to determine the correlation between the stocks and the market. For their pioneering work, Markowitz and Sharpe, along with Merton Miller, shared the 1990 Nobel Memorial Prize in Economic Sciences - the first time awarded for work in finance. Robert Merton and Paul Samuelson extended the then existing theory by replacing one-period models by continuous time, Brownian-motion models, and the quadratic utility function by a general increasing, concave utility function. In 1973, Fischer Black and Myron Scholes published their results on the modeling of financial markets with stochastic models. They changed the way mathematics of finance was viewed and since then, research in mathematics of finance has expanded significantly.

A large part of mathematics of finance is based on two theorems, called the fundamental theorems of asset pricing. This name was first coined by P.H. Dybvig and S. Ross in 1987 [53]. These theorems link the concepts of no arbitrage and market completeness to the theory of martingales.

The first fundamental theorem of asset pricing relates the principle of no arbitrage to the existence of an equivalent martingale measure. The no arbitrage principle takes on various forms, depending on, amongst other things, the completeness of the market and whether the underlying asset is bounded or not. The definition of no arbitrage is relatively weak and does not always imply the existence of an equivalent martingale measure. Hence, Kreps [107] introduced the concept of no free lunch, followed by Delbaen and Schachermayer's notion of no free lunch with vanishing risk [39]. These concepts allow for the first fundamental theorem of asset pricing to hold in a more general setting. The type of equivalent martingale measure (e.g. local martingale, sigma-martingale, ...), that the first fundamental theorem shows to exist, depends on the finiteness of the probability space, the boundedness and the continuity of the price process and whether time is considered to be discrete or continuous.

The existence of an equivalent martingale measure in the market model is important, as it is required to price derivatives. The price of a derivative is the discounted expected value of the future payoff of the derivative under some risk-neutral measure.

The second fundamental theorem of asset pricing connects market completeness to the uniqueness of the equivalent martingale measure. This theorem was first introduced by Harrison and Kreps [81] for finite probability spaces. Due to an example by Artzner and Heath [7], there was doubt whether the second fundamental theorem holds in general. However, Jarrow, Jin and Madan [88] showed that modifying the definition of completeness allows for the second fundamental theorem of asset pricing to hold in general.

In complete markets, the set of equivalent martingale measures consist of a unique element and hence, it is possible to get a perfect hedge that eliminates all the risk. Each contingent claim has a unique, preference-independent price, which is consistent with no arbitrage. In the real world, however, market frictions, for example transaction costs, non-traded assets and portfolio constraints, create an incomplete market, where perfect replication is not possible. Therefore, there are infinitely many equivalent martingale measures, resulting in a range of prices for each contingent claim, which are all consistent with no arbitrage.

The problem then, is to determine a criterion for selecting one of the equivalent martingale measures. There are various ways to tackle this: the super replication method, the use of a minimal distance criterion, convex risk measures or the maximisation of expected utility.

In incomplete markets, there are claims that have an intrinsic component of risk, which cannot be hedged. Therefore, it makes sense to look at the preference structure of the investors in the market and introduce a description of the investor in terms of his utility function. To use the utility maximisation approach, it is necessary to solve the stochastic control problem of finding the optimal investment. There are two methods to find the solution to the stochastic control problem. The first was developed by Merton [124], which reduces the problem to a partial differential equation (PDE) and solves it numerically. However, these PDE's are generally intractable if not virtually impossible to solve. The theory was extended by Davis [34, 35], Hendersen and Hobson [85] and others. The second method involves duality theory and stochastic calculus. It was first developed by Pliska [131] in 1984 and has since been improved and generalised by numerous others, most notably by Karatzas et al. [96] and Delbaen and Schachermayer [39, 40, 141].

Now, that a preference structure has been included in the valuation procedure, it is only natural to look at including it in other concepts of mathematical finance. To make this possible, it is necessary to look at a different mathematical setting. Instead of using the popular L_p-space setting, where $0 \leq p \leq \infty$, as has been done up until 2005, Biagini and Frittelli [15] consider a generalisation of the L_p-space: an Orlicz space, named after the Polish mathematician Wladyslaw Orlicz [128]. The utility function is used to generate the Orlicz space, and this space with its dual is then used to define a preference-dependent concept of no arbitrage and to prove a robust representation of convex risk measures.

Frittelli [64] uses the theory of Orlicz spaces to extend the concept of no free lunch to the Orlicz space setting, introducing a new notion called no market free lunch*. He, then, proves a new version of the first fundamental theorem of asset pricing, which states that no market free lunch*, defined on the Orlicz space, is equivalent to the existence of a sigma-martingale measure. For the proof of this theorem, it is necessary to adapt the minimax theorem to an Orlicz space setting. The minimax theorem plays an important role in the convex analysis of optimisation problems. In the financial setting it transforms the original expected utility maximisation problem into its dual form, which in some cases is easier to solve.

Now that we are looking at no arbitrage in an Orlicz space, it makes sense to look at risk measures in this setting as well. It has always been important to look at and understand the risks investors undertake when taking on certain financial positions. There are two main classifications of monetary risk measures: static risk measures and dynamic risk measures, the difference being that dynamic risk measures are time-dependent. In the 1960's Markowitz introduced the notion of using variance and correlation coefficients to quantify risk. As this is only accurate for elliptical distributions, Artzner et al. [6] presented an axiomatic approach to define risk measures. Their risk measures, called coherent risk measures, have to be monotonic, translation invariant, positively homogeneous and subadditive. Since risk does not necessarily increase in a linear manner, Heath [84] introduced convex risk measures in a finite probability space, which was later independently generalised by Föllmer and Schied [60] and Frittelli and Rosazza Gianin [62] to all probability spaces. Instead of the positive homogeneity and the subadditivity conditions, they require the risk measure to be convex. Coherent risk measures are thus, special cases of convex risk measures.

Cheridito and Li [26] looked at risk measures on Orlicz hearts, proving a robust representation for these risk measures. Their representation theorem is similar in form to those obtained by Artzner et al. [6] and Föllmer and Schied [60], amongst others. Biagini and Frittelli [16] proved a representation result for risk measures on Frechet lattices, which generalises Cheridito and Li's results. Orlicz spaces are examples of Frechet lattices.

In most of the literature, the risky portfolio under consideration is a given real-valued random variable and the risk measure is a map into $\mathbb{R}$. In other words, these risk measures do not consider portfolio aggregation. In reality, however, investors have access to different markets and form multi-asset portfolios. It is not always possible or desirable to transform a multi-dimensional portfolio into a position in one financial market, i.e. the position cannot be described by one real number. The reason for this could be transaction costs, liquidity bounds, fluctuating exchange rates, etc.

Thus, we actually require a risk measure that takes values in $\mathbb{R}^d$ and gives us a value in $\mathbb{R}^m$, where $m \leq d$. These m markets could, for example, be money market accounts in different currencies. In other words, it is necessary to look at risk measures in a set-valued setting.

Set-valued risk measures have gained in popularity over the past few years. Jouini et al. [91] were among the first to introduce the set-valued coherent risk measures. Since then, Hamel et al. [77, 78, 79] extended the approach of Jouini et al. to define set-valued convex risk measures and Konstantinides and Kountzakis [101] used the ideas from Stoica [149] and Jaschke and Küchler [89] to define a risk measure on partially ordered normed linear spaces.

Hamel et al. [77, 78, 79] defined convex set valued risk measures on the space $L_p(P, \mathbb{R}^d)$ of Bochner p-integrable functions with values in $\mathbb{R}^d$. Their method for the case $1 \leq p < \infty$ can be generalised to include spaces $H_\Phi(P, \mathbb{R}^d)$ of $\mathbb{R}^d$-valued Orlicz hearts. In view of the connection between utility functions and real-valued Orlicz spaces, as noted above, this extension to Orlicz spaces may be of interest. We use tensor products of Banach lattices and Banach spaces to generalise the results of Hamel et al. .

The thesis is set up as follows. For detailed information on the required results from functional analysis, tensor product theory and convex analysis, the reader is referred to the appendices. Chapter 2 considers scalar-valued Orlicz spaces. We define Young functions, the Orlicz space and heart and look at two norms on these spaces. We, then, look at the dual theory of and optimisation in Orlicz spaces. In Chapter 3, we look at the different variations of the definition of no arbitrage and the first fundamental theorem of asset pricing. We first provide an overview, which can be found in [115]. Next, equivalent martingale measures in Orlicz spaces are considered, before we look at Frittelli's concept of no market free lunch. Lastly, we look at how these concepts can be explained in an Orlicz space setting. In particular, we look at the minimax theorem and Frittelli's notion of no market free lunch*. Then, in Chapter 4, we move on to risk measures. We give a brief history of the literature and then look at Cheridito and Li's representation theorem of risk measures on Orlicz hearts. We end the chapter with some examples. Chapter 5 extends one of the examples given in the previous chapter. It looks at Choquet pricing and gives as an example, a complete description of the Wang transform, introduced by Wang [160] in 1997. The section on pricing exotic options using the Wang transform is new (see [117]). We, then, compare the Wang transform to the Esscher-Girsanov transform, the proof of which is new (see [113]). To conclude this chapter, we look at comonotonic convex risk measures and how the Wang transform can be used as a risk measure. Vector-valued Orlicz spaces and their duals are then defined and characterised in Chapter 6. Except for the known results, stated for the convenience of the reader, the material in this chapter is new and [114] is based on it. We describe the vector-valued Orlicz heart in terms of a suitable tensor product. We use this description to characterise martingale convergence in Banach space-valued Orlicz spaces and also to describe the Radon-Nikodým property in such spaces. Lastly, in Chapter 7, risk measures on vector-valued Orlicz hearts are considered. Except for the first section in this chapter, almost everything is new (see [116]). We show that the results of Hamel et al. [79] can be obtained via a tensor-product approach. In addition, by the tensor product approach and using a result from Labuschagne and Offwood [114], we get a representation of set-valued convex risk measures on vector-valued Orlicz hearts.

Chapter 2

Scalar-valued Orlicz spaces

Orlicz spaces are named after Wladyslaw Orlicz, who together with Zygmunt William Birnbaum, introduced them (see [18, 128]). The theory of Orlicz spaces was further developed in the 1930's by Orlicz and subsequently by many other mathematicians. The development of the theory and applications of Orlicz spaces was boosted by the works of Zaanen [166] and Krasnoselskii and Rutickii [106]. It is now a mature theory with many important applications (see [134] for more details).

Modern probability theory, which plays an important role in financial mathematics, is generally based in an L_p-space setting. However, L_p-spaces are in some cases not broad enough. To include utility functions into the theory of financial mathematics, one needs to consider a generalisation of L_p-spaces: Orlicz spaces.

The importance of Orlicz spaces in the study of mathematics of finance came into the spotlight in the 2000's, when Frittelli and his collaborators connected the theory of utility functions to Orlicz spaces (see [15, 16, 17, 65]). This will be explained in detail in Chapter 3.

In this chapter, we familiarise the reader with the theory of scalar-valued Orlicz spaces. We first define the Orlicz space and the Orlicz heart. We then look at two norms, the Luxemburg norm and the Orlicz norm, and give some of their properties. Duality and optimisation in Orlicz spaces are also discussed. Most of this chapter is based on [54], [134] and [167].

2.1 Definition

Consider a σ-finite measure space $(\Omega, \mathcal{F}, P)$ and let p and q be such that $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Let $u, v \geq 0$ and consider the function

$$\Phi(u) = \frac{u^p}{p}.$$

The derivative of this function is given by $\phi(u) = u^{p-1}$, the inverse of which is $\phi^{-1}(v) = v^{\frac{1}{p-1}} = v^{q-1}$, which in turn is the derivative of the function $\Psi(v) = \frac{v^q}{q}$. In other words, the properties of the L_p-space and its Banach dual, the L_q-space, are closely related to the functions u^p and v^q.

These properties motivated W.H. Young (1912) to develop the following generalisation. Instead of using $\phi(u) = u^{p-1}$, define a non-decreasing, left-continuous function $\phi(u)$ for $u \geq 0$ such that $\phi(0) = 0$ and $\phi(u) \to \infty$ as $u \to \infty$. Let $\psi(v)$ be the inverse function of $\phi(u)$. Then, define

$$\Phi(u) = \int_0^u \phi(t)\, dt \quad \text{and} \quad \Psi(v) = \int_0^v \psi(t)\, dt,$$

where $u, v \geq 0$. The functions Φ and Ψ are called *Young functions* or sometimes *Orlicz functions*.

Definition 2.1.1. A *Young function* is a function $\Phi : [0, \infty) \to [0, \infty]$ satisfying the following conditions:

(i) $\Phi(0) = 0$,

(ii) Φ is left-continuous, i.e. $\lim_{u \uparrow x} \Phi(u) = \Phi(x)$,

(iii) Φ is non-decreasing,

(iv) Φ is convex and

(v) Φ is non-trivial, i.e. $\Phi(u) > 0$ for some $u > 0$ and $\Phi(u) < \infty$ for some $u > 0$.

The convexity of the Young function Φ implies that Φ is continuous except possibly at a single point, where it jumps to ∞. So condition (ii) is only required at that one point.

For the dual space of an Orlicz space, we require the definition and some properties of complementary Young functions. Let $v = \phi(u)$, where $u \geq 0$, be a non-decreasing, left-continuous, real function such that $\phi(0) = 0$ and ϕ does not vanish identically. Let $u = \psi(v)$ denote the generalised left continuous inverse of ϕ given by

$$\psi(y) = \inf\{x \in (0,\infty) : \phi(x) \geq y\}.$$

This means that

- if ϕ is discontinuous at $u = a$, then $\psi(v) = a$ for $\phi(a^-) < v < \phi(a^+)$,
- if $\phi(u) = c$ for $a < u \leq b$ but $\phi(u) < c$ for $u < a$, then $\psi(c) = a$,
- if $\phi(u)$ has a finite limit l as $u \to \infty$, then $\psi(v) = \infty$ for $v > l$, and
- $\psi(0) = 0$.

The above defined function ψ is non-decreasing and its generalised left-continuous inverse is given by ϕ. The functions, defined by

$$\Phi(u) = \int_0^u \phi(t)\,dt \quad \text{and} \quad \Psi(v) = \int_0^v \psi(t)\,dt \quad \text{for all } u, v \geq 0,$$

are called *complementary* Young functions.

This leads to an important theorem, linking the multiplication of the domain to a sum in the range.

Theorem 2.1.2. *(Young's inequality) Let Φ and Ψ be complementary Young functions. Then for all $u, v \geq 0$, we have*

$$uv \leq \Phi(u) + \Psi(v),$$

where equality is obtained if and only if $v = \phi(u)$ or $u = \psi(v)$.

Corollary 2.1.3. *For all $u, v \geq 0$, we have*

$$\Phi(u) = \max\{uv - \Psi(v) : v \geq 0\},$$
$$\psi(v) = \sup\{uv - \Phi(u) : u \geq 0\},$$

where the sup *may be replaced by* max *if $\psi(v) < \infty$.*

Note that $\frac{\Phi(u)}{u} \to \infty$ as $u \to \infty$ if and only if $\Phi(u) \to \infty$ as $u \to \infty$ which in turn is equivalent to $\Psi(v) < \infty$ for all $v \geq 0$. If $\lim\limits_{u\to\infty} \Phi(u) = l$ is finite, then, as defined before, $\psi(v) = \Psi(v) = \infty$ and we will say that Ψ *jumps*.

Using Young functions, we can define the Young class.

Definition 2.1.4. The *Young class* $Y_\Phi(P) = Y_\Phi(\Omega, \mathcal{F}, P)$ is the set of all P-measurable functions f on Ω for which

$$M_\Phi(f) := \int_\Omega \Phi(|f|)\, dP < \infty.$$

The Young class $Y_\Psi(P)$ is defined similarly. These sets of functions are called *complementary Young classes* and M_Φ is called the *modular* of f for Φ.

The convexity of Φ implies that M_Φ is convex which in turn implies that the Young class $Y_\Phi(P)$ is a convex set.

In general Young classes fail to be vector spaces. Therefore, we need to consider a more general set of functions.

Definition 2.1.5.

(i) The *Orlicz space* for Φ, denoted by $L_\Phi(P) = L_\Phi(\Omega, \mathcal{F}, P)$, is the set of all measurable functions $f : \Omega \to \mathbb{R}$ such that $M_\Phi(\frac{f}{a}) < \infty$ for some $a > 0$. The space $L_\Psi(P)$ is defined similarly.

(ii) The *heart* of the Orlicz space, denoted by $H_\Phi(P) = H_\Phi(\Omega, \mathcal{F}, P)$, is the set of all measurable functions $f : \Omega \to \mathbb{R}$ such that $M_\Phi(\frac{f}{a}) < \infty$ for all $a > 0$. The space $H_\Psi(P)$ is defined similarly.

The spaces $L_\Phi(P)$ and $L_\Psi(P)$ are ideals in $L_0(P)$, as is readily verified, and are therefore Dedekind complete Riesz spaces. The spaces $L_\Phi(P)$ and $L_\Psi(P)$ are called *complementary Orlicz spaces*. Both Orlicz [128] and W.A.J. Luxemburg [120] introduced norms on these spaces. Although these norms are different, they are equivalent. We will first focus on the Luxemburg norm.

Definition 2.1.6. The *Luxemburg norm* of a measurable function $f \in L_\Phi(P)$ is defined by

$$\mathcal{N}_\Phi(f) = \inf\{a > 0 : M_\Phi\Big(\frac{f}{a}\Big) \leq 1\}.$$

Theorem 2.1.7. *The Luxemburg norm is a Riesz norm in $L_\Phi(P)$.*

See [167] for a proof of this theorem.

Note that if $f \in L_\Phi(P)$, then $M_\Phi(\frac{f}{n}) < \infty$ for some integer n. But $\frac{|f|}{n} \to 0$ a.e. as $n \to \infty$, so $\Phi(\frac{|f|}{n}) \to 0$ a.e.. By the dominated convergence theorem, $M_\Phi(\frac{f}{n}) \leq 1$ for some n and $\mathcal{N}_\Phi(f) < \infty$. Conversely, if $\mathcal{N}_\Phi(f) < \infty$, then clearly $f \in L_\Phi(P)$. Hence, another description of $L_\Phi(P)$ is the set of all f with $\mathcal{N}_\Phi(f) < \infty$. This looks very similar to the definition of an L_p-space.

Theorem 2.1.8.

(i) *If* $0 \leq \mathcal{N}_\Phi(f) < \infty$, *then* $M_\Phi(\frac{f}{\mathcal{N}_\Phi(f)}) \leq 1$.

(ii) *If* $\mathcal{N}_\Phi(f) \leq 1$, *then* $M_\Phi(f) \leq \mathcal{N}_\Phi(f) \leq 1$.

(iii) *If* $\mathcal{N}_\Phi(f) > 1$, *then* $M_\Phi(f) \geq \mathcal{N}_\Phi(f) > 1$.

Combining (ii) and (iii), we get that in $L_\Phi(P)$, $\mathcal{N}_\Phi(f) \leq 1$ if and only if $M_\Phi(f) \leq 1$ (see [167]).

Theorem 2.1.9.

(i) *The set* $L_\Phi(P)$ *is a Banach lattice.*

(ii) *The set* $H_\Phi(P)$ *is a closed vector subspace of* $L_\Phi(P)$ *and* $H_\Phi(P) \subseteq Y_\Phi(P) \subseteq L_\Phi(P)$.

(iii) *If* $(f_n) \subseteq L_\Phi(P)$ *is an increasing nonnegative sequence with* $\mathcal{N}_\Phi(f_n) \leq 1$ *for all* n, *then the pointwise limit* $f = \lim f_n$ *belongs to* $L_\Phi(P)$ *and* $\mathcal{N}_\Phi(f) \leq 1$.

$H_\Phi(P)$ has some properties that $L_\Phi(P)$ generally does not have.

Theorem 2.1.10. *Let* Φ *be a finite Young function.*

(i) $H_\Phi(P)$ *has an order continuous norm.*

(ii) $H_\Phi(P)$ *is the closure in* $L_\Phi(P)$ *of the integrable simple functions.*

Moreover, $H_\Phi([0,1])$ *is separable.*

Due to these properties, it is generally easier to work with $H_\Phi(P)$ than $L_\Phi(P)$.

2.2 Comparing Orlicz spaces with L_p-spaces

As mentioned earlier, L_p-spaces are special cases of Orlicz spaces. Let p be given with $1 \leq p < \infty$. Then

$$\Phi_p(x) = \frac{x^p}{p}$$

is a Young function with conjugate Φ_q, where $\frac{1}{p}+\frac{1}{q}=1$. Clearly, the Orlicz space $L_{\Phi_p}(P)$ and its heart $H_{\Phi_p}(P)$ are both equal to $L_p(P)$. The Luxemburg norm, however, is given by

$$\mathcal{N}_{\Phi_p}(f)=\frac{1}{p^{\frac{1}{p}}}\left\|f\right\|_p.$$

The following two propositions show how Orlicz spaces compare to $L_1(P)$ and $L_\infty(P)$.

Proposition 2.2.1. *Let Φ be a Young function. If $\Phi(u)=\infty$ for some u, then $L_\Phi(P)\subseteq L_\infty(P)$ and $H_\Phi(P)=\{0\}$.*

Proof. Suppose $\Phi(u_0)=\infty$ and let $f\in L_\Phi(P)$. Then $M_\Phi(\frac{f}{a})<\infty$ for some $a>0$. This implies that $\frac{|f|}{a}<u_0$ a.e. and hence, $f\in L_\infty(P)$.

If we take $f\in H_\Phi(P)$, then $M_\Phi(\frac{f}{a})<\infty$ for all $a>0$, which implies $|f|\leq au_0$ a.e. for all $a>0$. In particular, $|f|\leq nu_0$ for all $n\in\mathbb{N}$. As $L_0(P)$ is an Archimedean Riesz space, $|f|=0$, i.e. $f=0$. □

Proposition 2.2.2. *If $(\Omega,\mathcal{F},P)$ is a probability space, then*

$$L_\infty(P)\hookrightarrow L_\Phi(P)\hookrightarrow L_1(P),$$

i.e. then there exists constants $c,k\in\mathbb{R}$ such that $\mathcal{N}_\Phi(f)\leq k\left\|f\right\|_\infty$ for all $f\in L_\infty(P)$ and $\left\|f\right\|_1\leq c\mathcal{N}_\Phi(f)$ for all $f\in L_\Phi(P)$.

Proof. Let $f\in L_\infty(P)$. Then there exists $k\in\mathbb{N}$ such that $|f|\leq k\mathbf{1}$ a.e., where $\mathbf{1}(\omega)=1$ a.e. for all $\omega\in\Omega$. Let $l>0$ such that $\Phi(l)<\infty$. Then $\frac{|f|}{k/l}\leq l\mathbf{1}$ a.e. and

$$\int_\Omega\Phi\Big(\frac{|f|}{k/l}\Big)\,dP\leq\int_\Omega\Phi(l)\,dP\leq\Phi(l)P(\Omega)=\Phi(l)<\infty.$$

Hence, $f\in L_\Phi(P)$.

For the second part, choose an $a>0$ with $\Phi(a)>0$ and b with $0<b\leq\Phi(a)$. Consider the graphs of the Orlicz function $y=\Phi(x)$ and the straight line $y=\frac{\Phi(a)}{a}x$. By the convexity of Φ, we have for $u\geq a$ that $\Phi(u)\geq\frac{\Phi(a)}{a}u$. As $b<\Phi(a)$, it follows that $\frac{\Phi(a)}{a}u\geq\frac{b}{a}u$. Hence $\Phi(u)\geq\frac{b}{a}u$. Let $C=\frac{a}{b}+a$. Now, if $f\in L_\Phi(P)$, let $r=\mathcal{N}_\Phi(f)$ and compute

$$\begin{aligned}\frac{1}{r}\int|f|\,dP&=\int_{\{|f|\geq ar\}}\frac{|f|}{r}\,dP+\int_{\{|f|<ar\}}\frac{|f|}{r}\,dP\\&\leq\frac{a}{b}\int\Phi\Big(\frac{|f|}{r}\Big)\,dP+aP(\Omega)\\&\leq\frac{a}{b}+a=C.\end{aligned}$$

That is, $\left\|f\right\|_1\leq Cr=C\mathcal{N}_\Phi(f)$. □

2.3 The Δ_2-condition

In this section, we define the condition that results in $L_\Phi(P) = H_\Phi(P)$.

Definition 2.3.1. The Young function Φ is said to satisfy a Δ_2-*condition* if there exists a constant $M > 0$ such that

$$\Phi(2u) \leq M\Phi(u) \text{ for all } u \geq 0.$$

If there exists a $u_0 > 0$ such that the above holds for all $u \geq u_0$, then Φ is said to satisfy a Δ_2-*condition for large* u.

This Δ_2-condition is a sufficient condition for $Y_\Phi(P)$ to be a vector space.

Theorem 2.3.2. *If Φ satisfies the Δ_2-condition, then $Y_\Phi(P)$ is a vector space. If Ω is of finite measure and Φ satisfies Δ_2-condition for large u, then $Y_\Phi(P)$ is a vector space.*

Theorem 2.3.3. *If Φ satisfies the Δ_2-condition, then $L_\Phi(P) = H_\Phi(P)$.*

Corollary 2.3.4. *Let Φ be a strictly positive Young function satisfying the Δ_2-condition. Then the following statements hold:*

(i) Integrable simple functions are dense in $L_\Phi(P)$.

(ii) If (f_n) is an increasing sequence in $L_\Phi(P)$ and $\sup \mathcal{N}_\Phi(f_n) < \infty$, then (f_n) converges in norm.

(iii) Modular convergence is equivalent to norm convergence in $L_\Phi(P)$, i.e. $\mathcal{N}_\Phi(f - f_n) \to 0$ if and only if $M_\Phi(f - f_n) \to 0$.

Moreover, $L_\Phi([0, 1])$ is separable.

2.4 The Orlicz norm

Proposition 2.4.1. *Let Φ and Ψ be complementary Young functions. If $f \in L_\Phi(P)$ and $g \in L_\Psi(P)$, then the product fg is integrable and*

$$\int |fg| \, dP < 2\mathcal{N}_\Phi(f)\mathcal{N}_\Psi(g).$$

As mentioned before, Luxemburg, as well as Orlicz, introduced norms on the Orlicz space. We will now define the Orlicz norm.

Definition 2.4.2. The *Orlicz norm* of $g \in L_\Psi(P)$ is defined as the norm of the bounded linear functions:

$$\begin{aligned} \|g\|_\Psi &= \sup \Big\{ \Big| \int fg \, dP \Big| : f \in L_\Phi(P), \mathcal{N}_\Phi(f) \le 1 \Big\} \\ &= \sup \Big\{ \Big| \int fg \, dP \Big| : M_\Phi(f) \le 1 \Big\}. \end{aligned}$$

The next proposition shows some simple variants of the definition of the Orlicz norm.

Proposition 2.4.3. *Let $g \in L_\Psi(P)$. Then*

(i)

$$\|g\|_\Psi = \sup \Big\{ \int |fg| \, dP : f \in L_\Phi(P), \ \mathcal{N}_\Phi(f) \le 1 \Big\}.$$

(ii)

$$\|g\|_\Psi = \sup \Big\{ \int |fg| \, dP : f \ \textit{integrable simple}, \ \mathcal{N}_\Phi(f) \le 1 \Big\}.$$

(iii) If Φ is finite, then

$$\|g\|_\Psi = \sup \Big\{ \int |fg| dP : f \in H_\Phi(P), \ \mathcal{N}_\Phi(f) \le 1 \Big\}.$$

The Orlicz space $L_\Phi(P)$ carries two equivalent norms, that will be used in this thesis. The first one is the Luxemburg norm, given by

$$\mathcal{N}_\Phi(f) = \inf\{k : k > 0, \int \Phi\Big(\frac{|f|}{k}\Big) dP \le 1\},$$

and the second one is the Orlicz norm, given by

$$\|f\|_\Phi = \sup\{\int |fg| dP : g \in L_\Psi(P), \mathcal{N}_\Psi(g) \le 1\}.$$

An important relationship between the Orlicz norm and the Luxemburg norm, as well as the relationship between the Orlicz norm and the modular, are shown in the following theorem.

Theorem 2.4.4. *Let $g \in L_\Psi(P)$. Then*

(i) $\mathcal{N}_\Psi(g) \le \|g\|_\Psi \le 2\mathcal{N}_\Psi(g)$, i.e. the Luxemburg norm $\mathcal{N}_\Psi$ and the Orlicz norm $\|\cdot\|_\Psi$ are equivalent.

(ii) $\|g\|_\Psi \le 1 + M_\Psi(g)$.

2.5 Duality for Orlicz spaces

Remark 2.5.1. Since Orlicz spaces are Banach lattices, we know from Theorem A.4.4 that the Banach dual and the order dual coincide, i.e. $(L_\Phi(P))^* = (L_\Phi(P))^\sim$ and therefore so do the bands $(L_\Phi(P))^\sim_c$ and $(L_\Phi(P))^*_c$ of order continuous elements in $(L_\Phi(P))^\sim$ and $(L_\Phi(P))^*$. Added to this, it is known, (see [167]), that in an Orlicz space $L_\Phi(P)$, the spaces of order continuous and σ-order continuous functions coincide, i.e.

$$(L_\Phi(P))^\sim_n = (L_\Phi(P))^\sim_c.$$

Therefore, using Theorem A.2.5, we can decompose the Banach dual of the Orlicz space $L_\Phi(P)$ into the following:

$$(L_\Phi(P))^* = (L_\Phi(P))^*_c \oplus (L_\Phi(P))^*_s.$$

From the theory of Banach function spaces, $(L_\Phi(P))^*_c = (L_\Phi(P))^\sim_c$ is known as the *associate space* of $L_\Phi(P)$ and we will denote it by $L'_\Phi(P)$. The proofs of the following three theorems can be found in [167].

Theorem 2.5.2. *The space $L_\Phi(P)$ and the associate space $L'_\Psi(P)$ of $L_\Psi(P)$ contain the same functions, i.e. $L'_\Psi(P) = (L_\Psi(P))^\sim_c = L_\Phi(P)$. Similarly, the space $L_\Psi(P)$ and the associate space $L'_\Phi(P)$ of $L_\Phi(P)$ contain the same functions.*

Theorem 2.5.3. *If $L_\Phi(P)$ is equipped with the Luxemburg norm, then its associate space is $L_\Psi(P)$ with the Orlicz norm. If $L_\Phi(P)$ is equipped with the Orlicz norm, then its associate space is $L_\Psi(P)$ with the Luxemburg norm. Similar facts hold if Φ and Ψ are interchanged. Hence, the second associate space of $L_\Phi(P)$ with either the Luxemburg or the Orlicz norm is again $L_\Phi(P)$ with the same norm. Similarly for $L_\Psi(P)$.*

Theorem 2.5.4. *Let Φ and Ψ be finite complementary Young functions with Φ continuous and let $(\Omega, \mathcal{F}, P)$ be a σ-finite measure space. Then $(H_\Phi(P))^* = L_\Psi(P)$ and for each $x^* \in (H_\Psi(P))^*$, there is a unique $g_{x^*} \in L_\Phi(P)$ such that*

$$\begin{aligned} x^*(f) &= \int_\Omega f g_{x^*}\, dP, \quad f \in H_\Phi(P), \\ \|x^*\| &= \sup\{|x^*(f)| : \mathcal{N}_\Phi(f) \le 1, f \in H_\Phi(P)\} = \|g_{x^*}\|_\Psi, \\ \|x^*\|' &= \sup\{|x^*(f)| : \|f\|_\Phi \le 1, f \in H_\Phi(P)\} = \mathcal{N}_\Psi(g_{x^*}). \end{aligned}$$

Thus $\|x^*\|$ and $\|x^*\|'$ are equivalent norms for $(H_\Phi(P))^*$.

Note that we can equivalently write the norm of $x^* \in (H_\Phi(P))^*$ as

$$\|x^*\| = \sup\{|x^*(f)| : M_\Phi(f) \leq 1, f \in H_\Phi(P)\},$$

using the remark after Theorem 2.1.8.

Corollary 2.5.5. *Suppose Φ and Ψ are finite complementary Young functions. If Ψ satisfies the Δ_2-condition, then the Banach dual of $L_\Psi(P)$ is $L_\Phi(P)$. If Φ satisfies the Δ_2-condition, then the bidual of $H_\Psi(P)$ is $L_\Psi(P)$. If both Ψ and Φ satisfy the Δ_2-condition, then $L_\Phi(P)$ and $L_\Psi(P)$ are reflexive, i.e. $L_\Phi(P) = (L_\Phi(P))^{**}$ and $L_\Psi(P) = (L_\Psi(P))^{**}$.*

The next theorem is an important result, which shows that Orlicz spaces are in fact semi-M-spaces (see Appendix A.5 for more details on semi-M-spaces).

Theorem 2.5.6. *The Orlicz space $L_\Phi(P)$ is a semi-M-space.*

Proof. [167, Theorem 133.6] As before, the Luxemburg norm and the Orlicz norm of $L_\Phi(P)$ are denoted by $\mathcal{N}_\Phi$ and $\|\cdot\|_\Phi$ respectively. By Theorem 2.4.4, we have $\mathcal{N}_\Phi(f) \leq \|f\|_\Phi$ and $\|f\|_\Phi \leq 1 + M_\Phi(f)$ for every $f \in L_\Phi(P)$. Hence, $\mathcal{N}_\Phi(f) \leq 1 + M_\Phi(f)$. Assume that nonnegative functions u_1 and u_2 are given such that $\mathcal{N}_\Phi(u_1) = \mathcal{N}_\Phi(u_2) = 1$. Thus, by Theorem 2.1.8, we have $M_\Phi(u_1) \leq 1$ and $M_\Phi(u_2) \leq 1$. Let (v_n) be a sequence of functions in $L_\Phi(P)$ such that $u_1 \vee u_2 \geq v_n \downarrow 0$ and let $u = u_1 \vee u_2$. The function u is equal to u_1 on a certain subset A of Ω and equal to u_2 on the complementary subset A^c. Hence,

$$\begin{aligned} M_\Phi(u) &= \int_\Omega \Phi[u(x)]dP \\ &= \int_A \Phi[u_1(x)]dP + \int_{A^c} \Phi[u_2(x)]dP \\ &\leq M_\Phi(u_1) + M_\Phi(u_2) \\ &\leq 2. \end{aligned}$$

Since $u \geq v_n \downarrow 0$, we have by the continuity of Φ that $\Phi[v_n(x)] \downarrow 0$ for almost every $x \in \Omega$ and therefore $M_\Phi(v_n) \downarrow 0$. Following from this, we have that

$$\lim_{n\to\infty} \mathcal{N}_\Phi(v_n) \leq 1 + \lim_{n\to\infty} M_\Phi(v_n) = 1.$$

This shows that $L_\Phi(P)$ is a semi-M-space. $\square$

The next result is a consequence of Theorem A.5.3.

Corollary 2.5.7. *The disjoint complement of the set of all σ-order continuous linear functions $(L_\Phi(P))^*_s$ is an AL-space.*

T. Ando [1] was the first to prove that for an Orlicz space $L_\Phi(P)$, the space $(L_\Phi(P))^*_s$ is an AL-space.

Now that we have classified the dual space of an Orlicz space, we can look at decomposing it. Denote the annihilator of $H_\Phi(P)$ by $(H_\Phi(P))^{\text{annh}}$, which by definition is the set of elements of $(L_\Phi(P))^*$ that vanish on $H_\Phi(P)$, i.e.

$$(H_\Phi(P))^{\text{annh}} = \{z \in (L_\Phi(P))^* : z(f) = 0 \text{ for all } f \in H_\Phi(P)\}.$$

We could not find a proof for the following result in the literature. The proof presented here is due to the author.

Theorem 2.5.8. *If Φ is continuous and finite, then*

$$(L_\Phi(P))^* = (H_\Phi(P))^* \oplus (H_\Phi(P))^{\text{annh}}.$$

Proof. Using Remark 2.5.1, we can decompose $(L_\Phi(P))^*$ into

$$(L_\Phi(P))^* = (L_\Phi(P))^*_c \oplus (L_\Phi(P))^*_s.$$

By Theorem 2.5.2 and Theorem 2.5.4, we have that $(L_\phi(P))^*_c = (H_\Phi(P))^*$.

It remains to show that $(L_\Phi(P))^*_s = (H_\Phi(P))^{\text{annh}}$. Let $\gamma \in (H_\Phi(P))^{\text{annh}}$. By the definition of $(H_\Phi(P))^{\text{annh}}$, we have that $\gamma \in (L_\Phi(P))^*$. Thus, we can write $\gamma = \gamma_c + \gamma_s$, where $\gamma_c \in (L_\Phi(P))^*_c$ and $\gamma_s \in (L_\Phi(P))^*_s$. Then, for all $h \in H_\Phi(P)$, $\gamma(h) = \gamma_c(h) + \gamma_s(h)$, where $\gamma(h) = 0$ by definition and γ_c and γ_s are disjoint. Hence, $\gamma_c = -\gamma_s$ on $H_\Phi(P)$. This implies that $\gamma_s \in (L_\Phi(P))^*_c = (H_\Phi(P))^*$. However, $(L_\Phi(P))^*_c$ and $(L_\Phi(P))^*_s$ are disjoint and so we must have $\gamma_s = 0$ on $H_\Phi(P)$. Hence, $\gamma_c = 0$ on $H_\Phi(P)$. Thus, we can write $\gamma = \gamma_s \in (L_\Phi(P))^*_s$, i.e. $(H_\Phi(P))^{\text{annh}} \subseteq (L_\Phi(P))^*_s$.

Let $g \in (L_\Phi(P))^*_s$. Since $(L_\Phi(P))^*_s$ is a band projection, there exists a projection $P_s : (L_\Phi(P))^* \to (L_\Phi(P))^*_s$. By the surjective property of this projection, there exists $h \in (L_\Phi(P))^*$ such that $P_s(h) = g$. There also exists a band projection $P_c : (L_\Phi(P))^* \to (L_\Phi(P))^*_c = (H_\Phi(P))^*$ such that $h = P_c h + P_s h = P_c h + g$. Hence, if $x \in H_\Phi(P)$, then $h(x) = P_c h(x)$, i.e. $(h - P_c h)(x) = 0$. But $(h - P_c h)(x) = P_s h(x) = g(x) = 0$. Thus $g \in (H_\Phi(P))^{\text{annh}}$ and $(L_\Phi(P))^*_s \subseteq (H_\Phi(P))^{\text{annh}}$. □

Remark 2.5.9. It can be shown, without the use of Theorem A.2.5, that $(H_\Phi(P))^{\text{annh}}$ is norm and order isomorphic to $(L_\Phi(P)/H_\Phi(P))^*$. See [134] for more details.

To conclude this section, we will look at some simple but interesting examples of Orlicz spaces. The following examples also show how the choice of Young function influences where the L_p-spaces fit into the Orlicz space setting.

Example 2.5.10. (These examples are taken from [26]).

1. For $\Phi(x) = x$, we have
$$\Psi(y) = \begin{cases} 0 & \text{for } y \leq 1 \\ \infty & \text{for } y > 1 \end{cases}$$
and
$$H_\Phi(P) = L_\Phi(P) = L_1(P), \qquad \mathcal{N}_\Phi(\cdot) = \|\cdot\|_1, \qquad L_\Psi(P) = L_\infty(P), \qquad \|\cdot\|_\Psi = \|\cdot\|_\infty .$$

2. If $\Phi(x) = x^p$ for $1 < p < \infty$, then $\Psi(y) = p^{1-q}q^{-1}y^q$, and we have
$$H_\Phi(P) = L_\Phi(P) = L_p(P), \qquad \mathcal{N}_\Phi(\cdot) = \|\cdot\|_p, \qquad L_\Psi(P) = L_q(P), \qquad \|\cdot\|_\Psi = \|\cdot\|_q .$$

3. If $\Phi(x) = e^{\lambda x} - 1$ for $\lambda > 0$, then
$$\Psi(y) = \begin{cases} 0 & \text{for } y \leq \lambda \\ \frac{y}{\lambda}\log(\frac{y}{\lambda}) - \frac{y}{\lambda} + 1 & \text{for } y > \lambda \end{cases}$$
and $L_\infty(P) \subseteq H_\Phi(P) \subseteq L_p(P) \subseteq L_\Psi(P) \subseteq L_1(P)$ for all $1 < p < \infty$.

2.6 Optimisation in Orlicz spaces

The Orlicz space duality can be used to solve some optimisation problems. One such problem will be discussed in detail in Chapter 3. For more details regarding convex optimisation, the reader is referred to Appendix C.

Let (Φ, Ψ) be complementary Young functions and consider their Orlicz spaces $L_\Phi(P)$ and $L_\Psi(P)$. If $F : \mathbb{R} \to [-\infty, \infty]$ and $F^* : \mathbb{R} \to [-\infty, \infty]$ are lower semi-continuous, convex and not identically equal to ∞, then
$$K_F(x) = \mathbb{E}[F(x)], \text{ for all } x \in H_\Phi(P)$$
and
$$K_{F^*}(x^*) = \mathbb{E}[F^*(x^*)], \text{ for all } x^* \in L_\Psi(P)$$
define convex functions $K_F : H_\Phi(P) \to [-\infty, \infty]$ and $K_{F^*} : L_\Psi(P) \to [-\infty, \infty]$ respectively. The question that arises is: If F and F^* are conjugates, are K_F and K_{F^*} conjugates? Both Rockafellar [135] and Kozek [103] showed that this is true, Rockafellar solved it for L_p-spaces and Kozek for Orlicz spaces. They both, however, addressed it in more generality. For the details, see the above mentioned references.

Theorem 2.6.1. *Let (Φ, Ψ) be complementary Young functions. Let $F : \mathbb{R} \to [-\infty, \infty]$ and $F^* : \mathbb{R} \to [-\infty, \infty]$ be lower semi-continuous convex functions, which are not identically equal to ∞. If F and F^* are conjugates, then K_F and K_{F^*} are conjugate to each other.*

The next theorem is required to prove the minimax theorem, stated and proved later in the thesis.

Theorem 2.6.2. *Let (Φ, Ψ) be complementary Young functions. Let $F : \mathbb{R} \to [-\infty, \infty]$ and $F^* : \mathbb{R} \to [-\infty, \infty]$ be lower semi-continuous convex functions, which are not identically equal to ∞. If F and F^* are conjugates, then the conjugate of K_F is given by*

$$(K_F)^*(x^*) = K_{F^*}(x_c^*) + \sup\{x_s^*(x) : x \in C\},$$

where $x^* = x_c^* + x_s^*$ is the decomposition of x^* into its σ-order continuous part, x_c^*, and its singular part, x_s^*, and $C = \mathrm{dom}(K_F)$.

2.7 Conditional expectations on Orlicz spaces

To be able to use Orlicz spaces in place of L_p-spaces, we need to define conditional expectations on Orlicz spaces. If $(\Omega, \mathcal{F}, P)$ is a probability space and $\Phi : [0, \infty) \to [0, \infty]$ is a Young function, it was noted earlier that

$$L_\infty(P) \hookrightarrow L_\Phi(P) \hookrightarrow L_1(P).$$

Theorem 2.7.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{F}_1$ a sub σ-algebra of $\mathcal{F}$ and $\Phi : [0, \infty) \to [0, \infty]$ a Young function. Then the restriction of $\mathbb{E}[\,\cdot\,|\mathcal{F}_1] : L_1(\mathcal{F}, P) \to L_1(\mathcal{F}_1, P)$ to $L_\Phi(P)$, again denoted by $\mathbb{E}[\,\cdot\,|\mathcal{F}_1]$, is a map $\mathbb{E}[\,\cdot\,|\mathcal{F}_1] : L_\Phi(\mathcal{F}, P) \to L_\Phi(\mathcal{F}_1, P)$ such that $\|\,\mathbb{E}[\,\cdot\,|\mathcal{F}_1]\,\| = 1$. A similar statement holds if $L_\Phi(P)$ is replaced by $H_\Phi(P)$.*

Proof. Let $f \in L_\Phi(P)$. Consider the convex function $\phi = \Phi \circ |\cdot|$. By Jensen's inequality $\Phi\big(\big|\,\mathbb{E}[f|\mathcal{F}_1]\,\big|\big) = \phi\big(\mathbb{E}[f|\mathcal{F}_1]\big) \leq \mathbb{E}[\phi(f)|\mathcal{F}_1] = \mathbb{E}\big[\Phi(|f|)\big|\mathcal{F}_1\big]$ a.s.. But then

$$M_\Phi\big(\mathbb{E}[f|\mathcal{F}_1]\big) = \int_\Omega \Phi\big(\big|\,\mathbb{E}[f|\mathcal{F}_1]\,\big|\big)dP \leq \int_\Omega \mathbb{E}\big[\Phi(|f|)\big|\mathcal{F}_1\big]dP = M_\Phi(f).$$

Consequently, $\mathbb{E}[f|\mathcal{F}_1] \in L_\Phi(P)$. It also follows from

$$\begin{aligned} \mathcal{N}_\Phi(f) \leq 1 \quad &\Rightarrow \quad M_\Phi(f) \leq \mathcal{N}_\Phi(f) \leq 1 \\ &\Rightarrow \quad M_\Phi\big(\mathbb{E}[f|\mathcal{F}_1]\big) \leq 1 \\ &\Rightarrow \quad \mathcal{N}_\Phi\big(\mathbb{E}[f|\mathcal{F}_1]\big) \leq 1, \end{aligned}$$

that $\mathcal{N}_\Phi(\mathbb{E}[f|\mathcal{F}_1]) \leq \mathcal{N}_\Phi(f)$. The latter implies that $\|\mathbb{E}[\,\cdot\,|\mathcal{F}_1]\| \leq 1$. However, as $\mathbb{E}[\,\cdot\,|\mathcal{F}_1]$ is a projection, we must have that $\|\mathbb{E}[\,\cdot\,|\mathcal{F}_1]\| \geq 1$. Thus $\|\mathbb{E}[\,\cdot\,|\mathcal{F}_1]\| = 1$. $\square$

Chapter 3

No free lunch in Orlicz spaces

The concept of the first fundamental theorem of asset pricing plays a very important role in mathematical finance. There are various versions of this fundamental theorem, but all come down to a similar form: The absence of some form of arbitrage is equivalent to the existence of an 'equivalent martingale measure' for the stochastic processes representing the discounted prices of the financial securities in the market. Section 3.1 gives part of the history of the first fundamental theorem of asset pricing.

In Section 3.2, we take a look at equivalent martingale measures in an Orlicz space setting.

In an incomplete market, the pricing of contingent claims can be done via various techniques, one of which is by considering the preference structure of the investor. In other words, utility functions are used to measure an investor's preference for wealth and how much risk they are willing to undertake to gain more wealth. It, thus, makes sense to look at how this preference structure can be included into the first fundamental theorem. In this regard, Frittelli [63] introduced a concept, which he called 'no market free lunch'. This will be defined in Section 3.4 and we give the relationship between Frittelli's concept and the original definitions of no arbitrage. Frittelli shows that there exists an equivalent separating measure if and only if there is no market free lunch with respect to monotone concave utility functions.

However, no market free lunch is only equivalent to the existence of a separating measure. To find a condition, that includes the preference structure, and is equivalent to the existence of a sigma-martingale measure, Frittelli [64] introduces the concept of 'no market free lunch*'. He then uses Orlicz space theory and utility maximisation to show that it is equivalent to the existence of a sigma-martingale measure. This is discussed in Section 3.5.

3.1 No free lunch

This section is based on [41], see also [115].

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a filtered probability space, where we assume that the filtration satisfies the usual conditions of right continuity and completeness, as defined in Appendix A. The variable T denotes a fixed time horizon, which can take on the value ∞. The $\mathbb{R}^d$-valued càdlàg (see Definition A.6.2) semimartingale $S = (S_t)_{t\in[0,T]}$, where $S_t = \{S_t^1, S_t^2, \ldots, S_t^d\}$, represents the discounted price process of d tradeable assets. Let $\mathbb{P}$ be the set of all probability measures equivalent to P.

The investor has an initial endowment of x and there are no restrictions on the quantities he can buy, sell or short sell. A predictable process $H = (H_t)_{t\in[0,T]}$, where $H_t = \{H_t^1, H_t^2, \ldots, H_t^d\}$, gives the amounts invested in each tradeable asset respectively.

We denote by $\int_0^t H_s \cdot dS_s$ the Itô integral of H_t with respect to S_t.

Definition 3.1.1. An $\mathbb{R}^d$-valued S-integrable predictable process H is called *admissible* if there exists a constant c such that for all $t \in [0,T]$,

$$(H \cdot S)_t = \int_0^t H_s \cdot dS_s \geq c \quad P\text{-a.s.}.$$

We denote the set of all admissible processes by

$$\mathcal{H}^1 = \{H \in \mathfrak{P}(S) : (H \cdot S)_t \geq -c \text{ for all } t \in [0,T] \text{ and for some } c > 0\},$$

where $\mathfrak{P}(S)$ denotes the set of predictable and S-integrable processes.

The financial interpretation of c is a finite credit line, which the investor must respect in his trading. This lower bound on the investors losses, allows $(H \cdot S)_\infty$ to exist and bans so-called doubling strategies, where the losses are not bounded below. This restriction traces back to Harrison and Pliska [82], and is now a standard assumption in the literature. This also implies that any wealth process $X_T = x + \int_0^T H_s \cdot dS_s$ is bounded below.

Set

$$K = \{(H \cdot S)_T : H \in \mathcal{H}^1\}$$

and

$$\begin{aligned} C &= (K - L_0^+(P)) \cap L_\infty(P) \\ &= \{f \in L_\infty(P) : f \leq g \text{ for some } g \in K\}. \end{aligned}$$

The set K represents the cone of all claims that are replicable at zero cost via admissible trading strategies. The set $(K - L_0^+(P))$, defined by

$$(K - L_0^+(P)) = \{f \in L_0(P) : f \leq g \ P\text{-a.s. for some } g \in K\},$$

is the cone of all claims in $L_0(P)$ that can be dominated by a replicable claim. In other words, it is the cone of super-replicable claims and consequently, C is the cone of bounded super-replicable claims. A contingent claim $g \in L_\infty(P)$ is *super-replicable at price* 0 if we can achieve some other contingent claim f, $f \geq g$, with zero net investment by pursuing some predictable trading strategy H. Thus it might be necessary to 'throw away money' to arrive at g (also known as 'free disposal').

The notion of separating measures, which will be defined next, was introduced by Bellini and Frittelli [11].

Definition 3.1.2. A P-absolutely continuous probability measure Q is called a *separating measure* if $K \subseteq L_1(\Omega, \mathcal{F}, Q)$ and $\mathbb{E}_Q[k] \leq 0$ for all $k \in K$. It is called an *equivalent separating measure* if in addition $Q \in \mathbb{P}$. We denote the set of separating measures by $\mathcal{M}$.

The set $\mathcal{M}$ of separating measures can also be written as

$$\mathcal{M} = \{Q << P : \mathbb{E}_Q[w] \leq 0 \text{ for all } w \in C\}.$$

If S is bounded (resp. locally bounded), then $\mathcal{M}$ reduces to

$$\mathcal{M} = \{Q << P : S \text{ is a } Q\text{-martingale (resp. local martingale)}\}.$$

Generally, if $\mathcal{M} \neq \emptyset$, then the set of sigma-martingale measures $\mathbb{M}_\sigma(S) \neq \emptyset$ and the $L_1(P)$-norm closure of $\mathbb{M}_\sigma(S)$ is equal to $\mathcal{M}$. See Appendix A for more details on sigma-martingales.

The ability to price a contingent claim is based on the first fundamental theorem of asset pricing. The first fundamental theorem of asset pricing essentially states that the existence of an equivalent separating measure is equivalent to a properly defined condition of no arbitrage or no free lunch. The different notions of no arbitrage depend on different closures of C.

Definition 3.1.3. A semimartingale S satisfies the condition of *no arbitrage* (NA) if

$$K \cap L_0^+(P) = \{0\},$$

or equivalently

$$C \cap L_\infty^+(P) = \{0\}.$$

The NA property has an obvious interpretation: the terminal payoffs cannot be positive for all admissible trading strategies as this implies that the investor will make a profit with probability 1, i.e. he makes a profit without taking on any risk. This cannot be allowed.

Harrison and Pliska proved that for a finite probability space and a discrete time filtration, we have NA equivalent to the existence of an equivalent martingale measure. Dalang, Morton and Willinger extended this to infinite probability spaces.

In continuous time, the no arbitrage condition is too weak to imply the existence of an equivalent martingale measure. Kreps realised that the purely algebraic notion of no arbitrage has to be complemented with a topological notion. Thus, he introduced the notion of no free lunch.

Definition 3.1.4. A semimartingale S satisfies the condition of *no free lunch* (NFL) if

$$\overline{C}^* \cap L_\infty^+(P) = \{0\},$$

where $\overline{C}^*$ denotes the weak-star closure of C.

The process S admits a free lunch, if there exists a random variable $f \in L_\infty^+(P)\backslash\{0\}$ and a net $(f_\alpha)_{\alpha\in I} = (g_\alpha - h_\alpha)_{\alpha\in I}$ such that $g_\alpha = \int_0^T H_t^\alpha \, dS_t$, for some admissible trading strategy H^α, $h_\alpha \geq 0$ and $(f_\alpha)_{\alpha\in I}$ converges to f in the weak-star topology of $L_\infty(P)$. Economically, this implies that although f itself is not allowed to be of the form $\int_0^T H_t \, dS_t$, for some admissible H (this would be an arbitrage), it is required that f can be approximated by f_α in a suitable topology. In this approximation, people are allowed to 'throw money away', which is represented by the h_α.

Kreps then proved the following theorem. He used a separability assumption in his proof, which is not necessary, as Yan [165] showed. Hence, Delbaen and Schachermayer named the theorem after both these authors.

Theorem 3.1.5 (Kreps-Yan theorem)**.** *A locally bounded stochastic process S satisfies the condition of no free lunch if and only if there exists an equivalent local martingale measure.*

Delbaen and Schachermayer found the economic interpretation of NFL problematic, as there is no control on the maximal loss obtained when using the trading strategy that gives the gain of g_α. They asked themselves the following [41]:

(i) Can we find a requirement being closer to the original notion of NA and such that a version of the fundamental theorem of asset pricing still holds?

(ii) Can the weak* topology be replaced by a finer topology?

(iii) Is it possible to replace the net $(f_\alpha)_{\alpha\in I}$ by a sequence $(f_n)_{n=0}^\infty$?

(iv) Is it really necessary to allow for the 'throwing away of money'?

Added to the above drawbacks, if the semimartingale is not locally bounded, the Kreps-Yan theorem no longer holds. Hence, Delbaen introduced no free lunch with bounded risk, followed shortly by Delbaen and Schachermayer's [39, 40] introduction of the notion of no free lunch with vanishing risk.

Definition 3.1.6. A semimartingale S satisfies the condition of *no free lunch with bounded risk* (NFLBR) if

$$\tilde{C} \cap L_\infty^+(P) = \{0\},$$

where $\tilde{C}$ is the set of all limits of weak-star converging sequences of elements of C.

The process (S_t) satisfies NFLBR if and only if there does not exist a $(0, \infty]$-valued random variable f and a sequence $(f_n) \subseteq K$ such that $f_n \geq -1$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} f_n = f$ P-a.e.. An arbitrageur knows that in any case he can at most lose one unit of money, while as n increases, the net gain f_n becomes pointwise arbitrarily close to f.

Shortly thereafter, Delbaen and Schachermayer introduced no free lunch with vanishing risk.

Definition 3.1.7. A semimartingale S satisfies the condition of *no free lunch with vanishing risk* (NFLVR) if

$$\overline{C} \cap L_\infty^+(P) = \{0\},$$

where $\overline{C}$ denotes the closure of C with respect to the norm topology of $L_\infty(P)$.

In NFLVR, the weak star topology is replaced by the topology of uniform convergence. Now S allows for a free lunch with vanishing risk, if there exists $f \in L_\infty^+(P) \backslash \{0\}$ and sequences $(f_n)_{n=0}^\infty = ((H^n \cdot S)_\infty)_{n=0}^\infty \in K$, where $(H^n)_{n=0}^\infty$ is a sequence of admissible integrands, and $(g_n)_{n=0}^\infty$ satisfying $g_n \leq f_n$, such that

$$\lim_{n\to\infty} \|f - g_n\|_\infty = 0.$$

The term *vanishing risk* is explained by the fact that the negative parts $((f_n)^-)_{n=0}^\infty$ and $((g_n)^-)_{n=0}^\infty$ tend to zero uniformly.

With this new concept of NFLVR and the notion of sigma-martingales, Delbaen and Schachermayer proved the fundamental theorem of asset pricing for the most general semimartingale market model.

Theorem 3.1.8. *For any semimartingale S, the following are equivalent:*

(i) S satisfies NFLVR,

(ii) $\mathbb{M}_\sigma(S) \cap \mathbb{P} \neq \emptyset$.

As $C \subseteq \overline{C} \subseteq \tilde{C} \subseteq \overline{C}^*$, we have that NFL $\Rightarrow$ NFLBR $\Rightarrow$ NFLVR $\Rightarrow$ NA. Surprisingly, we actually have that NFL $\Leftrightarrow$ NFLBR $\Leftrightarrow$ NFLVR. This is due to the following theorem, proved by Kabanov [93].

Theorem 3.1.9. *Under NFLVR,* $C = \overline{C}^*$.

3.2 Equivalent martingale measures on Orlicz hearts

The existence of an equivalent martingale measure is of vital importance in mathematics of finance, as it is required to price financial instruments in a risk neutral setting. In this section, we show the necessary and sufficient conditions to ensure the existence of an equivalent martingale measure on an Orlicz heart.

3.2.1 Yan's theorem in Banach lattices

If E is a Banach lattice and $e \in E^+$, let

$$E_e := \bigcup_{n \in \mathbb{N}} [-ne, ne],$$

where

$$[u, v] := \{x \in E : u \leq x \leq v\} \text{ for all } u, v \in E.$$

Let

$$p_e(x) = \inf\{\lambda > 0 : x \in [-\lambda e, \lambda e]\} \text{ for all } x \in E_e.$$

It is well-known that p_e is an M-norm on E_e. If $\overline{E}_e = E$, then e is called a *quasi-interior point* of E, where $\overline{E}_e$ denotes the norm closure of E_e in E.

The following theorem by Nagel [125, 126] is based on Kakutani's M-space and L-space characterizations (see Kakutani [94, 95]).

Theorem 3.2.1. *Let E be a Banach lattice with order continuous norm and with a quasi-interior point $e \in E^+$. Then, there exists a probability space $(\Omega, \mathcal{F}, P)$ such that*

$$L_\infty(P) \hookrightarrow E \hookrightarrow L_1(P) \text{ and } L_\infty(P) \hookrightarrow E^* \hookrightarrow L_1(P),$$

where $L_\infty(P)$ and E are dense order-ideals in E and $L_1(P)$ respectively, and $L_\infty(P)$ and E^* are dense order-ideals in E^* and $L_1(P)$ respectively, where the duality is given by

$$\langle x, x^* \rangle = \int_\Omega x^* x \, dP.$$

Moreover, $L_\infty(P)$ is the Banach lattice (E_e, p_e) and e is the function $\mathbf{1} : \Omega \to \mathbb{R}$, which is 1 a.s..

The following theorem was proved by Yan for $E = L_1(P)$ [165] and by Ansel and Stricker for $E = L_p(P)$, where $1 \leq p < \infty$ [2]. We generalise it to Banach lattices with order continuous norm and quasi-interior points, by means of Nagel's theorem. The proof is adapted from Ansel and Stricker [2].

Theorem 3.2.2. *Let K be a convex subset of E such that $0 \in K$. The following statements are equivalent.*

(i) There exists $z^* \in E^*$ such that $z^* > 0$ a.s. and

$$\sup_{\xi \in K} z^*(\xi) < \infty.$$

(ii) For all $\eta \in E^+$ with $\eta \neq 0$, there exists $c > 0$ such that $c\eta \notin \overline{K - L_\infty^+(P)}$.

(iii) For all $A \in \mathcal{F}$ with $P(A) > 0$, there exists $c > 0$ such that $c\mathbb{1}_A \notin \overline{K - L_\infty^+(P)}$.

If, in addition K is a cone, then the following statements are equivalent to $(i), (ii)$ and (iii).

(iv) For all $A \in \mathcal{F}$ with $P(A) > 0$, we have $\mathbb{1}_A \notin \overline{K - L_\infty^+(P)}$.

(v) $E^+ \cap (\overline{K - L_\infty^+(P)}) = \{0\}$.

Proof. It is clear that (ii) $\Rightarrow$ (iii).

(iii) $\Rightarrow$ (i): Suppose condition (iii) is satisfied. Consider $A \in \mathcal{F}$ such that $P(A) > 0$. By assumption, there exists a real-number $c > 0$ such that $c\mathbb{1}_A \notin \overline{K - L_\infty^+(P)}$. Since $\overline{K - L_\infty^+(P)}$ is convex, we have by the Hahn Banach theorem that there exists $y^* \in E^*$ such that

$$\sup_{\xi \in K, \eta \in L_\infty^+(P)} y^*(\xi - \eta) < cy^*(\mathbb{1}_A). \tag{3.1}$$

Let $\xi = 0 \in K$ and $\eta = a\mathbb{1}_{\{y^* < 0\}}$, where $a > 0$. Then we get

$$-ay^*(\mathbb{1}_{\{y^*<0\}}) < cy^*(\mathbb{1}_A),$$

i.e.

$$-y^*(\mathbb{1}_{\{y^*<0\}}) < \frac{c}{a} y^*(\mathbb{1}_A).$$

By the Archimedean property of $\mathbb{R}$, we have $y^*(\mathbb{1}_{\{y^*<0\}}) \geq 0$ and therefore $y^* \geq 0$ a.s..

Letting $\eta = 0$, we get

$$\sup_{\xi \in K} y^*(\xi) \le c y^*(\mathbb{1}_A) < \infty.$$

Consequently

$$H := \{x^* \in (E^*)^+ : \sup_{\xi \in K} x^*(\xi) < \infty\}$$

is non-empty, as $y^* \in H$ from our above proof. Let

$$C = \{\{z^* = 0\} : z^* \in H\}.$$

We want to show that if $z_n^* \in H$ for all $n \in \mathbb{N}$ with $\{z_n^* = 0\} \in C$, then $\bigcap_n \{z_n^* = 0\} \in C$.

Let (z_n^*) be any sequence in H. By (3.1), the sequences $(\sup_{\xi \in K} z_n^*(\xi))_{n \in \mathbb{N}}$ and $(||z_n^*||_{E^*})_{n \in \mathbb{N}}$ are bounded. Let

$$c_n = \sup_{\xi \in K} z_n^*(\xi) \quad \text{and} \quad d_n = ||z_n^*||_{E^*}.$$

Define

$$z^* = \sum_n b_n z_n^*,$$

where the b_n's are such that $b_n \ge 0$ for all n, $\sum_n b_n c_n < \infty$ and $\sum_n b_n d_n < \infty$. As $||z^*||_{E^*} < \sum_n b_n d_n < \infty$ and $z^* \ge 0$, we have $z^* \in (E^*)^+$. Since

$$\sup_{\xi \in K} z^*(\xi) \le \sum_n b_n \sup_{\xi \in K} z_n^*(\xi) = \sum_n b_n c_n < \infty,$$

it follows that $z^* \in H$. By the construction of z^*, we have $\{z^* = 0\} = \bigcap_n \{z_n^* = 0\}$.

Hence, there exists $z^* \in H$ such that $P(z^* = 0) = \inf_{c \in C} P(c)$.

We will show that $z^* > 0$ a.s.. Suppose that $P(z^* = 0) > 0$. Let $y^* \in H$ satisfy (3.1) with $A = \{z^* = 0\}$. As shown above, we have $0 < y^*(\mathbb{1}_A) = y^*(\mathbb{1}_{\{z^*=0\}})$ and the random variable $y^* + z^* \in H$ with

$$P(y^* + z^* = 0) = P(z^* = 0) - P(z^* = 0, y^* > 0) < P(z^* = 0),$$

which contradicts the fact that $P(z^* = 0) = \inf_{c \in C} P(c)$. Therefore, $z^* > 0$ a.s., proving (iii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ii): Suppose (ii) is not satisfied. Then there exists $\eta \in E^+$ with $\eta \ne 0$ such that for all $n \in \mathbb{N}$, we have $n\eta \in \overline{K - L_\infty^+(P)}$. Hence, there exists $(z_n) \subseteq K - L_\infty^+(P)$ such that for $n \in \mathbb{N}$

$$||n\eta - z_n||_E \le \frac{1}{n}.$$

Also, there exists $k_n \in K$ and $l_n \in L_\infty^+(P)$ such that $z_n = k_n - l_n$. Let $\delta_n = n\eta - k_n + l_n$, then $||\delta_n||_E \le \frac{1}{n}$.

If z^* is a random variable in E^* such that $z^* > 0$ a.s., then we have

$$\begin{aligned} z^*(k_n) &= nz^*(\eta) + z^*(l_n) - z^*(\delta_n) \\ &\geq nz^*(\eta) - z^*(\delta_n). \end{aligned}$$

By definition, we have

$$||z^*||_{E^*} = \sup\{|z^*(z)| : ||z||_E \leq 1\} \geq |z^*(n\delta_n)| \geq nz^*(\delta_n).$$

Thus,

$$\begin{aligned} z^*(k_n) &\geq nz^*(\eta) - z^*(\delta_n) \\ &\geq nz^*(\eta) - \frac{1}{n}||z^*||_{E^*}. \end{aligned}$$

Hence, $\sup_{\xi \in K} z^*(\xi) = \infty$, showing that condition (i) is not satisfied.

Next, assume that K is a cone.

(v) $\Rightarrow$ (ii): Assume that $E^+ \cap (\overline{K - L_\infty^+(P)}) = \{0\}$. We will prove this by contradiction. Assume there exists $\eta \in E^+$ such that for all $c > 0$, $c\eta \in \overline{K - L_\infty^+(P)}$. This implies that $c\eta \notin E^+$, which is a contradiction as E^+ is a Banach lattice.

(ii) $\Rightarrow$ (v): Assume that (ii) holds. Once again, we will prove this by contradiction. Assume there exists $a \in E^+ \cap (\overline{K - L_\infty^+(P)})$. Then $a \in E^+$, which, by (ii), implies that for all $c > 0$ $ca \notin \overline{K - L_\infty^+(P)}$, and $a \in \overline{K - L_\infty^+(P)}$. However, as K is cone, we have that for all $c > 0$, $ca \in \overline{K - L_\infty^+(P)}$, which is a contradiction.

(iii) $\Leftrightarrow$ (iv): Follows similarly. □

3.2.2 Equivalent martingale measures

In this section, we take a look at an application of Theorem 3.2.2, which will give us a condition for the existence of an equivalent martingale measure. We prove the theorems for Banach lattices. The application of these theorems to Orlicz hearts will be a special case. The proofs are adapted from Stricker [150].

Let H be a predictable, simple process, i.e. $H = \sum_{i=1}^{n-1} \lambda_i \mathbb{1}_{(t_i, t_{i+1}]}$, where $0 \leq t_0 < t_1 < \cdots < t_n \leq 1$ and $\lambda_i = (\lambda_i^1, \ldots, \lambda_i^d) \in \mathbb{R}^d$ is a random $\mathcal{F}_t$-measurable vector. We denote by $(H \cdot X)_t$ the stochastic integral of the predictable process $H = (H^1, \ldots, H^d) \in \mathbb{R}^d$ with respect to the semimartingale $X = (X^1, \ldots, X^d) \in \mathbb{R}^d$. We suppose that $(H \cdot X)_0 = 0$. If H is locally bounded, then it is well known that

$$(H \cdot X)_t = \sum_{j=i}^{d} \int_0^t H_s^j \, dX_s^j.$$

Let X be a $\mathbb{R}^d$-valued, càdlàg, adapted semimartingale with $X_t \in E$ for all $t \in [0,1]$. In this section, we consider the following special case of the convex set K of the previous section. Let

$$K = \{(H \cdot X)_1 : H \text{ is predictable, simple and bounded}\}.$$

By considering this cone K, we may add another equivalent condition to those in Theorem 3.2.2.

Theorem 3.2.3. *The following are equivalent.*

(v) $E^+ \cap (\overline{K - L^+_\infty(P)}) = \{0\}$.

(vi) There exists a probability measure Q equivalent to P and with density $\frac{dQ}{dP} \in E^*$ such that X is a Q-martingale.

Proof. (v) $\Rightarrow$ (vi): Assume that condition (v) holds, then condition (i) of Theorem 3.2.2 also holds. Thus, by Theorem 3.2.2, there exists a random variable $z^* \in (E^*)^+$ such that $\sup\limits_{\xi\in K} z^*(\xi) < \infty$. Hence, $z^*(k) \leq \sup_{\xi\in K} z^*(\xi)$ for all $k \in K$. Since K is a cone, we get for all $n \in \mathbb{N}$ that

$$z^*(k) \leq \frac{1}{n} \sup_{\xi\in K} z^*(\xi).$$

Thus, by the Archimedean property, we have that $z^*(k) = 0$ for all $k \in K$. In other words, $z^*((H \cdot X)_1) = 0$ for all predictable, simple and bounded processes H. Let $dQ = z^* dP$. Then $\frac{dQ}{dP} \in E^*$ and X is a martingale under Q. Thus, we have shown that there exists a probability measure Q equivalent to P and with Radon-Nikodým derivative $\frac{dQ}{dP} \in E^*$ such that X is a Q-martingale.

(vi) $\Rightarrow$ (v): If Q is equivalent to P, the Radon-Nikodým derivative $\frac{dQ}{dP} = z^*$ belongs to E^* and $z^* > 0$. If X is a martingale under Q, then $z^*((H \cdot X)_1) = 0$ for all previsible, simple, bounded H. Thus $z^*(k) = 0$ for all $k \in K$, i.e.

$$\sup_{k\in K} z^*(k) = 0 < \infty.$$

Hence, condition (i) of Theorem 3.2.2 is verified with $z^* = \frac{dQ}{dP}$ and (v) holds. ☐

If X is continuous we can weaken condition (v) of Theorem 3.2.3 by replacing $\overline{K - L^+_\infty(P)}$ with $\overline{K}$, as is shown in the next theorem.

Theorem 3.2.4. *Let X be a continuous, adapted process with values in $\mathbb{R}^d$. Let $X_t \in E$ for all $t \in [0,1]$. Then there exists a probability Q equivalent to P and with Radon-Nikodým derivative $\frac{dQ}{dP} \in E^*$ such that X is a Q-martingale if and only if $E^+ \cap \overline{K} = \{0\}$.*

Proof. $\Rightarrow$: Assume that there exists a probability Q equivalent to P and with Radon-Nikodým derivative $\frac{dQ}{dP} \in E^*$ such that X is a Q-martingale. Then, by Theorem 3.2.3, we have that $E^+ \cap (\overline{K - L^+_\infty(P)}) = \{0\}$. As $\overline{K} \subseteq \overline{K - L^+_\infty(P)}$, we must have that $E^+ \cap \overline{K} = \{0\}$.

$\Leftarrow$: Conversely, suppose that $E^+ \cap \overline{K} = \{0\}$. We show, using a contradiction argument, that $\mathbb{1}_A \notin \overline{K - L^+_\infty(P)}$ is satisfied for all $A \in \mathcal{F}$ with $P(A) > 0$. Assume there exists $A \in \mathcal{F}$ with $P(A) > 0$ such that $\mathbb{1}_A \in \overline{K - L^+_\infty(P)}$. Then we can find a sequence of positive, bounded, random variables (B_n), a sequence of previsible, simple, bounded processes (H_n) and a set $A \in \mathcal{F}$ with $P(A) > 0$, such that $(H_n \cdot X)_1 - B_n$ converges to $\mathbb{1}_A$.

We will construct two sequences (U_n) and (B'_n) such that

- $B'_n \in L^+_\infty(P)$,
- U_n is a previsible, simple, bounded process,
- $((U_n \cdot X)_1)$ is bounded in E and
- $((U_n \cdot X)_1 - B'_n)$ converges to $\mathbb{1}_A$.

To construct these sequences, we introduce the stopping time

$$T_n = \begin{cases} 1 & \text{if } (H_n \cdot X)_t < 1 \text{ for all } t \in [0,1] \\ \inf\{t > 0 : (H_n \cdot X)_t \geq 1\} & \text{otherwise.} \end{cases}$$

Note that $(H_n \cdot X)^-_1 \geq (H_n \cdot X)^-_{T_n} \geq 0$. In fact $(H_n \cdot X)^-_1$ tends to 0 in E. We also have that $0 \leq (H_n \cdot X)^+_{T_n} \leq 1$. Thus, the sequence $(H_n \cdot X)_{T_n} = (H_n \cdot X)^+_{T_n} - (H_n \cdot X)^-_{T_n}$ is bounded in E.

However, the process $\mathbb{1}_{[0,T_n]} H_n$ is not simple.

It is well known that there exists a decreasing sequence of stopping times $(T_m)_{m \in \mathbb{N}}$ with $\lim\limits_{m \to \infty} T_m = T_n$. Let

$$T'_m = \begin{cases} T_m & \text{if } |(H_n \cdot X)_{T_n} - (H_n \cdot X)_{T_m}| \leq 1 \\ 1 & \text{otherwise.} \end{cases}$$

Then, we see that

$$\lim_{m \to \infty} \|(H_n \cdot X)_{T_n} - (H_n \cdot X)_{T'_m}\| = 0. \tag{3.2}$$

We can thus choose m such that $\|(H_n \cdot X)_{T_n} - (H_n \cdot X)_{T'_m}\| \leq \frac{1}{n}$.

Let $U_n = \mathbb{1}_{[0,T'_m]} H_n$ and $B'_n = \mathbb{1}_{\{T_n < 1\} \cap A^c} + \mathbb{1}_{\{T_n = 1\}} B_n$. Then U_n is a simple and bounded process and $B'_n \in L^+_\infty(P)$. But

$$(H_n \cdot X)_{T_n} - B'_n = \begin{cases} (H_n \cdot X)_1 - B_n & \text{on } \{T_n = 1\} \\ 0 & \text{on } A^c \cap \{T_n < 1\} \\ 1 & \text{on } A \cap \{T_n < 1\}. \end{cases}$$

Therefore, $(H_n \cdot X)_{T_n} - B'_n$ converges to $\mathbb{1}_A$ in E. From (3.2), we see that $(H_n \cdot X)_{T'_m} - B'_n$ also converges to $\mathbb{1}_A$.

Since $((U_n \cdot X)_1)$ is bounded, we can extract a sub-sequence which converges weakly to a random variable Y. As $\overline{K}$ is convex and closed in the strong topology, $Y \in \overline{K}$. But $(H_n \cdot X)_1 - B_n$ converges to $\mathbb{1}_A$ in E, with the result that the weak convergence of $(H_n \cdot X)_1$ leads to the weak convergence of B_n to a random variable $B \geq 0$. Thus, $Y = B + \mathbb{1}_A \neq 0$. This shows that $Y \in E^+$, contradicting our assumption that $E^+ \cap \overline{K} = \{0\}$.

Hence, $\mathbb{1}_A \notin \overline{K - L_\infty^+(P)}$. Hence, there exists a probability Q equivalent to P and with Radon-Nikodým derivative $\frac{dQ}{dP} \in E^*$ such that X is a Q-martingale. □

Let (Φ, Ψ) be complementary finite Young functions. As $H_\Phi(P)$ is a Banach lattice with order continuous norm, $e = \mathbf{1}$ is a quasi-interior point of $H_\Phi(P)$ and we have that $L_\infty(P) \subseteq H_\Phi(P) \subseteq L_1(P)$, we can specialise Theorem 3.2.2 and Theorem 3.2.3 to Orlicz hearts by letting $E = H_\Phi(P)$.

In incomplete markets, there does not exist a unique equivalent martingale measure that can be used to price contingent claims. One method of choosing an equivalent pricing measure, is to incorporate a preference structure into the pricing techniques. This is done via utility funtions, which will be introduced in the next section.

3.3 Utility functions

There are two different ways to model the prices of assets: via a no arbitrage model or via a capital asset pricing model. The latter is based on balancing supply with demand among investors who have utility functions that convert units of consumption into units of satisfaction. In other words, utility functions are used to measure an investor's preference for wealth and how much risk he is willing to undertake to gain more wealth. While the no arbitrage model allows for precise quantitative insights into the market in a complete market setting, the utility based methods are currently the only theoretically defensible way to model in incomplete markets [144].

Definition 3.3.1. A *utility function* $u : \mathbb{R} \to [-\infty, \infty]$ is a non-decreasing, twice differentiable function, with the following properties:

(i) non-satiation, i.e. $u'(x) > 0$, where u' denotes the first derivative of u.

(ii) risk aversion, i.e. $u''(x) < 0$, where u'' denotes the second derivative of u.

The non-satiation property implies that the the utility increases with wealth, i.e. that more wealth is preferred to less wealth, and that the investor never has enough wealth.

The risk aversion property implies that the utility function is concave, i.e. that the marginal utility of wealth decreases as wealth increases. This concavity of the utility function captures the trade-off between risk and return. In other words, the gain of one dollar to someone who only has one dollar is worth more than to someone who already has a million dollars.

Different investors can have different utility functions, as long as they satisfy the above mentioned properties.

By comparing expected utility of payoffs instead of expected payoffs, and choosing the utility function judiciously, it is possible to capture an investor's attitude towards the trade-off between risk and return. Therefore, in the theory of portfolio optimisation, a rational investor will always try to maximise his expected utility of wealth.

Definition 3.3.2. The *principle of expected utility maximisation* involves finding the optimal portfolio, by solving the following:

$$\max_{I\in\mathcal{H}} \mathbb{E}[u(X(I))],$$

where $\mathcal{H}$ is the set of all feasible investment options and $X(I)$ is the terminal value of the investment I after the given time period.

Note, that this expectation is computed under the real-world probability measure and not the risk-neutral one. It would not make sense to work under the risk-neutral measure. Under this measure, the stock and the money market account have the same expected rate of return and hence, the investor would only invest in the money market.

Example 3.3.3 (A fair game)**.** Consider the function defined by

$$u(x) = \sqrt{x}.$$

As

$$u'(x) = 0.5x^{-0.5} > 0 \text{ and } u''(x) = -0.25x^{-1.5} < 0,$$

we have that u is a utility function. Assume that the investor's initial wealth is \$5 and assume that there is only one investment available. In this investment a fair coin is flipped. If it comes up heads, the investor wins \$4, increasing his wealth to \$9. If it comes up tails, the investor loses \$4, decreasing his wealth to \$1. This is called a fair game as the expected gain is $0.5(4) + 0.5(-4) = 0$. Will the investor choose to play this game if he follows the principle of expected utility maximisation?

If he refuses to play the game, he has an expected utility of $\sqrt{5} = 2.24$. If he decides to play the game, his expected utility becomes $0.5\sqrt{1} + 0.5\sqrt{9} = 2$. Since 2.24 is greater than 2, the investor will refuse to play the game.

In general, a risk-averse investor will never choose an investment whose expected return is 0%. In other words, the property of risk aversion implies that investors attach more weight to losses than they do to gains of equal magnitude.

If we change the probability of a good outcome to 75%, then the expected outcome is \$7, the expected return is 40% and the expected utility would be $0.75\sqrt{9} + 0.25\sqrt{1} = 2.5$. Since 2.5 is greater than 2.24, the investor would be willing to make the investment. The expected return of 40% is a 'risk premium', which compensates him for undertaking the risk of the investment. □

Definition 3.3.4.

(i) The *certainty equivalent* for an investment with outcome given by a random variable x, is denoted by c and is defined by

$$u(c) = \mathbb{E}(u(x)).$$

(ii) The *risk premium* for an investment with outcome given by a random variable x, is denoted by ρ and is defined by

$$\rho(x) = \mathbb{E}[x] - c.$$

An investor with current wealth less than the certainty equivalent will consider the investment attractive, while an investor with current wealth greater than the certainty equivalent will not. Note that since the utility function is an increasing function, maximising expected utility is equivalent to maximising the certainty equivalent.

Definition 3.3.5. Suppose that u is a twice continuously differentiable utility function on Ω. Then

$$\alpha(x) = -\frac{u''(x)}{u'(x)}$$

is called the *Arrow-Pratt coefficient of absolute risk aversion* of u at level x.

The most general class of utility functions used in practice, is called the *hyperbolic absolute risk aversion class*, also known as the *HARA* class. A utility function falls into the

HARA class, if its Arrow-Pratt coefficient of absolute risk aversion is hyperbolic, i.e.

$$\alpha(x) = \frac{1}{ax + b}$$

for $a, b \in \mathbb{R}$. This class is obtained as follows. Let $c \in \mathbb{R}$ and $p \in [0, 1)$. For $p \neq 0$, define

$$u_p(x) = \begin{cases} \frac{1}{p}(x-c)^p & \text{if } x > c \\ 0 & \text{if } 0 < p < 1 \text{ and } x = c \\ -\infty & \text{if } p < 0 \text{ and } x = c \\ -\infty & \text{if } x < c. \end{cases}$$

For $p = 0$, define the logarithmic utility function

$$u_0(c) = \begin{cases} \ln(x-c) & \text{if } x > c \\ -\infty & \text{if } x \leq c. \end{cases}$$

We end this section with another example.

Example 3.3.6 (Optimising a portfolio)**.** Consider an investment, which returns -10% with probability $\frac{1}{2}$ and $+20\%$ with probability $\frac{1}{2}$. The investor has the choice to invest any part of his total wealth of \$100 in the risky asset. We will consider the following utility function. For any $\lambda < 1, \lambda \neq 0$, let

$$u_\lambda(x) = \frac{x^\lambda - 1}{\lambda}.$$

Then,

$$\begin{aligned} u'_\lambda(x) &= x^{\lambda-1} > 0 \\ u''_\lambda(x) &= (\lambda - 1)x^{\lambda-2} < 0. \end{aligned}$$

Let θ be the amount invested in the risky asset and the investor does nothing with the remaining $100 - \theta$. The two possible outcomes are:

$$\text{Bad outcome: } x = 0.9\theta + (100 - \theta) = 100 - 0.1\theta.$$

$$\text{Good outcome: } x = 1.2\theta + (100 - \theta) = 100 + 0.2\theta.$$

The expected utility of the outcome is

$$\begin{aligned} f(\theta) &= \frac{1}{2}u(100 - 0.1\theta) + \frac{1}{2}u(100 + 0.2\theta) \\ &= \frac{(100 - 0.1\theta)^\lambda - 1}{2\lambda} + \frac{(100 + 0.2\theta)^\lambda - 1}{2\lambda} \\ &= \frac{1}{2\lambda}[(100 - 0.1\theta)^\lambda + (100 + 0.2\theta)^\lambda - 2]. \end{aligned}$$

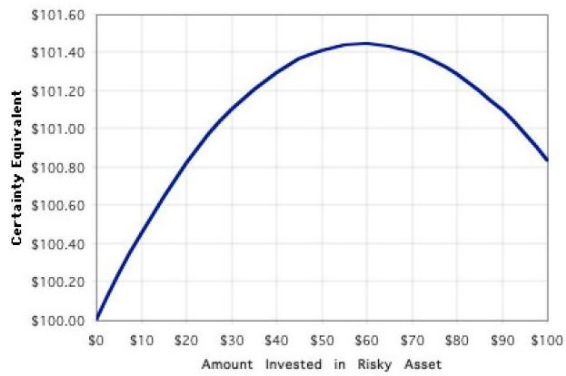

Figure 3.1: Utility Hill for $\lambda = -3$.

Figure 3.1 shows the certainty equivalent of the function $u^{-1}(f(\theta))$ for $\lambda = -3$. Since maximising expected utility is equivalent to maximising the certainty equivalent, Figure 3.1 indicates that the optimal amount to invest in the risky asset in the case of $\lambda = -3$ is approximately \$59.

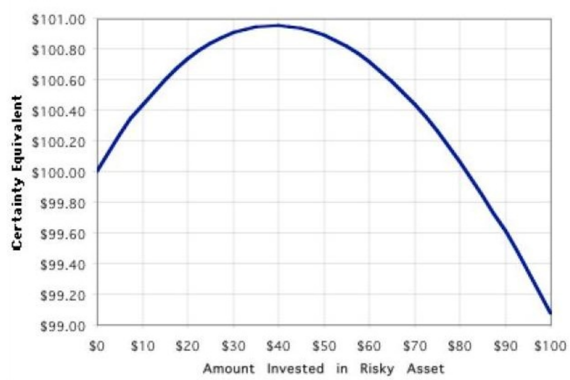

Figure 3.2: Utility Hill for $\lambda = -5$.

Figure 3.2 shows the corresponding graph for $\lambda = -5$. For this more risk averse investor, the optimal amount to invest in the risky asset is approximately \$39.

Since, this example is relatively simple, we can find the exact optimal portfolio. The value of θ that maximises $f(\theta)$, is given by

$$\theta = \frac{100(2^{\frac{1}{A}} - 1)}{0.2 + 0.1(2^{\frac{1}{A}})},$$

where $A = 1 - \lambda$. The number A is called the *coefficient of risk aversion*. As λ decreases, investors become more risk averse and vice versa.

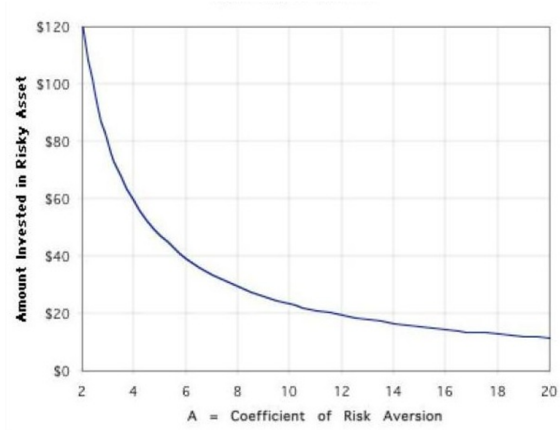

Figure 3.3: Amount invested in risky asset as coefficient of risk aversion changes.

Figure 3.3 shows how much the investor should invest in the risky asset as a function of his risk aversion. It is interesting to note that an investor with a very low coefficient of risk aversion, should actually invest more than his total wealth in the risky asset.

3.4 No market free lunch

In this section we will consider the positive cone of $L_\infty(P)$ excluding zero, i.e. $L_\infty^+(P)\backslash\{0\}$. Note that the set

$$\{w \in L_\infty(P) : P(w \geq 0) = 1 \text{ and } P(w > 0) > 0\},$$

as considered in [63], is equal to $L_\infty^+(P)\backslash\{0\}$.

We will interpret each element $w \in L_\infty^+(P)\backslash\{0\}$ as the time T payoff of a claim.

If we short sell this claim, then we will receive a positive amount now, but will have to pay back $-w$ at time T. But today we could also choose an admissible trading strategy, with zero (or negative) initial cost, that might 'hedge' the claim w. At time t, our payoff will then be $f - w$, where $f \in C$.

Lemma 3.4.1. *If for some $w \in L_\infty^+(P)\backslash\{0\}$ and $f \in C$*

$$f - w \geq 0 \ P\text{-}a.s.,$$

then the no arbitrage condition is violated.

Proof. Select $w \in L_\infty^+(P)\backslash\{0\}$ and $f \in C = (K - L_0^+(P)) \cap L_\infty(P)$ such that $f - w \geq 0$ P-a.s.. Since both $f, w \in L_\infty(P)$, we have $f - w \in L_\infty^+(P)$.

Since $f \in (K - L_0^+(P))$, we can write $f = k - l$ where $k \in K$ and $l \in L_0^+(P)$. Then

$$f - w = k - (l + w),$$

where $l + w \in L_0^+(P)$. Hence $f - w \in (K - L_0^+(P))$ and thus $f - w \in C$. In other words, $f - w \in C \cap L_\infty^+(P)$, which implies that arbitrage is possible. □

Definition 3.4.2. The *essential infimum*, denoted by $ess \inf f$, is given by

$$ess \inf f = \sup\{z : f \geq z \text{ a.e.}\}.$$

Lemma 3.4.3. *The following are equivalent.*

(i) There is FLVR.

(ii) There exist $w \in L_\infty^+(P)\backslash\{0\}$ and a sequence $(f_n) \subseteq C$ such that

$$\lim_{n\to\infty} \|f_n - w\|_\infty = 0. \tag{3.3}$$

(iii) There exist $w \in L_\infty^+(P)\backslash\{0\}$ and $f \in C$ such that

$$\sup_{f\in C}\{ess \inf_\Omega (f - w)\} \geq 0. \tag{3.4}$$

Proof. To show (i) ⇔ (ii) is trivial.

(ii) ⇒ (iii): Suppose there exist $w \in L_\infty^+(P)\backslash\{0\}$ and a sequence $(f_n) \subseteq C$ such that $\lim\limits_{n\to\infty} \|f_n - w\|_\infty = 0$. Let $\tilde{f}_n = f_n - (f_n - w)^+$ for all $n \in \mathbb{N}$. Then $(\tilde{f}_n) \subseteq C$ and

$$-\frac{1}{n} \leq \tilde{f}_n - w \leq (f_n - w) - (f_n - w)^+ \leq 0.$$

Hence,

$$ess \inf_\Omega (\tilde{f}_n - w) \geq -\frac{1}{n},$$

and so

$$\sup_{f\in C}\{ess \inf_\Omega (f - w)\} \geq 0.$$

(iii) ⇒ (ii): Suppose there exist $w \in L_\infty^+(P)\backslash\{0\}$ and $f \in C$ such that $\sup\limits_{f\in C}\{ess \inf\limits_\Omega (f - w)\} \geq 0$. Then, for all $n \in \mathbb{N}$, there exists $f_n \in C$ such that

$$ess \inf_\Omega (f_n - w) \geq -\frac{1}{n}.$$

Let $\tilde{f}_n = f_n - (f_n - w)^+$. Then $\tilde{f}_n \in C$ and

$$-\frac{1}{n} \leq \tilde{f}_n - w \leq (f_n - w) - (f_n - w)^+ \leq 0 \leq \frac{1}{n}.$$

Hence, we have that

$$\lim_{n\to\infty} ||\tilde{f}_n - w||_\infty = 0.$$

□

Let $\mathbb{U}$ be a certain set of utility functions $u : \mathbb{R} \to [-\infty, \infty]$. We assume that the preference " $\succeq$ " of the investors in the market under consideration can be represented by the expected utility, i.e.

$$f_1 \succeq f_2 \quad \Leftrightarrow \quad \mathbb{E}_Q[\, u(f_1)\,] \geq \mathbb{E}_Q[\, u(f_2)\,],$$

where $Q \in \mathbb{P}$, $u \in \mathbb{U}$ and $f_1,\ f_2 \in L_0(P)$.

Frittelli [63] introduced the notion of a market free lunch that depends on the preferences of the investors in the market. Market free lunch with respect to $\mathbb{U}$ is defined as follows.

Definition 3.4.4. [63, Definition 3] There is a *market free lunch* with respect to $\mathbb{U}$ if for all $P \in \mathbb{P}$ and $u \in \mathbb{U}$, there exists $w \in L_\infty^+(P)\backslash\{0\}$ such that

$$\sup_{f\in C} \mathbb{E}_P[\, u(f - w)\,] \geq u(0). \tag{3.5}$$

Hence, there is *no market free lunch* (NMFL($\mathbb{U}$)) with respect to $\mathbb{U}$ if for all $w \in L_\infty^+(P)\backslash\{0\}$ there exist $P \in \mathbb{P}$ and $u \in \mathbb{U}$ such that

$$\sup_{f\in C} \mathbb{E}_P[\, u(f - w)\,] < u(0).$$

This definition clearly depends on the set of utility functions, which we choose. The above definition only makes economical sense if our utility function is non-decreasing on $\mathbb{R}$. Consider $w \in L_\infty^+(P)\backslash\{0\}$ such that Equation 3.5 holds. A market free lunch implies that all investors in the market, who are represented by their beliefs $P \in \mathbb{P}$ and their preferences $u \in \mathbb{U}$, regard the risk w as a free lunch, as each investor can hedge the risk g in such a way that their preferences and beliefs are not compromised.

Consider the following families of utility functions

$$\mathbb{U}_0 = \{u : \mathbb{R} \to [-\infty, \infty] : u \text{ is non-decreasing on } \mathbb{R}\},$$
$$\mathbb{U}_1 = \{u \in \mathbb{U}_0 : u \text{ is left continuous at } 0 \in \text{int}(\text{dom}(u))\} \text{ and}$$
$$\mathbb{U}_2 = \{u \in \mathbb{U}_0 : u \text{ is finite-valued and concave on } \mathbb{R}\}.$$

Note that $\mathbb{U}_2 \subseteq \mathbb{U}_1 \subseteq \mathbb{U}_0$ and $NMFL(\mathbb{U}_0) \Rightarrow NMFL(\mathbb{U}_1) \Rightarrow NMFL(\mathbb{U}_2)$.

Proposition 3.4.5.

(i) $NMFL(\mathbb{U}_0) \Leftrightarrow NA.$

(ii) $NMFL(\mathbb{U}_1) \Leftrightarrow NFLVR.$

Proof. ([63, Proposition 5]) (i): Assume that an arbitrage opportunity exists. We need to show that there exists $w \in L_\infty^+(P)\backslash\{0\}$ such that (3.5) holds. Consider $w \in C \cap L_\infty^+(P)\backslash\{0\}$. Since $L_\infty^+(P)\backslash\{0\} \subseteq L_\infty(P)$, we have that $w \in C \cap L_\infty(P)$, which shows that w is an arbitrage opportunity. Then, for $u \in \mathbb{U}_0$

$$\sup_{f\in C} \mathbb{E}_P[\, u(f-w)\,] \geq \mathbb{E}_P[\, u(w-w)\,] = u(0),$$

i.e. there is a MFL($\mathbb{U}_0$).

Conversely, assume there exists a MFL($\mathbb{U}_0$), i.e. there exists $w \in L_\infty^+(P)\backslash\{0\}$ such that $\sup_{f\in C} \mathbb{E}_P[\, u(f-w)\,] \geq u(0)$ holds for all $P \in \mathbb{P}$ and $u \in \mathbb{U}_0$. Take $P \in \mathbb{P}$ and define $u \in \mathbb{U}_0$ by

$$u(x) = \begin{cases} 0 & \text{for } x \geq 0 \\ -\infty & \text{for } x < 0. \end{cases}$$

Thus, $\sup_{f\in C} \mathbb{E}_P[\, u(f-w)\,] \geq 0$ and due to the definition of u, there must exist $f \in C$ such that $P(f - w \geq 0) = 1$, i.e. by Lemma 3.4.1, there is an arbitrage opportunity.

(ii): Let $n \geq 1$. Suppose there is a free lunch with vanishing risk. Then, by Lemma 3.4.3, there exist $w \in L_\infty^+(P)\backslash\{0\}$ and a sequence $(f_n) \subseteq C$ such that (3.4) is satisfied. From the left continuity of u at 0, there exists $\epsilon_n > 0$ such that if $-\epsilon_n < x \leq 0$, then $u(x) > u(0) - \frac{1}{n}$.

For each n, there exists $\delta_n > 0$ such that $\lim_{n\to\infty} \|f_n - w\|_\infty > \delta_n$. Set $\tilde{f}_n := f_n - (f_n - w)^+$. Then $\tilde{f}_n \in C$ and $-\delta_n < (\tilde{f}_n - w) \leq 0$ P-a.s.. Hence, $u(\tilde{f}_n - w) > u(0) - \frac{1}{n}$ P-a.s.. Thus, we can conclude that $\sup_{f\in C} \mathbb{E}_P[\, u(f-w)\,] \geq u(0)$, i.e. a MFL($\mathbb{U}_1$) exists.

Conversely, suppose that there exists a MFL($\mathbb{U}_1$). Take $P \in \mathbb{P}$ and for all $n \geq 1$, define u_n by

$$u_n(x) = \begin{cases} 0 & \text{for } x > -\frac{1}{n} \\ -\infty & \text{for } x \leq -\frac{1}{n}. \end{cases}$$

Then $u_n \in \mathbb{U}_1$. By assumption there exists $w \in L_\infty^+(P)\backslash\{0\}$ such that $\sup_{f\in C} \mathbb{E}_P[\, u_n(f-w)\,] \geq u_n(0) = 0$ for all $n \geq 1$. Hence, there exists $f_n \in C$ such that $P(f_n - w \leq -\frac{1}{n}) = 0$, i.e. such that $ess\inf(f_n - w) \geq -\frac{1}{n}$ for all $n \geq 1$. Thus, $\sup_{f\in C}\{ess\inf(f-w)\} \geq 0$, which, by Lemma 3.4.3, implies that there exists a FLVR. □

This proposition shows that the difference, from an economic perspective, between NA and NFLVR is due to the differing preferences of the investors.

Under this new concept of NMFL, Bellini and Frittelli [11] proved another version of the fundamental theorem of asset pricing.

Theorem 3.4.6. *For any semimartingale S, the following are equivalent:*

(i) S *satisfies NMFL*$(\mathbb{U}_2)$,

(ii) $\mathcal{M} \cap \mathbb{P} \neq \emptyset$.

We will not prove this theorem here, as we will be considering an alternative version of it in Section 3.5, which we will then prove. See [14] for a proof of Theorem 3.4.6.

If S is not locally bounded, then $\mathcal{M}$ cannot be reduced to a simpler form. Hence, we need to introduce a new setup, introduced by Frittelli [64], which allows us to work with unbounded processes.

3.5 No free lunch in Orlicz spaces

3.5.1 The optimisation problem

The analysis of any optimisation problem depends greatly on the definition of the domain of optimisation and the objective function. In a financial setting, it is very helpful to ensure that the optimal value is finite. Therefore, the utility maximisation problem, given by

$$\sup_{k \in D} \mathbb{E}[\ u(x+k)\ ],$$

where u is the utility function, x is the initial wealth of the investor and D is an appropriate set of admissible strategies, requires the specification of

1. the financial market model and the admissible terminal wealths,
2. the technical assumptions on the utility function and
3. some joint conditions on the market model and the utility function.

The market model and the admissible terminal wealths have been described in Section 3.1.

The utility function $u : \mathbb{R} \to [-\infty, \infty]$ is increasing and concave on (a, ∞), where $a \in [-\infty, 0)$ and, if a is finite, then $u(x) = -\infty$ for all $x \leq a$. Hence, $u(x)$ must satisfy $\lim_{x \to -\infty} u(x) = -\infty$.

Two widely used utility functions satisfy the above conditions. They are stated in the following example.

Example 3.5.1.

1. The logarithmic utility function is given by

$$u(x) = \begin{cases} \ln(1+x) & \text{for } x > -1 \\ -\infty & \text{for } x \leq -1. \end{cases}$$

   In this example, we have $a = -1$.

2. The exponential utility function is given by

$$u(x) = -e^{-x} \quad \text{for } x \in \mathbb{R}.$$

   Here $a = -\infty$.

Figure 3.4 shows the general shape of a utility function, that satisfies these conditions.

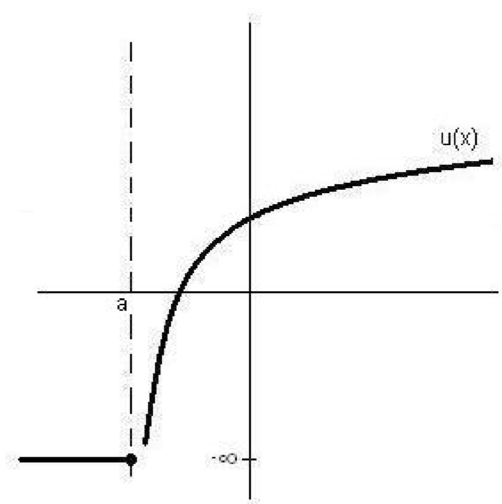

Figure 3.4: General shape of a Utility function.

In general, our class of utility functions includes all those functions, for which

$$\lim_{x \downarrow 0} \frac{u(x)}{x} < \infty$$

and those functions, which are constant for $x \geq x_0$, where $x_0 \in \mathbb{R}$ is fixed.

Definition 3.5.2.

(i) The utility function u satisfies the *Inada conditions* if

$$\lim_{x\downarrow-\infty} u'(x) = -\infty \qquad \text{and} \qquad \lim_{x\uparrow\infty} u'(x) = 0.$$

(ii) The utility function u satisfies the *reasonable asymptotic elasticity* condition $RAE(u)$ introduced in [104, 141] if

$$\liminf_{x\to-\infty} \frac{xu'(x)}{u(x)} > 1 \qquad \text{and} \qquad \limsup_{x\to\infty} \frac{xu'(x)}{u(x)} < 1.$$

When working with a locally bounded semimartingale, the above two technical conditions ensure that the optimisation problem has a solution. This is shown in [141]. However, it has been shown [13, 105] that within a specific market model one may state more general necessary and sufficient conditions on the utility function that allow the dual approach of the optimisation to work and ensure the existence of an optimal investment. The advantages of the RAE(u) condition is that it is easily verified and the most commonly used utility functions satisfy it.

Lastly, we need to look at the joint conditions between the market model and the utility function. The duality approach to the optimisation problem requires the convex conjugate Φ of the utility function u, given by

$$\Phi(y) = \sup_{x\in\mathbb{R}}\{u(x) - xy\}.$$

This is just the Legendre-transform of $x \mapsto -u(-x)$. Note that $\Phi' = -(u')^{-1}$.

Definition 3.5.3. A probability $Q << P$ has *finite generalised entropy* if its density $\frac{dQ}{dP}$ satisfies the integrability condition

$$\mathbb{E}[\ \Phi(\frac{dQ}{dP})\ ] < \infty,$$

where Φ denotes the convex conjugate of the utility function u.

Given Φ, we denote the set of pricing measures with finite generalised entropy by M_Φ.

In general, the optimisation problem above does not admit an optimal solution if $D = K$. It was first shown in Bellini and Frittelli [11] that if

$$\sup_{k\in K} \mathbb{E}_P[u(x+k)] < \lim_{y\to\infty} u(y),$$

then the following duality relation holds

$$\sup_{k\in K}\mathbb{E}_P[u(x+k)] = \min_{Q\in\mathcal{M}}\min_{\lambda>0}\Big\{\lambda x + \mathbb{E}_P\Big[\Phi(\lambda\frac{dQ}{dP})\Big]\Big\},$$

without any further assumptions on the utility function. In this section, we show a generalised version of this relation as done in [15].

Frittelli proves a clear financial interpretation of the set M_Φ. Pricing by $Q \in M_\Phi$ guarantees that the investor cannot reach his maximum possible utility.

The utility maximisation problem can take on various forms. We will state it now as in the work of Biagini and Frittelli [15, 64]. We are trying to solve the following problem

$$\sup_{H\in\mathcal{H}}\mathbb{E}_P[u(x+(H\cdot S)_T)], \tag{3.6}$$

where

- u is the utility function of the investor, which is assumed to be concave and increasing over its proper domain,
- x is the initial endowment of the investor, $x \in \mathrm{dom}(u)$,
- T is the time horizon,
- $\mathcal{H}$ is the proper class of admissible $\mathbb{R}^d$-valued predictable processes, which represent the allowed trading strategies,
- P is the real-world probability measure,
- S is an $\mathbb{R}^d$-valued càdlàg semimartingale defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, where the filtration $\mathbb{F}$ satisfies the usual assumptions of right continuity and completeness and $\mathcal{F}_0$ is trivial, i.e. consists only of null sets and their complements,
- $(H\cdot S)_T$ is the terminal gain of the investor when following strategy H.

Initially, this maximisation was done over $\mathcal{H}^1$ (see [104, 141, 13]). However, when S is non-locally bounded, $\mathcal{H}^1$ may reduce to the null set, making the maximisation trivial. This may happen if, for example, S is a compound Poisson process with unbounded jump sizes. Therefore, we need to enlarge this set by introducing the less-restrictive notion of W-admissible strategies. This extension of the notion of admissibility was first used by Delbaen and Schachermayer [40] in the context of the fundamental theorem of asset pricing, and subsequently by Biagini and Frittelli [15] in the context of utility maximisation.

Definition 3.5.4. Let $W \in L_0(P)$ be a fixed random variable. The predictable and S-integrable process H is W-*admissible* if there exists a nonnegative constant c such that for all $t \leq T$,

$$(H \cdot S)_T \geq -cW.$$

The set of all W-admissible trading strategies is defined by

$$\mathcal{H}^W = \{H \in \mathcal{H} : (H \cdot S)_t \geq -cW \text{ for some } c > 0 \text{ and for all } t \in [0, T]\}.$$

In other words, the random variable W controls the losses in trading and, in order to build a reasonable utility maximisation, should satisfy two conditions. These two conditions are stated next.

Definition 3.5.5. A random variable $W \geq 1$ is *suitable* for the process S if for each $i = 1, \ldots, d$, there exists a process $H^i \in \mathcal{H}$ such that

$$P(\{\omega : H^i_t(\omega) = 0 \text{ for some } t \geq 0\}) = 0$$

and

$$|(H^i \cdot S^i)_t| \leq W \quad \text{for all } t \in [0, T].$$

The set of suitable random variables is denoted by $\mathbb{S}$.

The suitability condition changes the integrability of the stochastic integral.

Definition 3.5.6. A positive random variable W is *compatible* with the utility function u if

$$\mathbb{E}_P[\, u(-\alpha W)\,] > -\infty \text{ for all } \alpha > 0$$

and *weakly compatible* with u if

$$\mathbb{E}_P[\, u(-\alpha W)\,] > -\infty \text{ for some } \alpha > 0.$$

We call a suitable and compatible random variable a *loss variable*.

Stochastic integrals formed with W-admissible strategies enjoy good mathematical properties, when the random variable W satisfies suitability with the market and compatibility with the preference structure.

The problem we are trying to solve is written in terms of an optimisation over stochastic processes, however, to use the duality arguments, we need it to be written as an optimisation over random variables. Therefore, given a suitable and compatible random variable W, we define the set of terminal values obtained from W-admissible trading strategies by

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

The primal optimisation problem can then be written as

$$\sup_{k \in K^W} \mathbb{E}_P[\, u(x+k)\,].$$

Note that for the duality, W does not necessarily have to be suitable. But this property is very desirable, as it makes the domain of maximisation K^W non-trivial.

The next step is to identify an appropriate cone C^W, related to K^W, and invoke Fenchel's duality theorem (see Appendix C). To do this, we need to choose an appropriate Banach space and its order dual, so that we can define the polar set $(C^W)^\circ$. Classically, the spaces $(L_1(P), L_\infty(P))$ were used to deal with locally bounded traded assets. Biagini and Frittelli [64], however, wanted to accommodate more general markets and decided to use an appropriate Orlicz space, which is defined in the next section. The following choice was mainly made because of the similarity between the compatibility condition and the heart of the appropriate Orlicz space.

3.5.2 The Orlicz space associated with u

To create an Orlicz space using a utility function we need to find a convex function which is related to the concave utility function.

Lemma 3.5.7. *If u is a utility function, then $\hat{u} : [0,\infty) \to [0,\infty]$ defined by*

$$\hat{u}(x) = -u(-x) + u(0),$$

is a Young function.

Figures 3.4 and 3.5 can be used to aid in the proof of Lemma 3.5.7.

The Orlicz space induced by $\hat{u}$ is given by

$$L_{\hat{u}}(P) = \{f \in L_0(P) : \mathbb{E}[\, \hat{u}(\alpha|f|)\,] < \infty \ \text{ for some } \alpha > 0\}$$

and we endow it with the Luxemburg norm

$$\mathcal{N}_{\hat{u}}(f) = \inf\{a > 0 : M_{\hat{u}}(\frac{f}{a}) \leq 1\}.$$

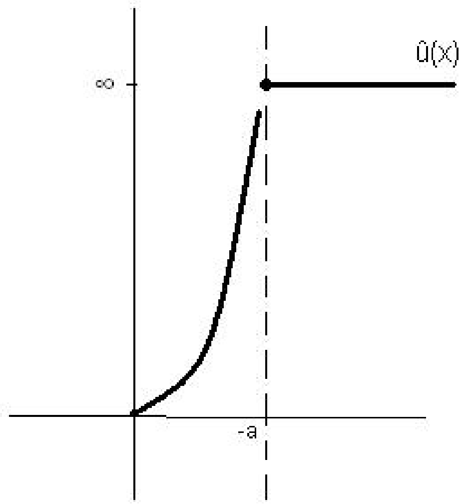

Figure 3.5: General shape of $\hat{u}$.

Using Corollary 2.1.3, the convex conjugate function $\hat{\Phi}$ of $\hat{u}$ is given by

$$\hat{\Phi}(y) = \sup_x \{xy - \hat{u}(x)\},$$

which is also a Young function, and therefore we can introduce the Orlicz space $L_{\hat{\Phi}}(P)$ given by

$$L_{\hat{\Phi}}(P) = \{f \in L_0(P) : \mathbb{E}[\ \hat{\Phi}(\alpha|f|)\ ] < \infty \ \text{ for some } \alpha > 0\},$$

which we endow with the Orlicz norm

$$\|f\|_{\hat{\Phi}} = \sup\{\mathbb{E}[\ |fg|\ ] : g \in L_{\hat{u}}(P) \text{ and } \mathbb{E}[\ \hat{u}(g)\ ] \leq 1\}.$$

The heart of the Orlicz space $L_{\hat{u}}(P)$ is given by

$$H_{\hat{u}}(P) = \{f \in L_0(P) : \mathbb{E}[\ \hat{u}(\alpha|f|)\ ] < \infty \ \text{ for all } \alpha > 0\}.$$

The utility function is increasing and concave on (a, ∞). We will consider the two possible cases: either a is finite or a is infinite.

- **The case $a < 0$ is finite.**
  Since $\hat{u}$ is a Young function, we have that $-u(-x) + u(0) \geq 0$ for all x. Therefore, for all $|x| \leq -a$ and $y > 0$, we have

$$\begin{aligned} xy - \hat{u}(x) &= xy + u(-x) - u(0) \\ &\leq xy \\ &\leq -ay. \end{aligned}$$

Hence, $\hat{\Phi}(y) \leq -ay$, and in particular $\hat{\Phi}(\alpha|y|) \leq -a\alpha|y|$. Taking expectations on both sides results in

$$\mathbb{E}[\ \hat{\Phi}(\alpha|y|)\ ] \leq \mathbb{E}[\ -a\alpha|y|\ ] = -a\alpha\mathbb{E}[\ |y|\ ].$$

Now, if we take $f \in L_1(P)$, then

$$\mathbb{E}[\ \hat{\Phi}(\alpha|f|)\ ] \leq -a\alpha\mathbb{E}[\ |f|\ ] < \infty$$

and $f \in L_{\hat{\Phi}}(P)$. This and Proposition 2.2.2 show that $L_{\hat{\Phi}}(P) = L_1(P)$.

Since $\hat{u}(y) = \infty$ for some y, we have by Proposition 2.2.1 and 2.2.2, that $L_{\hat{u}}(P) = L_\infty(P)$ and $H_{\hat{u}}(P) = \{0\}$. Since $L_{\hat{u}}(P)$, $L_{\hat{\Phi}}(P)$, $L_1(P)$ and $L_\infty(P)$ are Banach spaces, we also have that the Luxemburg norm $\mathcal{N}_{\hat{u}}$ and the Orlicz norm $\|\cdot\|_{\hat{\Phi}}$ are equivalent to the $\|\cdot\|_\infty$-norm and the $\|\cdot\|_1$-norm respectively. This follows from a well known result in functional analysis, see Theorem A.3.2.

- **The case $a = -\infty$.**
  The function $\hat{u}$ is continuous and consequently the subspace $H_{\hat{u}}(P)$ is a Banach space with the inherited $\hat{u}$-norm. We will look at two examples, one with the exponential utility and one with a linear utility.

Example 3.5.8.

1. Let $u(x) = -e^{-x}$, then $\hat{u}(x) = e^x - 1$, $\Phi(y) = y \ln y - y$ and

$$\hat{\Phi}(y) = (|y| \ln |y| - |y| + 1)\mathbb{1}_{\{|y| \geq 1\}}.$$

   Therefore

$$\begin{aligned} L_{\hat{u}}(P) &= \Big\{ f \in L_0(P) : \mathbb{E}[\, e^{\alpha|f|}\,] < \infty \ \text{ for some } \alpha > 0 \Big\}, \\ H_{\hat{u}}(P) &= \Big\{ f \in L_0(P) : \mathbb{E}[\, e^{\alpha|f|}\,] < \infty \ \text{ for all } \alpha > 0 \Big\}, \text{ and} \\ L_{\hat{\Phi}}(P) &= \Big\{ g \in L_0(P) : \mathbb{E}[\, |g| \ln |g|\,] + (\ln \alpha - 1)\mathbb{E}[\, |g|\,] < \infty \text{ for some } \alpha > 0 \Big\}. \end{aligned}$$

2. Let $u(x) = x$, then $\hat{u}(x) = x$, $\Phi(y) = \infty$ for all $y \geq 0$ and

$$\hat{\Phi}(y) = (\infty)\mathbb{1}_{\{|y| \geq 1\}} = \delta_{\{|y| \leq 1\}}.$$

   So, $L_{\hat{u}}(P) = L_1(P) = H_{\hat{u}}(P)$ and $L_{\hat{\Phi}}(P) = L_\infty(P)$.

Note that the Young function $\hat{u}$ carries information about the utility function at large losses, in the sense that for $\alpha > 0$ we have that

$$\mathbb{E}_P[\ \hat{u}(\alpha|f|)\ ] < \infty \qquad \Leftrightarrow \qquad \mathbb{E}_P[\ u(-\alpha|f|)\ ] > -\infty.$$

Hence, a random variable W is compatible with a utility function u if and only if $W \in H_{\hat{u}}(P)$ and weakly compatible if and only if $W \in L_{\hat{u}}(P)$. The suitability condition shows that $(H \cdot S)_t$ is an element of the ideal generated by W in $L_{\hat{u}}(P)$.

Next we need to have a look at the duality. As mentioned before, we can generally decompose the Banach dual of the Orlicz space $L_{\hat{u}}(P)$ into the following:

$$(L_{\hat{u}}(P))^* = (L_{\hat{u}}(P))^*_c \oplus (L_{\hat{u}}(P))^*_s.$$

Depending on the nature of u, this decomposition can be reduced to more familiar spaces.

- If a is finite, then $L_{\hat{u}}(P) = L_\infty(P)$ and the above decomposition reduces to the Yosida-Hewitt decomposition for elements of $(L_\infty(P))^*$, i.e.

$$(L_\infty(P))^* = L_1(P) \oplus (L_{\hat{u}}(P))^*_s.$$

- If $a = -\infty$, then $\hat{u}$ is continuous and $\hat{u}(x) = 0$ if and only if $x = 0$, and by Theorem 2.5.8 we have

$$(L_{\hat{u}}(P))^* = (H_{\hat{u}}(P))^* \oplus (H_{\hat{u}}(P))^{\text{annh}} = L_{\hat{\Phi}}(P) \oplus (H_{\hat{u}}(P))^{\text{annh}}.$$

Note that Biagini and Frittelli proved in [15, Proposition 11], that $(H_{\hat{u}}(P))^{\text{annh}}$ is an AL-space. However, it is just a special case of Corollary 2.5.7.

Another thing to note is that $H_{\hat{u}}(P) = \overline{L_\infty(P)}^{\hat{u}}$ and consequently

$$z \in (L_{\hat{u}}(P))^*_s \Leftrightarrow z(f) = 0 \text{ for all } f \in L_\infty(P).$$

Therefore, we can identify $z_c \in (L_{\hat{u}}(P))^*_c$ of any $z \in (L_{\hat{u}}(P))^*_+$ with its Radon-Nikodým derivative $\frac{dz_c}{dP} \in L_{\hat{\Phi}}(P)$ and we write its action on $f \in L_{\hat{u}}(P)$ as

$$\begin{aligned} z_c(f) &= \int_\Omega f\, dz_c \\ &= \mathbb{E}_{z_c}[f]. \end{aligned}$$

Now that we have defined and characterised the appropriate Orlicz space and its dual, we can return to the optimisation problem.

3.5.3 The minimax theorem

The minimax theorem, established by John von Neumann [155], is a decision rule used in decision theory, statistics and philosophy for minimising the maximum possible loss. Informally, the minimax theorem states that for every two-person, zero-sum game with finite strategies, there exists a value V and a mixed strategy for each player, such that

(a) given player 2's strategy, the best payoff possible for player 1 is V, and

(b) given player 1's strategy, the best payoff possible for player 2 is -V.

The name minimax arises because each player minimises the maximum payoff possible for the other, and, since the game is zero-sum, he also maximises his own minimum payoff.

Before we can generalise the minimax theorem to Orlicz spaces, some theory is needed. Firstly, we need to state the monotone convergence theorem and an extension thereof, as we will require these theorems in the proof of a lemma, that allows us to change the set over which we are optimising.

Theorem 3.5.9 (Monotone convergence theorem)**.** *Consider the measure space* $(\Omega, \mathcal{F}, P)$*. Let* (h_i) *be an increasing sequence of nonnegative Borel measurable functions, and let* $h(\omega) = \lim_{n\to\infty} h_n(\omega)$ *for* $\omega \in \Omega$*. Then*

$$\int_\Omega h_n \, dP \to \int_\Omega h \, dP.$$

This theorem is restricted to nonnegative sequences. If the nonnegativity assumption is dropped, then the following extension, taken from [8, Theorem 1.6.7], allows one to make the same conclusion.

Theorem 3.5.10 (Extended monotone convergence theorem)**.** *Consider the measure space* $(\Omega, \mathcal{F}, P)$*. Let* $g_1, g_2, \ldots, g, h$ *be Borel measurable.*

(i) If $h \leq g_n$ *for all* n*, where* $\int_\Omega h \, dP > -\infty$ *and* $g_n \uparrow g$*, then*

$$\int_\Omega g_n \, dP \uparrow \int_\Omega g \, dP.$$

(ii) If $g_n \leq h$ *for all* n*, where* $\int_\Omega h \, dP < \infty$ *and* $g_n \downarrow g$*, then*

$$\int_\Omega g_n \, dP \downarrow \int_\Omega g \, dP.$$

The following lemma is also required to prove that we can change the set over which we are optimising.

Lemma 3.5.11. *Let $W \neq 0$ and $W \in L^+_{\hat{u}}(P)$. If $n \in \mathbb{N}$ and $f \in K^W$, then*

$$f \wedge n\mathbf{1} \in (K^W - L_0^+(P)) \cap L_{\hat{u}}(P).$$

Proof. Since $f \in K^W$, there exists $c > 0$ such that $f \geq -cW$, i.e. $-f \leq cW$. But $cW > 0$, and hence $f^- = (-f) \vee 0 \leq cW$. Since $\hat{u}$ is an increasing function and $W \in L^+_{\hat{u}}(P)$,

$$\hat{u}(f^-) \leq \hat{u}(cW) \Rightarrow \int \hat{u}(f^-)\, dP \leq \int \hat{u}(cW)\, dP < \infty,$$

showing that $f^- \in L^+_{\hat{u}}(P)$. It follows from

$$0 \leq f^+ \wedge n\mathbf{1} \leq n\mathbf{1},$$

that $f^+ \wedge n\mathbf{1} \in L_\infty(P) \subseteq L_{\hat{u}}(P)$. Using $f^+ \wedge n\mathbf{1}, f^- \in L_{\hat{u}}(P)$ and the identity

$$f \wedge n\mathbf{1} = f^+ \wedge n\mathbf{1} - f^-$$

established in Lemma A.1.10, it follows that $f \wedge n\mathbf{1} \in L_{\hat{u}}(P)$. Furthermore,

$$f \wedge n\mathbf{1} = f - (f - f \wedge n\mathbf{1}),$$

where $f \in K^W$ and $f - f \wedge n\mathbf{1} \in L_0^+(P)$. Hence, $f \wedge n\mathbf{1} \in K^W - L_0^+(P)$. Therefore,

$$f \wedge n\mathbf{1} \in (K^W - L_0^+(P)) \cap L_{\hat{u}}(P),$$

which completes the proof. □

The next lemma proves that we can change the set over which we are optimising.

Lemma 3.5.12. *Let $W \neq 0$ and $W \in L^+_{\hat{u}}(P)$ and let x be such that $-\infty < \mathbb{E}[\, u(x)\,] < \infty$. Then*

$$\sup_{f \in K^W} \mathbb{E}[\; u(x+f)\;] = \sup_{f^* \in (K^W - L_0^+(P)) \cap L_{\hat{u}}(P)} \mathbb{E}[\; u(x+f^*)\;].$$

Proof. Since $f \in K^W$ can be written as $f = f - 0$, it follows that $K^W \subseteq K^W - L_0^+(P)$, which implies that

$$\sup_{f \in K^W} \mathbb{E}[\, u(x+f)\,] \leq \sup_{f \in K^W - L_0^+(P)} \mathbb{E}[\, u(x+f)\,].$$

Let $f \in K^W - L_0^+(P)$, then f can be written as $f = k - l$, where $k \in K^W$ and $l \in L_0^+(P)$. Then

$$\begin{aligned}\mathbb{E}[\, u(x+f)\,] &= \mathbb{E}[\, u(x+k-l)\,] \\ &\leq \mathbb{E}[\, u(x+k)\,] \\ &\leq \sup_{f \in K^W} \mathbb{E}[\, u(x+f)\,].\end{aligned}$$

Consequently,

$$\sup_{f\in K^W - L_0^+(P)} \mathbb{E}[\,u(x+f)\,] \leq \sup_{f\in K^W} \mathbb{E}[\,u(x+f)\,].$$

Hence,

$$\begin{aligned}\sup_{f\in K^W} \mathbb{E}[\,u(x+f)\,] &= \sup_{f\in K^W - L_0^+(P)} \mathbb{E}[\,u(x+f)\,] \\ &\geq \sup_{f^*\in (K^W - L_0^+(P))\cap L_{\hat{u}}(P)} \mathbb{E}[\,u(x+f^*)\,].\end{aligned}$$

To prove the reverse inequality we need to use the extension of the monotone convergence theorem, stated above in Theorem 3.5.10. Firstly, note that by Lemma A.1.10

$$f \wedge n\mathbf{1} \uparrow f,$$

since $\mathbf{1}$ is a weak order unit of $L_0(P)$. Hence, $x+f\wedge n\mathbf{1} \uparrow x+f$ and, since u is left-continuous, $u(x+f\wedge n\mathbf{1}) \uparrow u(x+f)$.

But $0 \in K^W$, so

$$\sup_{f\in K^W} \mathbb{E}[\,u(x+f)\,] \geq \mathbb{E}[\,u(x)\,] > -\infty.$$

Pick any $f \in K^W$ satisfying $\mathbb{E}[\,u(x+f)\,] > -\infty$. Consider $f_n = f \wedge n$, which is in $(K^W - L_0^+(P)) \cap L_{\hat{u}}(P)$ by Lemma 3.5.11. Then

$$\begin{aligned}u(x+f_n) &= u(x+f^+\wedge n)\mathbb{1}_{\{f\geq 0\}} + u(x-f^-)\mathbb{1}_{\{f<0\}} \\ &\geq u(x)\mathbb{1}_{\{f\geq 0\}} + u(x-f^-)\mathbb{1}_{\{f<0\}}.\end{aligned}$$

Let $h = u(x)\mathbb{1}_{\{f\geq 0\}} + u(x-f^-)\mathbb{1}_{\{f<0\}}$ and $g_n = u(x+f_n)$. Then $h \leq g_n$ for all n and $\mathbb{E}[h] \geq -\infty$. The extended monotone convergence theorem gives

$$\mathbb{E}[\,u(x+f_n)\,] \uparrow \mathbb{E}[\,u(x+f)\,],$$

and since $f_n = f\wedge n \in (K^W - L_0^+(P)) \cap L_{\hat{u}}(P)$,

$$\sup_{f\in K^W} \mathbb{E}[\,u(x+f)\,] \leq \sup_{f^*\in (K^W - L_0^+(P))\cap L_{\hat{u}}(P)} \mathbb{E}[\,u(x+f^*)\,],$$

which completes the proof. □

Therefore, given a loss variable $W \in L_{\hat{u}}^+(P)$, we define

$$C^W(P) = (K^W - L_0^+(P)) \cap H_{\hat{u}}(P).$$

This cone represents those random variables that can be super-replicated by trading strategies in $\mathcal{H}^W$ and have the same type of boundedness as W.

We also define the concave function $K_u : L_{\hat{u}}(P) \to [-\infty, \infty)$ by

$$K_u(f) = \mathbb{E}[\,u(f)\,]$$

and let $\mathcal{D}$ be its proper domain, i.e.

$$\mathcal{D} = \{f \in L_{\hat{u}}(P) : \mathbb{E}[\,u(f)\,] > -\infty\}.$$

Proposition C.13 can be used to characterise the above function. Note that a similar result holds for concave functions. This proposition allows us to prove the following proposition characterising K_u.

Proposition 3.5.13. *The concave function K_u on $L_{\hat{u}}(P)$ is proper and it is norm-continuous on the interior of its proper domain, which is not empty. Moreover, there exists a norm continuity point of K_u that belongs to $C^W(P)$.*

Proof. [15, Proposition 16] By Proposition C.13, the first statement is equivalent to the existence of a non-empty open set $\mathcal{O}$ on which K_u is not everywhere equal to ∞ and is bounded below by a constant $c \in \mathbb{R}$.

Firstly, we show that on the unit ball U of $L_{\hat{u}}(P)$, the function K_u is everywhere less than ∞. If $x \in U$, then by Jensen's inequality and the fact that $L_{\hat{u}}(P) \subseteq L_1(P)$, it follows that $|\mathbb{E}[x|]| \leq \mathbb{E}[\,|x|\,] < \infty$. So, using Jensen's inequality again and the definition of the utility function, $K_u(x) = \mathbb{E}[\,u(x)\,] \leq \mathbb{E}[\,u(|x|)\,] \leq u(\mathbb{E}[x]) < \infty$.

Secondly, we show that K_u is uniformly bounded below on the unit ball $U \in L_{\hat{u}}(P)$. For all $x \in U$, $\mathbb{E}[\,\hat{u}(x)\,] \leq 1$ and $\mathbb{E}[\,\hat{u}(x^-)\,] \leq 1$. Hence

$$\begin{aligned}
-K_u(-x^-) &= -\mathbb{E}[\,u(-x^-)\,] \\
&= \mathbb{E}[\,-u(-x^-)\,] \\
&= \mathbb{E}[\,\hat{u}(x^-) - u(0)\,] \\
&= \mathbb{E}[\,\hat{u}(x^-)\,] - u(0) \\
&\leq 1 - u(0)
\end{aligned}$$

and so $K_u(x) \geq K_u(-x^-) \geq u(0) - 1$, i.e. K_u is uniformly bounded below.

The second statement follows from the following. Let $z \in -U^+$, then $z \in L_{\hat{u}}(P)$. Also, z can be written as $z = 0 - (-z)$, which shows that $z \in K^W - L_0^+(P)$. Hence, $z \in C^W(P)$ and $-U^+ \subseteq C^W(P)$. □

Note that the above proof shows that K_u is finite on the unit ball U.

Proposition 3.5.14. *If $z \in (L_{\hat{u}}(P))^*_s$ with $z \geq 0$, then*

$$\|z\|_{(L_{\hat{u}})^*} = \sup\{z(f) : f \geq 0, f \in L_{\hat{u}}(P) \text{ such that } \mathbb{E}[\,\hat{u}(f)\,] < \infty\}. \tag{3.7}$$

Proof. By definition,

$$\|z\|_{(L_{\hat{u}})^*} = \sup\{\|z(f)\| : \mathcal{N}_{\hat{u}}(f) \leq 1, f \geq 0\}.$$

Since z is nonnegative and using Theorem 2.1.8, we can complete the proof. □

For the convenience of the reader, we recall that the utility function $u : \mathbb{R} \to [-\infty, \infty]$ is increasing and concave on (a, ∞), where $a \in [-\infty, 0)$.

Lemma 3.5.15. *Let $z \in ((L_{\hat{u}}(P))^*_s)^+$. Then*

$$\|z\|_{(L_{\hat{u}})^*} = \sup_{f \in \mathcal{D}} z(-f).$$

In the case where a is finite,

$$\|z\|_{(L_{\hat{u}})^*} = -az(\Omega).$$

Proof. Since $z \geq 0$,

$$\sup_{f \in \mathcal{D}} z(-f) = \sup_{(-f) \geq 0,\, f \in \mathcal{D}} z(-f).$$

Now, $f \in \mathcal{D}$ implies that $\mathbb{E}[u(f)] > -\infty$, which is equivalent to $\mathbb{E}[\hat{u}(-f)] < \infty$, where $-f$ is nonnegative. Therefore,

$$\sup_{f \in \mathcal{D}} z(-f) = \sup\{z(-f) : -f \geq 0, (-f) \in L_{\hat{u}}(P) \text{ such that } \mathbb{E}[\hat{u}(-f)] < \infty\},$$

and using Proposition 3.5.14 completes the first part of the proof.

For a finite, we have

$$\|z\|_{(L_{\hat{u}})^*} = \sup\{z(f) : f \in L^+_{\hat{u}}(P),\, f < -a\} = -az(\Omega),$$

as $\sup_{f \in L^+_{\hat{u}}(P),\, f < \mathbf{1}} z(f) = z(\Omega)$. □

To describe the dual variables, we define the *negative polar cone* by

$$(C^W(P))^\circ_- = \{z \in (H_{\hat{u}}(P))^* = L_{\hat{\Phi}}(P) : z(f) \leq 0 \text{ for all } f \in C^W(P)\}$$

and the subset of normalised functions in $(C^W(P))^\circ_-$ by

$$\mathcal{M}^W(P) = \{Q \in (C^W(P))^\circ_- : Q(\mathbb{1}_\Omega) = 1\}.$$

Thus, the elements of $\mathcal{M}^W(P)$ are probability measures, which are absolutely continuous with respect to P.

Remark 3.5.16. By definition, if $Q \in \mathcal{M}^W(P)$ and $f \in C^W(P)$, then $\mathbb{E}_Q[f] \leq 0$.

In general,

$$\mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P) \subseteq \mathcal{M}^W(P),$$

and it can be shown, see [15], that if $W \in \mathbb{S} \cap H_{\hat{u}}(P)$, then

$$\mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P) = \mathcal{M}^W(P). \tag{3.8}$$

Next, we will state the special case of Rockafellar and Kozek's theorem, Theorem 2.6.2, required for the proof of Theorem 3.5.18.

Theorem 3.5.17. *Suppose that $F : \mathbb{R} \to (-\infty, \infty]$ and $F^* : \mathbb{R} \to (-\infty, \infty]$ are convex, lower semi-continuous conjugate functions, not identically equal to ∞, and that there exists $f \in L_{\hat{u}}(P)$ such that $K_F(f) = \mathbb{E}[F(f)] < \infty$. If $K_{F^*}(g) < \infty$ for some $g \in L_{\hat{\Phi}}(P)$, then the convex conjugate $K_F^* : (L_{\hat{u}}(P))^* \to (-\infty, \infty]$ of K_F is given by*

$$K_F^*(z) = K_{F^*}\Big(\frac{dz_r}{dP}\Big) + \sup\{z_s(f) : f \in \mathrm{dom}(K_F)\}.$$

Now we can state and prove the required version of the minimax theorem, as proved by Biagini and Frittelli (see [15, Theorem 21]).

Theorem 3.5.18. *Let $u : \mathbb{R} \to [-\infty, \infty]$ be increasing and concave on the interior (a, ∞), $a \in [-\infty, 0)$, of its effective domain and let $\lim_{x \to -\infty} u(x) = -\infty$.*

(i) If there exists $W \in L_{\hat{u}}^+(P)$ such that $\sup_{f \in K^W} \mathbb{E}[\, u(x + f)\,] < \lim_{y \to \infty} u(y)$ for some $x > a$, then $\mathcal{M}^W$ is not empty and

$$\begin{aligned} U^W(x) &:= \sup_{H \in \mathcal{H}^W} \mathbb{E}[\, u(x + (H \cdot S)_T)\,] \\ &= \sup_{f \in K^W} \mathbb{E}[\, u(x + f)\,] \\ &= \sup_{f^* \in C^W(P)} \mathbb{E}[\, u(x + f^*)\,] \\ &= \min_{\lambda > 0, Q \in \mathcal{M}^W(P)} \Big\{\lambda(x + \|Q_s\|) + \mathbb{E}[\, \Phi\Big(\lambda \frac{dQ_c}{dP}\Big)\,]\Big\}, \end{aligned} \tag{3.9}$$

where $Q = Q_c + Q_s$ is the decomposition of Q into its σ-order continuous and singular part.

(ii) If $W \in H_{\hat{u}}(P)$, then the set $\mathcal{M}^W(P)$ can be replaced by the set $\mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P)$ and no singular term appears in the duality formula above.

Proof. (i): Firstly, we will prove the result for $x = 0$. Using Theorem 3.5.17, the concave conjugate function of K_u is given by

$$\begin{aligned} J_u(z) &= -K_\Phi(z_c) - \sup_{f\in\mathcal{D}}\{z_s(-f)\} \\ &= -\mathbb{E}[\,\Phi(z_c)\,] - \sup_{f\in\mathcal{D}}\{z_s(-f)\} \\ &= -\mathbb{E}[\,\Phi(\frac{dz_c}{dP})\,] - \sup_{f\in\mathcal{D}}\{z_s(-f)\}. \end{aligned}$$

From Lemma 3.5.15,

$$J_u(z) = -\mathbb{E}[\,\Phi(\frac{dz_c}{dP})\,] - \|z_s\|\,.$$

Due to Proposition 3.5.13, we know that an interior point of $C^W(P)$ exists and hence, we can apply the Fenchel duality theorem, or more specifically we can use Equation (C.1), to get

$$\begin{aligned} \sup_{f\in C^W(P)} K_u(f) &= \sup_{f\in C^W(P)} \mathbb{E}[\,u(f)\,] \\ &= \min_{z\in(C^W(P))^\circ} -J_u(z) \\ &= \min_{z\in(C^W(P))^\circ} \Big\{\mathbb{E}[\,\Phi\Big(\frac{dz_c}{dP}\Big)\,] + \|z_s\|\,\Big\}. \end{aligned} \tag{3.10}$$

Let $\tilde{z} \in (C^W(P))^\circ$ be such that

$$\min_{z\in(C^W(P))^\circ} \Big\{\mathbb{E}[\,\Phi\Big(\frac{dz_c}{dP}\Big)\,] + \|z_s\|\,\Big\} = \mathbb{E}[\,\Phi\Big(\frac{d\tilde{z}_c}{dP}\Big)\,] + \|\tilde{z}_s\|\,.$$

We next prove the claim that $\mathcal{M}^W(P)$ is not empty. Assume that $\tilde{z}_c = 0$. Then

$$\sup_{f\in C^W(P)} \mathbb{E}[\,u(f)\,] = \Phi(0) + \|\tilde{z}_s\| \geq \lim_{x\to\infty} u(x),$$

since $\Phi(0) = \lim_{x\to\infty} u(x)$. This contradicts the assumption that $\sup_{k\in K^W} \mathbb{E}[u(k)] < \lim_{x\to\infty} u(x)$. Hence, $\tilde{z}_c \neq 0$. Let $\hat{z} = \frac{\tilde{z}_c}{\|\tilde{z}_c\|}$, then $\|\hat{z}\| = 1$ and consequently $\frac{d\hat{z}}{dP} \in \mathcal{M}^W(P)$, i.e. $\mathcal{M}^W(P)$ is not empty.

Returning to (3.10) and reparametrising our optimisation problem via $\mathcal{M}^W$, which consist of all the functions of norm 1, we have that

$$\sup_{f\in C^W(P)} K_u(f) = \min_{\lambda>0, Q\in\mathcal{M}^W(P)} \Big\{\mathbb{E}[\,\Phi\Big(\lambda\frac{dQ_c}{dP}\Big)\,] + \lambda\,\|Q_s\|\,\Big\}.$$

Next we look at the case where $x \neq 0$. Let $u_x(f) = u(x+f)$. Then u_x is finite on (a_x, ∞), where $a_x = a - x < 0$, and K_{u_x} is given by $K_{u_x}(f) = \mathbb{E}[u(x+f)]$ with $\mathcal{D}_x = \mathcal{D} - x$

as its domain. Now, we can use the same method as when $x = 0$, noting that the conjugate of u_x is given by $\Phi_x(g) = xg + \Phi(g)$ and thus, the concave conjugate of K_{u_x} is given by

$$J_{u_x}(z) = -xz_c(\Omega) - \mathbb{E}[\Phi(\frac{dz_c}{dP})] - \sup_{f \in \mathcal{D}_x} z_s(-f).$$

However, $\sup_{f\in\mathcal{D}_x} z_s(-f) = \sup_{g\in\mathcal{D}} z_s(-g) + xz_s(\Omega)$, so

$$J_{u_x}(z) = -\mathbb{E}[\Phi(\frac{dz_c}{dP})] - xz(\Omega) - \|z_s\|.$$

Using this in Fenchel's theorem and simplifying as in the case where $x = 0$, completes the proof.

(ii): The singular term disappears due to the fact that the dual of $H_{\hat{u}}(P)$, which is $L_{\hat{\Phi}}(P)$, has no singular part. Using (3.8), we can replace $\mathcal{M}^W(P)$ by $\mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P)$. □

The following corollary is a reformulation of Theorem 3.5.18, required for the preference-dependent version of the first fundamental theorem of asset pricing presented in the next section. The constant initial endowment x is replaced by a random variable B, which is assumed to be bounded. For a more general version of Theorem 3.5.18, where the random variable is not necessarily bounded, the reader is referred to [17].

If $\mathbb{S} \cap H_{\hat{u}}(P)$ is not empty, define

$$\mathcal{H}^{\hat{u}} = \bigcup_{W\in\mathbb{S}\cap H_{\hat{u}}(P)} \mathcal{H}^W.$$

Then, $\mathcal{H}^{\hat{u}}$ does not depend on any particular $W \in \mathbb{S} \cap H_{\hat{u}}(P)$.

Corollary 3.5.19. *Assume $B \in L_\infty(P)$ and suppose that there exists $W \in H_{\hat{u}}(P)$ with $W \geq 1$ and satisfying*

$$\sup_{H\in\mathcal{H}^W} \mathbb{E}_P[\, u(B + (H\cdot S)_T)\,] < \lim_{y\to\infty} u(y).$$

Then

(i) $\mathcal{M}^W(P)$ is not empty and

$$\begin{aligned} U^W(B) &:= \sup_{H\in\mathcal{H}^W} \mathbb{E}_P[\, u(B + (H\cdot S)_T)\,] \\ &= \min_{\lambda>0, Q\in\mathcal{M}^W(P)} \Big\{ \lambda\mathbb{E}_Q[B] + \mathbb{E}_P[\,\Phi(\lambda\frac{dQ}{dP})\,]\Big\}. \end{aligned}$$

(ii) If in addition, $W \in \mathbb{S} \cap H_{\hat{u}}(P)$, then $\mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P)$ is not empty and $\mathcal{H}^W$ can be replaced by $\mathcal{H}^{\hat{u}}$ and $\mathcal{M}^W(P)$ by $\mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P)$.

Now that we have stated and proved the minimax theorem, we can state and prove Frittelli's new version of the fundamental theorem of asset pricing.

3.5.4 No market free lunch*

In this section, we specialise the definition of a market free lunch to a subset $\mathcal{P}_{\hat{u}}$ of the probability measures $\mathbb{P}$. Define

$$\mathcal{P}_{\hat{u}} = \{Q \sim P : \mathbb{S} \cap H_{\hat{u}}(Q) \neq \emptyset\}.$$

Then, similar to Section 3.4, Frittelli [64] defines a market free lunch* with respect to $\mathbb{U}$, which will be based on the Orlicz space.

Definition 3.5.20. There is a *market free lunch** with respect to $\mathbb{U}$ if for all $P \in \mathcal{P}_{\hat{u}}$ and $u \in \mathbb{U}$, there exists $w \in L^+_\infty(P)\backslash\{0\}$ such that

$$\sup_{f \in \mathcal{H}^{\hat{u}}} \mathbb{E}_P[\, u(f - w)\,] \geq u(0). \tag{3.11}$$

Hence, there is *no market free lunch** (NMFL*($\mathbb{U}$)) with respect to $\mathbb{U}$ if for all $w \in L^+_\infty(P)\backslash\{0\}$, there exists $P \in \mathcal{P}_{\hat{u}}$ and $u \in \mathbb{U}$ such that

$$\sup_{f \in \mathcal{H}^{\hat{u}}} \mathbb{E}_P[\, u(f - w)\,] < u(0). \tag{3.12}$$

We need the following result, due to Halmos and Savage [75] to prove the preference dependent version of the first fundamental theorem of asset pricing.

Theorem 3.5.21 (Halmos-Savage theorem)**.** *Let M be a set of P-absolutely continuous probability measures on $\mathcal{F}$, which is closed under countable convex combinations. Suppose that for each set $A \in \mathcal{F}$ with $P(A) > 0$, there exists $Q \in M$ with $Q(A) > 0$. Then, there exists Q_0 such that, for all sets $A \in \mathcal{F}$ with $P(A) > 0$, $Q_0(A) > 0$; that is, Q_0 and P are equivalent probability measures.*

Theorem 3.5.22. *NMFL*($\mathbb{U}_2$) $\Longleftrightarrow \mathbb{M}_\sigma(S) \cap \mathbb{P} \neq \emptyset$.*

Proof. $\Leftarrow$: Assume there exists $Q \in \mathbb{M}_\sigma(S) \cap \mathbb{P}$. We need to show that for all $w \in L^+_\infty\backslash\{0\}$, there exist $P \in \mathcal{P}_{\hat{u}}$ and $u \in \mathbb{U}_2$ such that (3.12) holds. By Proposition A.8.6, since Q is a sigma-martingale measure, there exist a d-dimensional Q-martingale N and a positive predictable N-integrable process φ such that $S^i = \varphi \cdot N^i$. Then $N^i = \varphi^{-1} \cdot S^i$ is a martingale and its maximal process $(N^i)^* = \sup_{t\le T} |N^i_t|$ is Q-integrable. Let

$$W = \mathbf{1} + \sum_{i=1}^{d} \sup_{t\le T}\{|(\varphi^{-1}\cdot S^i)_t|\}.$$

Then $W \in L_1(Q)$. By definition $|(\varphi^{-1}\cdot S^i)_t| \le W$ for all t and i, where φ^{-1} is a positive predictable S^i-integrable process. Hence W is suitable.

Consider $u \in \mathbb{U}_2$ such that it is strictly increasing and $\lim_{x\to-\infty} u(x) = x$. It follows that $\lim_{x\to\infty} \hat{u}(x) = x + u(0)$. Since $\hat{u}(0) = 0$, we must have that

$$\hat{u}(x) \le \begin{cases} x + u(0) & \text{if } u(0) \ge 0 \\ x & \text{if } u(0) < 0. \end{cases}$$

Now, let $f \in L_1(Q)$, i.e. $\int |f|\, dQ < \infty$. Then, for all $a > 0$,

$$\int \hat{u}\Big(\frac{|f|}{a}\Big)\, dQ \le \begin{cases} \int \Big(\frac{|f|}{a} + u(0)\Big)\, dQ < \infty & \text{if } u(0) \ge 0 \\ \int \frac{|f|}{a}\, dQ < \infty & \text{if } u(0) < 0, \end{cases}$$

i.e. $f \in H_{\hat{u}}(Q)$. Hence, in conjunction with Proposition 2.2.2, $H_{\hat{u}}(Q) = L_1(Q)$.

Since W is integrable with respect to Q, we have that $W \in L_1(Q) = H_{\hat{u}}(Q)$, which shows that W is compatible. Therefore, $W \in H_{\hat{u}}(Q) \cap \mathbb{S}$ and by definition, $Q \in \mathcal{P}_{\hat{u}}$.

Since $W \in H_{\hat{u}}(Q) \cap \mathbb{S}$, we have by Equation (3.8) that

$$\mathcal{M}^W(Q) = \mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(Q) = \mathbb{M}_\sigma(S) \cap L_\infty(P).$$

The last equality follows from the fact that $H_{\hat{u}}(Q) = L_1(Q)$ and $(L_1(Q))^* = L_\infty(Q)$. Since $Q \sim P$, we also have that $L_\infty(Q) = L_\infty(P)$.

Since $\mathbf{1} \in L_{\hat{\Phi}}(P)$, we have that $Q \in \mathbb{M}_\sigma(S) \cap L_\infty(P)$ and hence, $Q \in \mathcal{M}^W(Q)$. Thus, using Remark 3.5.16, it follows that $\mathbb{E}_Q[f] \le 0$ for all $f \in C^W(Q)$.

Let $g \in L^+_\infty(P)\backslash\{0\}$. Using Jensen's inequality and the fact that u is increasing,

$$\mathbb{E}_Q[\,u(f-g)\,] \le u(\mathbb{E}_Q[f] - \mathbb{E}_Q[g]) \le u(-\mathbb{E}_Q[g]) \le u(0). \tag{3.13}$$

Thus, we can deduce that

$$\sup_{f\in C^W(Q)} \mathbb{E}_Q[\,u(f-g)\,] < u(0),$$

which implies

$$\sup_{f\in \mathcal{H}^{\hat{u}}(Q)} \mathbb{E}_Q[\,u(f-g)\,] < u(0),$$

completing the first part of the proof.

$\Rightarrow$: Assume that for all $g \in L_\infty^+ \backslash \{0\}$ there exist $u \in \mathbb{U}_2$, $P \in \mathcal{P}_{\hat{u}}$ and $W \in H_{\hat{u}}(P) \cap \mathbb{S}$ such that

$$\sup_{f \in C^W(P)} \mathbb{E}_P[\, u(f-g)\,] < u(0) \leq u(\infty).$$

From Corollary 3.5.19, we can deduce that there exist $Q_g \in \mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P)$ and $\lambda_g > 0$ that attain the minimum in the dual problem, i.e.

$$\begin{aligned} u(0) > & \sup_{f \in \mathcal{H}^{\hat{u}}(P)} \mathbb{E}_P[\, u(f-g)\,] \\ & = \min_{\lambda > 0, Q \in \mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P)} \Big\{ -\lambda \mathbb{E}_Q[g] + \mathbb{E}_P[\Phi\Big(\lambda \frac{dQ}{dP}\Big)] \Big\} \\ & = -\lambda_g \mathbb{E}_{Q_g}[g] + \mathbb{E}_P[\Phi\Big(\lambda_g \frac{dQ_g}{dP}\Big)] \\ & \geq -\lambda_g \mathbb{E}_{Q_g}[g] + u(0), \end{aligned}$$

where the last inequality follows from

$$\Phi(y) = \sup_{x \in \mathbb{R}} \{u(x) - xy\} \geq u(0) \text{ for all } y \geq 0.$$

In particular, let $g = \mathbb{1}_A$, where $A \in \mathcal{F}$ satisfies $P(A) > 0$. Using the above, we have that $0 > -\lambda_{\mathbb{1}_A} Q_{\mathbb{1}_A}(A)$. In other words, for all $A \in \mathcal{F}$ with $P(A) > 0$, there exists $Q_{\mathbb{1}_A} \in \mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P)$ such that $Q_{\mathbb{1}_A}(A) > 0$. By the Halmos-Savage theorem, there exists $Q_0 \in \mathbb{M}_\sigma(S) \cap L_{\hat{\Phi}}(P) \subseteq \mathbb{M}_\sigma(S)$ such that $Q_0 \sim P$, i.e. $\mathbb{M}_\sigma(S) \cap \mathbb{P} \neq \emptyset$. □

Corollary 3.5.23. *NMFL*($\mathbb{U}_2$) $\Leftrightarrow$ *NFLVR*

Proof. This follows easily from Theorem 3.1.8 and Theorem 3.5.22. □

Klein [100] states that NMFL($\mathbb{U}_2$) $\Leftrightarrow$ NFL. In our opinion there is a gap in the first part of the proof, as we cannot verify the first inequality on page 4 line 17. However, using the above corollary along with Theorem 3.1.9 proves her statement.

Chapter 4

Scalar-valued risk measures on Orlicz hearts

It is important, both theoretically and practically, to be able to quantify the risk involved in a financial position and hence, decide if it is acceptable or not. For this reason, different ways to quantify risk and various risk measures have been proposed in the literature.

Markowitz proposed to use the variance of the return distribution to measure the risk associated with each investment, and in the case of a portfolio, i.e. a combination of assets, to look at the covariance between all pairs of investments. The main innovation by Markowitz was to use the joint distribution of returns of all assets. A joint distribution is characterised by the marginal properties of each component random variable and by their dependence structure. Markowitz used the mean and the variance to describe the marginal properties and the linear correlation coefficient between each pair to describe the dependence structure. However, the only class of random variables for which the linear correlation coefficient can be used as a dependence measure, is the class of elliptic distributions, like the normal or t-distribution with finite variances.

Another disadvantage of using the variance as a risk measure, is the asymmetry in the financial interpretation, i.e. only the downside risk plays a role, whereas the variance gives both upside and downside risk.

A popular risk measures is *value at risk* (VaR). VaR has been around since the early 1900's but JP Morgan was the first to coin the term 'value at risk' in a report by 30 people entitled: *Derivatives: Practices and Principles* [92]. VaR with parameter α answers the following question: what are your expected losses in one day (week, year, ...) with a given probability α. In other words, what percentage of your investment is at risk? The disadvantages of VaR are that it is model dependent and that it does not decrease when we diversify the portfolio. See Section 4.4 for more details.

This lead Artzner et al. [4] to introduce an axiomatic definition of a risk measure, which they called a coherent risk measure. Their results are based in a finite probability space. This was later extended to an infinite probability space by Delbaen [42]. Independently around the same time, Wang et al. [158] published similar results on the closely related topic of insurance premia.

The theory of risk measures can be split into two parts: static risk measures, like value at risk and coherent risk measures, and dynamic risk measures, which were proposed by Cvitanic and Karatzas [33] and Wang [163]. In this chapter, we will only look at static risk measures.

4.1 Definition

Measuring risk is equivalent to establishing a correspondence ρ between the space $\mathcal{X}$ of random variables (for instance the returns of a given set of investments) and the real numbers, i.e. risk measures are defined as functions from some set of possible scenarios to the reals and they need to satisfy certain conditions of consistency, as defined below.

Definition 4.1.1. Let $\mathcal{X}$ be a vector subspace of $L_0(P)$ that contains all constant functions. A map $\rho : \mathcal{X} \to (-\infty, \infty]$ is called a *monetary risk measure* on $\mathcal{X}$ if it has the following properties:

(i) Finiteness at 0, i.e. $\rho(0) \in \mathbb{R}$.

(ii) Monotonicity, i.e. if $x \leq y$, then $\rho(x) \geq \rho(y)$ for all $x, y \in \mathcal{X}$.

(iii) Translation invariance, i.e. if $m \in \mathbb{R}$, then $\rho(x + m\mathbf{1}) = \rho(x) - m$ for all $x \in \mathcal{X}$, where $\mathbf{1}(\omega) = 1$ a.e. for all $\omega \in \Omega$.

The monetary risk measure is called *coherent* if it also satisfies

(iv) Positive homogeneity, i.e. if $\lambda \geq 0$, then $\rho(\lambda x) = \lambda \rho(x)$ for all $x \in \mathcal{X}$.

(v) Subadditivity, i.e. $\rho(x + y) \leq \rho(x) + \rho(y)$ for all $x, y \in \mathcal{X}$.

The translation invariance property (also known as the cash additivity property) implies that by adding a sure return to a random return, the risk will decrease by that sure amount.

If we are working with a loss variable x, then (iii) becomes $\rho(x + m\mathbf{1}) = \rho(x) + m$ for all $m \in \mathbb{R}$. From here onwards we will write only m for $m\mathbf{1}$ as is customary in the literature.

Subadditivity seems like a natural requirement as diversification should not increase the risk. Artzner et al. [6] explain it by the phrase 'a merger does not create extra risk'. Note that subadditivity implies $\rho(\lambda x) \leq \lambda\rho(x)$. The positive homogeneity gives us the other inequality. Artzner et al. [5, 6] justify the latter inequality by liquidity considerations. An investment λX would be less liquid and hence, more risky than λ smaller investments.

In many situations, however, the risk of a position might increase in a non-linear manner with the size of the position. This suggests that the conditions of positive homogeneity and subadditivity should be relaxed. Therefore, Heath [84] introduced convex risk measures in a finite probability space, which was later independently generalised by Föllmer and Schied [60] and Frittelli and Rosazza Gianin [62] to all probability spaces. Convex risk measures are defined as follows.

Definition 4.1.2. A map $\rho : \mathcal{X} \to (-\infty, \infty]$ is called a *convex risk measure* if it satisfies the conditions of monotonicity, translation invariance and

(vi) Convexity, i.e. $\rho(\lambda x + (1-\lambda)y) \leq \lambda\rho(x) + (1-\lambda)\rho(y)$ for any $\lambda \in (0, 1)$ and for all $x, y \in \mathcal{X}$.

Note that coherent risk measures are special cases of convex risk measures. It can be shown that any positively homogeneous function is convex if and only if it is subadditive. Convex risk measures are sometimes called *weakly coherent*.

If there is uncertainty in the interest rate, then El Karoui and Ravanelli [57] suggest replacing cash-additivity with cash-subadditivity.

Definition 4.1.3. A map $\rho : \mathcal{X} \to (-\infty, \infty]$ is *cash-subadditive* if for all $x \in \mathcal{X}$ and $m \in \mathbb{R}$

$$\rho(x + m) \geq \rho(x) - m. \tag{4.1}$$

Cash-subadditivity implies that when m is subtracted from a future position the present capital requirement cannot be increased by more than m. Note that (4.1) is equivalent to

$$\rho(x - m) \leq \rho(x) + m \text{ for all } x \in \mathcal{X}.$$

Cerreia-Vioglio et al. [22] then suggest that once cash-additivity has been replaced by cash-subadditivity, convexity should be replaced by quasiconvexity, in order to maintain the original interpretation in terms of diversification.

Definition 4.1.4. A map $\rho : \mathcal{X} \to (-\infty, \infty]$ is *quasiconvex* if for any $\lambda \in (0,1)$ and for all $x, y \in \mathcal{X}$

$$\rho(\lambda x + (1-\lambda)y) \leq \max(\rho(x), \rho(y)). \tag{4.2}$$

Cerreia-Vioglio et al. [22] state that quasiconvexity 'allows a complete disentangling between the diversification principle, which is arguably the central pillar of risk management, and the assumption of liquidity of the riskless asset, which is an abstract simplification'. The economic counterpart of quasiconvexity of risk measures is quasiconcavity of utility functions. In this thesis we only consider convex risk measures.

To compare all these different types of risk measures, representation theorems can be helpful. A lot of work has gone into finding the dual representation theorems of risk measures with different properties.

Artzner et al. [5, 6] and Föllmer and Schied [60] modelled future financial positions by elements of the set $L(\Omega)$ of all real-valued functions on a finite sample space Ω and a coherent, convex or monetary risk measure is a map $\rho : L(\Omega) \to \mathbb{R}$ satisfying certain properties. They show that every monetary risk measure can be expressed as

$$\rho(x) = \inf\{m \in \mathbb{R} : x + m \in \mathcal{A}\} \text{ for all } x \in \mathcal{X},$$

where $\mathcal{A} = \{x \in L(\Omega) : \rho(x) \leq 0\}$ is the set of acceptable positions, and that every convex risk measure has a convex dual representation of the form

$$\rho(x) = \sup_{Q \in \mathcal{M}_p} \{\mathbb{E}^Q[-x] - \alpha(Q)\} \text{ for all } x \in \mathcal{X}, \tag{4.3}$$

where $\mathcal{M}_p$ is the set of all probability measures on Ω and α is a function from $\mathcal{M}_p$ to $(-\infty, \infty]$. Note that $\mathcal{M}_p$ is independent of x. If ρ is coherent, then α can be chosen so that it only takes the values 0 or ∞, and thus (4.3) reduces to

$$\rho(x) = \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[-x] \text{ for all } x \in \mathcal{X},$$

where $\mathcal{Q} = \{Q \in \mathcal{M}_p : \alpha(Q) = 0\}$.

Economically, this representation tells us that $\rho(x)$ is the minimal amount which has to be added to a position x to make it acceptable. The function α is called the *penalty function* and it penalises different probabilities, depending on how likely they are to occur. In the literature, the standard proofs of the robust representations involve the separating hyperplane theorem and become more involved in a general framework. It is also noteworthy, that the representation changes depending on the set $L(\Omega)$. If, for example, one looks at $L_\infty(P)$, the robust representation formula involves finitely additive measures and to reduce it to σ-additive measures, additional continuity assumptions are required.

4.2 Acceptance sets

The acceptance set describes all those positions that the regulator or investor deems as acceptable in terms of risk.

Definition 4.2.1. Let $\mathcal{X}$ be a linear subspace of $L_0(P)$ that contains all constant functions. Let $\rho : \mathcal{X} \to (-\infty, \infty]$ be a monetary risk measure. The set $\mathcal{A}_\rho$ given by

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

is called the *acceptance set* of ρ.

The following propositions, taken from Cheridito and Li [26], shows the connection between risk measures and certain sets of random variables.

Proposition 4.2.2. *Let $\mathcal{X}$ be a linear subspace of $L_0(P)$ that contains all constant functions, and $\rho : \mathcal{X} \to (-\infty, \infty]$ a monetary risk measure with acceptance set $\mathcal{A}_\rho$. Then, for $x \in \mathcal{X}$,*

$$\rho(x) = \inf\{m \in \mathbb{R} : x + m \in \mathcal{A}_\rho\}.$$

Furthermore, the following properties are satisfied by $\mathcal{A}_\rho$:

(i) $\inf\{m \in \mathbb{R} : m \geq z$ *for some* $z \in \mathcal{A}_\rho\} \in \mathbb{R}$.

(ii) For all $x \in \mathcal{X}$, $\inf\{m \in \mathbb{R} : x + m \geq z$ *for some* $z \in \mathcal{A}_\rho\} \in (-\infty, \infty]$.

(iii) For all $x \in \mathcal{A}_\rho$, $\{y \in \mathcal{X} : y \geq x\} \subseteq \mathcal{A}_\rho$.

(iv) If (x_n) is a sequence in $\mathcal{A}_\rho$ such that $||x_n - x||_\infty \to 0$ for some $x \in \mathcal{X}$, then $x \in \mathcal{A}_\rho$, i.e. $\mathcal{A}_\rho$ is closed with respect to $||\cdot||_\infty$. Moreover, if ρ is convex, then so is $\mathcal{A}_\rho$. If ρ is coherent, then $\mathcal{A}_\rho$ is a convex cone.

(v) If ρ is real-valued, then $\mathcal{A}_\rho$ has the following property: For all $x \in \mathcal{X}$, $\inf\{m \in \mathbb{R} : x + m \geq z$ *for some* $z \in \mathcal{A}_\rho\} \in \mathbb{R}$.

Proof. Assume $x + m \in \mathcal{A}_\rho$. It follows from the definition of $\mathcal{A}_\rho$ and the translation invariance property that $\rho(x) \leq m$ for all $m \in \mathbb{R}$. Therefore,

$$\rho(x) = \inf\{m \in \mathbb{R} : x + m \in \mathcal{A}_\rho\}.$$

(iii): Let $y_0 \in \{y \in \mathcal{X} : y \geq x\}$. By the monotonicity, $\rho(y_0) \leq \rho(x) \leq 0$ and hence, $y_0 \in \mathcal{A}_\rho$.

(ii): Using (iii), we get that

$$\{m \in \mathbb{R} : x + m \geq z \text{ for some } z \in \mathcal{A}_\rho\} \subseteq \{m \in \mathbb{R} : x + m \in \mathcal{A}_\rho\}.$$

Since $x + m \geq x + m$, we have equality in the above. Therefore,

$$\rho(X) = \inf\{m \in \mathbb{R} : x + m \geq z \text{ for some } z \in \mathcal{A}_\rho\}$$

and since $\rho(x) \in (-\infty, \infty]$, we get (ii).

(i): $\rho(0) = \{m \in \mathbb{R} : m \geq z \text{ for some } z \in \mathcal{A}_\rho\} \in \mathbb{R}$.

(iv): Consider a sequence $(x_n) \subseteq \mathcal{A}_\rho$ with $||x_n - x||_\infty \to 0$ for some $x \in \mathcal{X}$. Then, for all $\epsilon > 0$, there exists $n \geq 1$ such that $x \geq x_n - \epsilon$. By the monotonicity and the translation invariance,

$$\rho(x) \leq \rho(x_n - \epsilon) = \rho(x_n) + \epsilon \leq \epsilon,$$

and thus, $x \in \mathcal{A}_\rho$.

Now, assume that ρ is convex and let $x, y \in \mathcal{A}_\rho$. Then for $\lambda \in (0, 1)$

$$\rho(\lambda x + (1 - \lambda)y) \leq \lambda\rho(x) + (1 - \lambda)\rho(y) \leq 0,$$

i.e. $\lambda x + (1 - \lambda)y \in \mathcal{A}_\rho$. Thus, $\mathcal{A}_\rho$ is convex.

It is obvious that if ρ is coherent, then $\mathcal{A}_\rho$ is a convex cone.

(v): Assume ρ is real-valued. By the proof of (ii), we have that

$$\rho(x) = \inf\{m \in \mathbb{R} : x + m \geq z \text{ for some } z \in \mathcal{A}_\rho\}.$$

□

Proposition 4.2.3. *Let $\mathcal{X}$ be a vector subspace of $L_0(P)$ that contains all constant functions, and $\mathcal{B}$ a subset of $\mathcal{X}$ with properties (i) and (ii) of Proposition 4.2.2.*

(i) Then

$$\rho_{\mathcal{B}}(x) = \inf\{m \in \mathbb{R} : x + m \geq z \textit{ for some } z \in \mathcal{B}\}$$

defines a monetary risk measure on $\mathcal{X}$, whose acceptance set $\mathcal{A}_{\rho_{\mathcal{B}}}$ is the smallest subset of $\mathcal{X}$ that contains $\mathcal{B}$ and satisfies (iii) and (iv) of Proposition 4.2.2.

(ii) If $\mathcal{B}$ is convex, then so is $\rho_{\mathcal{B}}$.

(iii) If $\mathcal{B}$ is a convex cone, then $\rho_{\mathcal{B}}$ is coherent.

(iv) If $\mathcal{B}$ satisfies condition (v) of Proposition 4.2.2, then $\rho_{\mathcal{B}}$ is real-valued.

Proof. (i): First we will show that $\rho_{\mathcal{B}}$ is a monetary risk measure. Consider

$$\rho_{\mathcal{B}}(0) = \inf\{m \in \mathbb{R} : m \geq z \text{ for some } z \in \mathcal{B}\}.$$

Since $\mathcal{B}$ satisfies property (i) of Proposition 4.2.2, we have that $\rho_{\mathcal{B}}(0) \in \mathbb{R}$. Next let $x \leq y$, then we have

$$\{m \in \mathbb{R} : x + m \geq z \text{ for some } z \in \mathcal{B}\} \subseteq \{m \in \mathbb{R} : y + m \geq z \text{ for some } z \in \mathcal{B}\}.$$

Thus, when looking at the infimum of each set, we have that $\rho_{\mathcal{B}}(x) \geq \rho_{\mathcal{B}}(y)$. It is easy to show that $\rho_{\mathcal{B}}$ has the translation invariance property. We have proved that $\rho_{\mathcal{B}}$ is a monetary risk measure.

Consider $x \in \mathcal{B}$. Then $x + m \geq x$ for all $m \geq 0$. Hence

$$\mathbb{R}_+ \subseteq \{m \in \mathbb{R} : x + m \geq z \text{ for some } z \in \mathcal{B}\}$$

and $\rho_{\mathcal{B}}(x) \leq \inf \mathbb{R}_+ = 0$, i.e. $x \in \mathcal{A}_{\rho_{\mathcal{B}}}$. Thus $\mathcal{B}$ is contained in $\mathcal{A}_{\rho_{\mathcal{B}}}$.

Now, consider a subset C of $\mathcal{X}$ with $\mathcal{B} \subseteq C$ and satisfying (iii) and (iv) of Proposition 4.2.2. We need to show that $\mathcal{A}_{\rho_{\mathcal{B}}} \subseteq C$. Thus, consider $x \in \mathcal{A}_{\rho_{\mathcal{B}}}$. Then

$$\inf\{m \in \mathbb{R} : x + m \geq z \text{ for some } z \in \mathcal{B}\} \leq 0.$$

Hence, there exists $z_n \in \mathcal{B} \subseteq C$ such that $x + \frac{1}{n} \geq z_n$ for $n \geq 1$. By (iii) of Proposition 4.2.2, we have $x + \frac{1}{n} \in C$ and since $x + \frac{1}{n} \to x$ as $n \to \infty$, we have by (iv) of Proposition 4.2.2 that $x \in C$, proving our claim.

(ii): Assume that $\mathcal{B}$ is convex. First, let $\lambda \in (0,1)$ and consider

$$\begin{aligned}\lambda\rho_{\mathcal{B}}(x) &= \inf\{\lambda m \in \mathbb{R} : x + m \geq z \text{ for some } z \in \mathcal{B}\} \\ &= \inf\{m_1 \in \mathbb{R} : \lambda x + m_1 \geq \lambda z \text{ for some } z \in \mathcal{B}\}.\end{aligned}$$

Let

$$U_1 = \{m_1 \in \mathbb{R} : \lambda x + m_1 \geq \lambda z \text{ for some } z \in \mathcal{B}\}$$

and

$$U_2 = \{m_2 \in \mathbb{R} : (1-\lambda)y + m_1 \geq (1-\lambda)z \text{ for some } z \in \mathcal{B}\}.$$

Take $u \in U_1 + U_2$. Then, $u = u_1 + u_2$, where for some $z_1, z_2 \in \mathcal{B}$, $\lambda x + u_1 \geq \lambda z_1$ and $(1-\lambda)y + u_2 \geq (1-\lambda)z_2$. Hence,

$$\lambda x + (1-\lambda)y + u \geq z,$$

where $z = \lambda z_1 + (1-\lambda)z_2 \in \mathcal{B}$, as $\mathcal{B}$ is convex. Thus

$$U_1 + U_2 \subseteq \{m \in \mathbb{R} : \lambda x + (1-\lambda)y + m \geq \lambda z \text{ for some } z \in \mathcal{B}\},$$

proving that $\rho_{\mathcal{B}}$ is convex.

(iii): Assume that $\mathcal{B}$ is a convex cone. By the convexity of $\mathcal{B}$ we get that $\rho_{\mathcal{B}}$ is convex. All we need to show is that $\rho_{\mathcal{B}}$ is positive homogeneous. Let $\lambda > 0$ and consider

$$\begin{aligned}
\rho_{\mathcal{B}}(\lambda x) &= \inf\{m \in \mathbb{R} : \lambda x + m \geq z \text{ for some } z \in \mathcal{B}\} \\
&= \inf\{m \in \mathbb{R} : x + \frac{m}{\lambda} \geq \frac{z}{\lambda} \text{ for some } z \in \mathcal{B}\} \\
&= \inf\{\lambda n \in \mathbb{R} : x + n \geq z_1 \text{ for some } z_1 \in \mathcal{B}\} \\
&= \lambda \inf\{n \in \mathbb{R} : x + n \geq z_1 \text{ for some } z_1 \in \mathcal{B}\} \\
&= \lambda \rho_{\mathcal{B}}(x).
\end{aligned}$$

Note that if $z \in \mathcal{B}$, we have $\frac{z}{\lambda} \in \mathcal{B}$, as $\mathcal{B}$ is a cone. Hence, we have shown that $\rho_{\mathcal{B}}$ is coherent.

(iv): If $\mathcal{B}$ satisfies (v) of Proposition 4.2.2, then $\rho_{\mathcal{B}}$ is by definition real-valued. □

The following two examples are taken from [61].

Example 4.2.4. Consider the *worst case measure* ρ_{max} defined by

$$\rho_{max}(x) = -\inf_{\omega \in \Omega} x(\omega) \quad \text{for all } x \in \mathcal{X}.$$

The corresponding acceptance set $\mathcal{A}$ is given by the convex cone of all nonnegative functions in $\mathcal{X}$. Thus, ρ_{max} is a coherent measure of risk. It is the most conservative measure of risk in the sense that any monetary risk measure ρ on $\mathcal{X}$ with $\rho(0) = 0$ satisfies

$$\rho(x) \leq -\inf_{\omega \in \Omega} x(\omega) = \rho_{max}(x).$$

Example 4.2.5. Let $\mathcal{Q}$ be a set of probability measures on $(\Omega, \mathcal{F})$, and consider a mapping $\gamma : \mathcal{Q} \to \mathbb{R}$ with $\sup_{Q \in \mathcal{Q}} \gamma(Q) < \infty$. Suppose that a position is acceptable if

$$\mathbb{E}_Q[x] \geq \gamma(Q) \quad \text{for all } Q \in \mathcal{Q}.$$

In other words, $\gamma(Q)$ is a lower bound for the expected value. The acceptance set $\mathcal{A}$ of such positions satisfies properties (ii) and (iii) of Proposition 4.2.2 and it is convex. Hence, the associated risk measure $\rho_{\mathcal{A}}$ is convex and is of the form

$$\rho_{\mathcal{A}}(x) = \sup_{Q \in \mathcal{Q}} \{\gamma(Q) - \mathbb{E}_Q[x]\}.$$

Alternatively, we can write

$$\rho_{\mathcal{A}}(x) = \sup_{Q \in \mathbb{P}} \{\mathbb{E}_Q[-x] - \alpha(Q)\},$$

where the *penalty function* $\alpha : \mathbb{P} \to (-\infty, \infty]$ is defined by $\alpha(Q) = -\gamma(Q)$ if $Q \in \mathcal{Q}$ and $\alpha(Q) = \infty$ otherwise. Note that $\rho_{\mathcal{A}}$ is coherent if $\gamma(Q) = 0$ for all $Q \in \mathcal{Q}$.

As mentioned before, representation theorems are important in the study of risk measures. In 2002, Föllmer and Schied [60] proved the following representation theorem for convex risk measures. They assume that $\rho : L_\infty(P) \to \mathbb{R}$.

Theorem 4.2.6. *[60, Theorem 6] Suppose $\mathcal{X} = L_\infty(\Omega, \mathcal{F}, P)$, $\mathbb{P}$ is the set of probability measures, which are absolutely continuous with respect to P, and $\rho : \mathcal{X} \to \mathbb{R}$ is a convex risk measure. Then the following properties are equivalent:*

(i) There is a function $\alpha : \mathbb{P} \to (-\infty, \infty]$ *such that*

$$\rho(x) = \sup_{Q \in \mathbb{P}} \{\mathbb{E}_Q[\,-x\,] - \alpha(Q)\} \text{ for all } x \in \mathcal{X}. \tag{4.4}$$

(ii) The acceptance set $\mathcal{A}_\rho$ associated with ρ is $\sigma(L_\infty(P), L_1(P))$-closed.

(iii) The risk measure ρ possesses the Fatou property, i.e. if the sequence $(x_n)_{n \in \mathbb{N}} \subseteq \mathcal{X}$ is uniformly bounded and x_n converges in probability to some $x \in \mathcal{X}$, then $\rho(x) \leq \liminf_n \rho(x_n)$.

(iv) If the sequence $x_n \subseteq \mathcal{X}$ decreases to $x \in \mathcal{X}$, then $\rho(x_n) \to \rho(x)$.

This considers bounded random variables. However, most models in financial mathematics allow for random variables, which are not bounded, and therefore, it makes sense to look at risk measures on larger sets than $L_\infty(P)$. This has been done by numerous people, amongst them Delbaen [42] and Cheridito et al. [25], who investigated convex risk measures on $L_0(P)$, Frittelli and Rosazza Gianin [62], who provide robust representations for real-valued risk measures on L_p-spaces, and various others (see [26] for more details). Cheridito and Li [26] extended the theory on risk measures to Orlicz spaces, which will be discussed in the next section.

4.3 Risk measures on Orlicz hearts

4.3.1 Monotone functionals on Banach lattices

Before we can state and prove the theorem which gives us the robust representation of a convex risk measure on an Orlicz heart, we need a few definitions and properties of monotone functionals on Banach lattices.

Definition 4.3.1. Consider two normed spaces $\mathcal{X}$ and $\mathcal{Y}$. A function $f : \mathcal{X} \to \mathcal{Y}$ is *Lipschitz-continuous* if there exists a positive constant $K \in \mathbb{R}$ such that

$$||f(x_1) - f(x_2)|| \leq K||x_1 - x_2||,$$

for all $x_1, x_2 \in \mathcal{X}$.

It can be shown that any monetary risk measure ρ is Lipschitz continuous with respect to the supremum norm $||\cdot||_\infty$, i.e.

$$|\rho(x) - \rho(y)| \leq ||x - y||_\infty.$$

Lemma 4.3.2. *If f is an increasing function from a Banach lattice $\mathcal{X}$ to $(-\infty, \infty]$, then* $\operatorname{core}(\operatorname{dom}(f)) = \operatorname{int}(\operatorname{dom}(f))$.

Proof. Since $\operatorname{int}(A) \subseteq \operatorname{core}(A)$ for every $A \subseteq \mathcal{X}$, we just have to show that $\operatorname{core}(\operatorname{dom}(f)) \subseteq \operatorname{int}(\operatorname{dom}(f))$. By way of contradiction, assume that f is real-valued on an algebraic neighbourhood of $x \in \mathcal{X}$ but not on a neighbourhood of x. Then there exist elements $z_n \in \mathcal{X}$, $n \geq 1$ with norm $||x - z_n|| \leq 4^{-n}$ and $f(z_n) = \infty$. We also have that $||x - z_n^+|| \leq 4^{-n}$ and $f(z_n^+) = \infty$. Let $y_n^+ = x - z_n^+$. Then $||y_n^+|| \leq 4^{-n}$ and $f(x + y_n^+) = \infty$.

Define $y = \sum_{n\geq 1} 2^n y_n^+$. By assumption, there exists an $\epsilon > 0$ such that $f(x + ty) \in \mathbb{R}$ for $t \in [0, \epsilon]$. It follows that for all n with $\epsilon 2^n \geq 1$, we have

$$\infty > f(x + \epsilon y) \geq f(x + \epsilon 2^n y_n^+) \geq f(x + y_n^+) = \infty,$$

a contradiction. So there has to exist a neighbourhood of x on which f is real-valued. □

For every proper convex function f, the conjugate f^*, as defined previously, is given by

$$f^*(x^*) = \sup_{x \in \mathcal{X}} \{x^*(x) - f(x)\}.$$

The conjugate is a $\sigma(\mathcal{X}^*, \mathcal{X})$-lower semi-continuous, convex function from $\mathcal{X}^*$ to $(-\infty, \infty]$. It is immediate from the definition of f^* that

$$f(x) \geq f^{**}(x) = \sup_{x^* \in \mathcal{X}^*} \{x^*(x) - f^*(x^*)\}.$$

The next proposition, taken from [169, Corollary 2.2.12], is required for the proof of Theorem 4.3.4.

Proposition 4.3.3. *Let $\mathcal{X}$ be a normed space and $f : \mathcal{X} \to [-\infty, \infty]$ a proper convex function. Suppose that $x_0 \in \mathrm{dom} f$ and for some $p > 0$ and $m \geq 0$*

$$f(x) \leq f(x_0) + m$$

for all $x \in \mathcal{X}$ such that $|x - x_0| \leq p$. Then for all $p' \in (0, p)$ and $x, y \in \mathcal{X}$ such that $|x - x_0| \leq p'$ and $|y - x_0| \leq p'$, we have

$$|f(x) - f(y)| \leq \frac{m}{p} \cdot \frac{p + p'}{p - p'} ||x - y||.$$

Now we can state and prove the following, as done by Cheridito and Li in [26].

Theorem 4.3.4. *Let f be an increasing, convex function from a Banach lattice $\mathcal{X}$ to $(-\infty, \infty]$. For all $x \in \mathrm{core}(\mathrm{dom}(f))$,*

(i) there exists a neighbourhood of x on which f is Lipschitz-continuous with respect to the norm on $\mathcal{X}$,

(ii) f is subdifferentiable at x, and

(iii) $f(x) = \max_{x^* \in \mathcal{X}^*} \{x^*(x) - f^*(x^*)\}$.

Proof. (ii): By Lemma 4.3.2, every $x \in \mathrm{core}(\mathrm{dom}(f))$ has a neighbourhood contained in $\mathrm{dom} f$. So it follows from Proposition C.14 that f is continuous and subdifferentiable at x.

(iii): Note that if f is increasing then f^* is finite at most over $(X^*)^+$. Since f is subdifferentiable, there exists a positive subgradient $x^* \in \mathcal{X}^*$ of f such that for all $y \in \mathrm{int}(\mathrm{dom} f)$ we have

$$f(x) - f(y) \geq x^*(x - y),$$

i.e.

$$x^*(y) - f(y) \geq x^*(x) - f(x).$$

As this holds for all $x \in \mathcal{X}$, we have

$$x^*(y) - f(y) = \max_{x \in \mathcal{X}} \{x^*(x) - f(x)\} = f^*(x^*).$$

This chain of equalities implies that

$$f(y) = x^*(y) - f^*(x^*) = \max_{x^* \in \mathcal{X}^*} \{x^*(y) - f^*(x^*)\},$$

where the last equality is satisfied because $f(y) \geq x^*(y) - f^*(x^*)$ automatically holds for any $x^* \in \mathcal{X}^*$.

(i): Since f is continuous at x, there exists a neighbourhood of x on which f is bounded and hence, (i) follows from Proposition 4.3.3. □

Remark 4.3.5. Theorem 4.3.4 is a special case of the extended Namioka-Klee theorem proved by Biagini and Frittelli [16]. Biagini and Frittelli prove this theorem for a proper, convex and monotone increasing function on a Frechet lattice, which is a generalisation of a Banach lattice. By using a Frechet lattice, Biagini and Frittelli obtain a representation result that holds for Orlicz spaces as well as Orlicz hearts.

4.3.2 Dual Representation

We will now look at risk measures on Orlicz hearts. Let (Φ, Ψ) be complementary finite Young functions. The Orlicz space, heart and norm and the Luxemburg norm are defined as in Chapter 2. We will identify a probability measure $Q \in \mathbb{P}$ on $(\Omega, \mathcal{F})$ with its Radon-Nikodým derivative $\xi = \frac{dQ}{dP} \in L_1(P)$. Define

$$\mathcal{R}(P) = \{\xi \in L_1(P) : \xi \geq 0, \, \mathbb{E}_P[\, \xi \,] = 1\},$$

and let $\mathcal{R}^\Psi(P) = \mathcal{R}(P) \cap L_\Psi(P)$.

Remark 4.3.6. Let $\mathcal{X}$ be an ordered locally convex topological vector space, endowed with a topology τ, for which $\mathcal{X}$ and its Banach dual space $\mathcal{X}^*$ form a dual system. The space of positive continuous linear functionals, i.e. the positive polar cone of $\mathcal{X}^+$ is given by

$$(\mathcal{X}^+)^\circ = \{x^* \in \mathcal{X}^* : x^*(x) \geq 0 \text{ for all } x \in \mathcal{X}^+\}.$$

If $z^* \in (\mathcal{X}^+)^\circ$ and $z^*(\mathbf{1}) = 1$, then z^* is a probability density in $(\mathcal{X}^+)^\circ$. By the Radon-Nikodým theorem, the probability density z^* can be identified with its associated probability measure Q by setting $z^* = \frac{dQ}{dP}$. Therefore, we have

$$z^*(x) = \mathbb{E}_P[\, z^* x \,] = \mathbb{E}_Q[x].$$

The concept of a penalty function is crucial to the theory of risk measures as it defines the robust representations.

Definition 4.3.7.

(i) A map $\alpha : \mathcal{R}^\Psi(P) \to (-\infty, \infty]$ is called a *penalty function* on $\mathcal{R}^\Psi(P)$ if it is bounded from below and not identically equal to infinity.

(ii) The penalty function $\alpha : \mathbb{P} \to (-\infty, \infty]$ satisfies the *growth condition* with respect to the norm $||\cdot||$ if there exist $a \in \mathbb{R}$ and $b > 0$ such that

$$\alpha(Q) \geq a + b\,\|Q\| \text{ for all } Q \in \mathbb{P}. \tag{4.5}$$

(iii) For any penalty function α on $\mathcal{R}^{\Psi}(P)$, we define for all $x \in \mathcal{X}$

$$\rho_{\alpha}(x) = \sup_{Q \in \mathcal{R}^{\Psi}(P)} \{\mathbb{E}_Q[\,-x\,] - \alpha(Q)\}, \tag{4.6}$$

and call the right hand side of (4.6) a *robust representation* of ρ_{α}.

Since the Orlicz norm and the Luxemburg norm are equivalent, a penalty function α satisfies the growth condition with respect to the Orlicz norm if and only if α satisfies the growth condition with respect to the Luxemburg norm.

Remark 4.3.8. We repeat here as a reminder some properties of the Orlicz and Luxemburg norms. Note that by the definitions of the L_1-norm and the L_∞-norm, we have for $Q \in L_{\Psi}(P)$ that $\mathbb{E}_Q[y] \leq ||y||_\infty$ and $\mathbb{E}_Q[-y] \leq ||-y||_\infty = ||y||_\infty$.

Also, by the definition of the Orlicz norm

$$\begin{aligned}
||Q||_{\Psi} &\geq \int |Q \cdot \frac{y-x}{\mathcal{N}_{\Phi}(y-x)}|dP \\
\mathcal{N}_{\Phi}(y-x)||Q||_{\Psi} &\geq \int |Q(y-x)|dP \\
&= \mathbb{E}_Q[|y-x|] \\
&= \mathbb{E}_Q[|x-y|] \\
&\geq \big|\mathbb{E}_Q[x-y]\big| \\
&\geq \mathbb{E}_Q[x-y],
\end{aligned}$$

where the second last inequality follows from Jensen's inequality.

Proposition 4.3.9. *The function ρ_{α} given by (4.6) defines a lower semi-continuous convex monetary risk measure on $H_{\Phi}(P)$ with values in $(-\infty, \infty]$.*

Proof. First, consider

$$\rho_{\alpha}(0) = \sup_{Q \in \mathcal{R}^{\Psi}(P)} \{-\alpha(Q)\} = -\inf_{Q \in \mathcal{R}^{\Psi}(P)} \alpha(Q) < \infty,$$

as α is bounded below.

Secondly, let $x \leq y$. Then $\mathbb{E}_Q[-x] - \alpha(Q) \geq \mathbb{E}_Q[-y] - \alpha(Q)$ for all $Q \in \mathcal{R}^\Psi(P)$, proving the monotonicity. Translation invariance is clear. To show that ρ_α is convex, let $\lambda \in (0,1)$. Then,

$$\begin{aligned}
&\lambda\rho_\alpha(x) + (1-\lambda)\rho_\alpha(y) \\
&\qquad = \sup_{Q\in\mathcal{R}^\Psi(P)} \{\mathbb{E}_Q[-\lambda x] - \lambda\alpha(Q)\} + \sup_{Q\in\mathcal{R}^\Psi(P)} \{\mathbb{E}_Q[-(1-\lambda)y] - (1-\lambda)\alpha(Q)\} \\
&\qquad \geq \sup_{Q\in\mathcal{R}^\Psi(P)} \{\mathbb{E}_Q[-\lambda x - (1-\lambda)y] - \alpha(Q)\} \\
&\qquad = \rho_\alpha(\lambda x + (1-\lambda)y).
\end{aligned}$$

To show that the risk measure ρ_α is lower semi-continuous at x, we have to show that for every $\epsilon > 0$, there exists $\delta > 0$ such that for all $y \in N_\delta(x)$, where $N_\delta(x)$ is the δ-neighbourhood of x, we have $\rho_\alpha(x) - \rho_\alpha(y) < \epsilon$. Consider

$$\begin{aligned}
\mathbb{E}[-x] &\leq \mathbb{E}[|x-y|] + \mathbb{E}[-y] \\
&= ||x-y||_1 + \mathbb{E}[-y] \\
&\leq K\mathcal{N}_\Phi(x-y) + \mathbb{E}[-y]
\end{aligned}$$

for some K, as we have the continuous embedding $L_\infty \hookrightarrow L_\Phi \hookrightarrow L_1$. Hence,

$$\begin{aligned}
\mathbb{E}[-x] - \alpha(Q) &\leq K\mathcal{N}_\Phi(x-y) + \mathbb{E}[-y] - \alpha(Q) \\
&\leq K\mathcal{N}_\Phi(x-y) + \rho_\alpha(y).
\end{aligned}$$

Thus,

$$\rho_\alpha(x) \leq K\mathcal{N}_\Phi(x-y) + \rho_\alpha(y),$$

i.e. $\rho_\alpha(x) - \rho_\alpha(y) \leq K\mathcal{N}_\Phi(x-y) < \epsilon$ if and only if $\mathcal{N}_\Phi(x-y) < \delta := \frac{\epsilon}{K}$. □

The next theorem characterises penalty functions, which satisfy the growth condition.

Theorem 4.3.10. *Let α be a penalty function on $\mathcal{R}^\Psi(P)$. The following are equivalent.*

(i) *α satisfies the growth condition with respect to $||\cdot||_\Psi$.*

(ii) $\mathrm{core}(\mathrm{dom}(\rho_\alpha)) \neq \emptyset$.

(iii) *ρ_α is real-valued and every $x \in H_\Phi(P)$ has a neighbourhood on which ρ_α is Lipschitz-continuous with respect to $\mathcal{N}_\Phi(\cdot)$.*

Proof. We show (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) $\Rightarrow$ (iii). The first implication is trivial.

(ii) $\Rightarrow$ (i): Assume that ρ_α is real-valued on an algebraic neighbourhood of $x \in H_\Phi(P)$. Since the mapping $y \to \rho_\alpha(-y)$ is increasing, we obtain from Lemma 4.3.2 that there exists an $\epsilon > 0$ such that ρ_α is real-valued on the closed ball $B_\epsilon(x)$ with radius ϵ around x. Since $L_\infty(P)$ is $\mathcal{N}_\Phi$-dense in $H_\Phi(P)$, there exists a sequence $(y_n)_{n\geq 1}$ of bounded random variables such that $\mathcal{N}_\Phi(y_n - x) \leq \epsilon 2^{-n-2}$. Assume that α does not satisfy the growth condition with respect to $||\cdot||_\Psi$. Then there exists a sequence of probability measures $(Q_n)_{n\geq 1}$ in $\mathcal{R}^\Psi(P)$ such that for all $n \geq 1$

$$\alpha(Q_n) < -n - ||y_n||_\infty + \epsilon 2^{-n-2}||Q_n||_\Psi.$$

Since $(L_\Psi(P), ||\cdot||_\Psi)$ is the norm dual of $(H_\Phi(P), \mathcal{N}_\Phi(\cdot))$ and by the definition of the Orlicz norm, there exists for every $n \geq 1$, $z_n \in H_\Phi(P)$ such that $z_n \leq 0$, $\mathcal{N}_\Phi(z_n) \leq 1$ and

$$\begin{aligned}\mathbb{E}_{Q_n}[-z_n] &= \int -z_n Q_n dP \\ &= \int |-z_n Q_n| dP \\ &\geq ||Q_n||_\Psi \\ &\geq \frac{1}{2}||Q_n||_\Psi.\end{aligned}$$

The random variable $z = \epsilon \sum_{n\geq 1} 2^{-n} z_n$ is in $H_\Phi(P)$ with norm $\mathcal{N}_\Phi(z) \leq \epsilon$. As $z_n \leq 0$ and ρ_α is monotonic and using Remark 4.3.8, we get

$$\begin{aligned}&\rho_\alpha(x+z) \\ &\quad\geq \rho_\alpha(x + \epsilon 2^{-n} z_n) \\ &\quad\geq \mathbb{E}_{Q_n}[-x - \epsilon 2^{-n} z_n] - \alpha(Q_n) \\ &\quad\geq \mathbb{E}_{Q_n}[-y_n] + \mathbb{E}_{Q_n}[y_n - x] + \epsilon 2^{-n}\mathbb{E}_{Q_n}[-z_n] + n + ||y_n||_\infty - \epsilon 2^{-n-2}||Q_n||_\Psi \\ &\quad\geq -||y_n||_\infty - \mathcal{N}_\Phi(y_n - x)||Q_n||_\Psi + \epsilon 2^{-n-1}||Q_n||_\Psi + n + ||y_n||_\infty - \epsilon 2^{-n-2}||Q_n||_\Psi \\ &\quad= ||Q_n||_\Psi\Big[-\mathcal{N}_\Phi(y_n - x) + \epsilon 2^{-n-1} - \epsilon 2^{-n-2}\Big] + n \\ &\quad= ||Q_n||_\Psi\Big[-\mathcal{N}_\Phi(y_n - x) + \epsilon 2^{-n-2}\Big] + n \\ &\quad\geq 0 + n \\ &\quad= n\end{aligned}$$

for all $n \geq 1$. But this contradicts the finiteness of ρ_α on $B_\epsilon(x)$. Therefore, α must satisfy the growth condition with respect to $||\cdot||_\Psi$, and (i) is proved.

(i) $\Rightarrow$ (iii): Assume there exist constants $a \in \mathbb{R}$ and $b > 0$ such that $\alpha(Q) \geq a + b||Q||_\Psi$ for all $Q \in \mathcal{R}^\Psi(P)$. Choose $x \in H_\Phi(P)$. There exists $\bar{x} \in L_\infty(P)$ with $\mathcal{N}_\Phi(x - \bar{x}) \leq b$, and we obtain

$$\begin{aligned}\mathbb{E}_Q[-x] - \alpha(Q) &= \mathbb{E}_Q[-\bar{x}] + \mathbb{E}_Q[\bar{x} - x] - \alpha(Q)\\ &\leq ||\bar{x}||_\infty + \mathcal{N}_\Phi(\bar{x} - x)||Q||_\Psi - a - b||Q||_\Psi\\ &\leq ||\bar{x}||_\infty - a\end{aligned}$$

for all $Q \in \mathcal{R}^\Psi(P)$. This shows that $\rho_\alpha(x) \leq ||\bar{x}||_\infty - a < \infty$. Hence, ρ_α is real-valued. The rest of (iii) follows from Theorem 4.3.4 (i) with $f(x) = \rho_\alpha(-x)$. □

Cheridito and Li [26] show that every convex monetary risk measure ρ on $H_\Phi(P)$ with $\operatorname{core}(\operatorname{dom}(\rho)) \neq \emptyset$ has a robust representation with penalty function

$$\rho^\#(Q) = \sup_{x \in \mathcal{A}_\rho} \mathbb{E}_Q[-x], \tag{4.7}$$

for $Q \in \mathcal{R}^\Psi(P)$ and that $\rho^\#$ is the minimal penalty function of ρ. Since $x + \rho(x) \in \mathcal{A}_\rho$ for all $x \in H_\Phi(P)$, Equation (4.7) can equivalently be written as

$$\rho^\#(Q) = \sup_{x \in H_\Phi(P)} \{\mathbb{E}_Q[-x] - \rho(x)\} \tag{4.8}$$

for $Q \in \mathcal{R}^\Psi(P)$.

Now we are ready to state and prove the representation theorem for risk measures on Orlicz hearts.

Theorem 4.3.11. *Let $\rho : H_\Phi(P) \to (-\infty, \infty]$ be a convex monetary risk measure with* $\operatorname{core}(\operatorname{dom}(\rho)) \neq \emptyset$. *Then*

(i) ρ is real-valued,

(ii) $\rho^\#$, defined by

$$\rho^\#(Q) = \sup_{x \in H_\Phi(P)} \{\mathbb{E}_Q[-x] - \rho(x)\},$$

is a penalty function on $\mathcal{R}^\Psi(P)$ satisfying the growth condition with respect to $||\cdot||_\Psi$, and

(iii) for all $x \in H_\Phi(P)$

$$\rho(x) = \max_{Q \in \mathcal{R}^\Psi(P)} \{\mathbb{E}_Q[\,-x\,] - \rho^\#(Q)\}. \tag{4.9}$$

Moreover, if $\rho = \rho_\alpha$ for a penalty function α on $\mathcal{R}^\Psi(P)$, then $\rho^\#$ is the greatest convex, $\sigma(L_\Psi(P), H_\Phi(P))$-lower semi-continuous minorant of α.

Proof. Consider the function $f(x) = \rho(-x)$ for $x \in H_\Phi(P)$. Then f is increasing and convex. Note that

$$\begin{aligned} f^*(Q) &= \sup_{x \in H_\Phi(P)} \{\mathbb{E}_Q[x] - f(x)\} \\ &= \sup_{x \in H_\Phi(P)} \{\mathbb{E}_Q[x] - \rho(-x)\} \\ &= \sup_{y \in H_\Phi(P)} \{\mathbb{E}_Q[-y] - \rho(y)\} \\ &= \sup_{y \in \mathcal{A}_\rho} \mathbb{E}_Q[-y] \\ &= \begin{cases} \rho^\#(Q) & \text{if } Q \in \mathcal{R}^\Psi(P) \\ \infty & \text{if } Q \in L_\Psi(P) \backslash \mathcal{R}^\Psi(P). \end{cases} \end{aligned}$$

To get the last equality, we adapt the proof of Lemma 4.30 of Föllmer and Schied [60]. Consider $Q \in L_\Psi(P) \backslash \mathcal{R}^\Psi(P)$, i.e. Q is not absolutely continuous with respect to P. Then there exists $A \in \mathcal{F}$ such that $Q(A) > 0$ but $P(A) = 0$. Take any $x \in \mathcal{A}_\rho$ and define $x_n = x - n\mathbb{1}_A$ for $n \in \mathbb{N}$. Then $\rho(x_n) = \rho(x)$, i.e. $x_n \in \mathcal{A}_\rho$. Hence,

$$\begin{aligned} f^*(Q) &\geq \mathbb{E}_Q[-x_n] - \rho(x_n) \\ &\geq \mathbb{E}_Q[-x] + nQ(A) \to \infty. \end{aligned}$$

Thus, it follows from Theorem 4.3.4 (iii), that for all $X \in \text{core}(\text{dom}(f))$,

$$\begin{aligned} \rho(x) = f(-x) &= \max_{Q \in L_\Psi(P)} \{\mathbb{E}_Q[-x] - f^*(Q)\} \qquad (4.10) \\ &\geq \max_{Q \in \mathcal{R}^\Psi(P)} \{\mathbb{E}_Q[-x] - f^*(Q)\} \\ &= \max_{Q \in \mathcal{R}^\Psi(P)} \{\mathbb{E}_Q[-x] - \rho^\#(Q)\}. \qquad (4.11) \end{aligned}$$

This shows that ρ has a continuous affine minorant. Therefore, the greatest lower semi-continuous minorant $\overline{\text{co}}\rho$ of ρ (defined in Appendix C), also called the lower semi-continuous hull of ρ, is proper. Since ρ is convex, we have that the greatest lower semi-continuous minorant $\overline{\text{co}}\rho$ is equal to the lower semi-continuous envelope $\bar{\rho}$ of ρ. Hence, we obtain from Theorem C.18 that $\bar{\rho} = \rho^{**}$, i.e.

$$\bar{\rho}(x) = \sup_{Q \in \mathcal{R}^\Psi(P)} \{\mathbb{E}_Q[-x] - \rho^\#(Q)\}$$

for all $x \in H_\Phi(P)$.

Since $\bar{\rho}$ is proper for all $x \in \mathcal{X}$ and $\rho(0) \in \mathbb{R}$, we get $-\infty < \bar{\rho}(0) < \rho(0) < \infty$. Also,

$$\bar{\rho}(0) = - \inf_{Q \in \mathcal{R}^\Psi(P)} \rho^{\#}(Q).$$

This implies that $\rho^{\#}$ is bounded below and not identically equal to ∞, and hence $\rho^{\#}$ is a penalty function.

Furthermore, $\mathrm{core}(\mathrm{dom}(\bar{\rho})) \supseteq \mathrm{core}(\mathrm{dom}(\rho)) \neq \emptyset$. So, Theorem 4.3.10 yields that $\bar{\rho}$ is real-valued and $\rho^{\#}$ satisfies the growth condition.

(Note that a set is nowhere dense if and only if its closure has non-empty interior. Also, a convex set has the same interior points as its closure. Thus, if the interior of a convex set is non-empty, then the set is dense.) By Lemma 4.3.2, the interior of the convex set $\mathrm{dom}\rho$ is non-empty, thus we have that $\mathrm{dom}\rho$ is dense in $H_\Phi(P)$.

Now assume that there exists $y \in H_\Phi(P) \backslash \mathrm{dom}\rho$, i.e. $\rho(y) = \infty$. Then, by Eidelheit's separation theorem, Theorem C.3, there exists $\xi \in L_\Psi(P) \backslash \{0\}$ such that

$$\sup_{x \in \mathrm{dom}\rho} \mathbb{E}_P[x\xi] \leq \mathbb{E}_P[y\xi].$$

Then select $z \in H_\Phi(P)$ such that

$$\sup_{x \in \mathrm{dom}\rho} \mathbb{E}_P[x\xi] \leq \mathbb{E}_P[y\xi] < \mathbb{E}_P[z\xi].$$

However, since $\mathrm{dom}\rho$ is dense in $H_\Phi(P)$ and $z \in H_\Phi(P)$, we have that there exists $z_0 \in \mathrm{dom}\rho$ such that for all $\epsilon > 0$

$$\begin{aligned} \mathbb{E}_P[z\xi] &= \mathbb{E}_P[z\xi - z_0\xi] + \mathbb{E}[z_0\xi] \\ &\leq K\epsilon + \sup_{x \in \mathrm{dom}\rho} \mathbb{E}_P[x\xi] \\ &\leq \sup_{x \in \mathrm{dom}\rho} \mathbb{E}_P[x\xi] \\ &< \mathbb{E}_P[z\xi]. \end{aligned}$$

This is a contradiction of the density of $\mathrm{dom}\rho$ in $H_\Phi(P)$ and the continuity of $\mathbb{E}_P$. Hence, ρ must be real-valued. The equality in representation (4.9) now follows from (4.11) and the fact that ρ is real-valued.

To prove the last part of the theorem, let α be a penalty function on $\mathcal{R}^\Psi(P)$ with $\rho = \rho_\alpha$. Denote by $\hat{\alpha}$ the function from $L_\Psi(P)$ to $(-\infty, \infty]$ which is equal to α on $\mathcal{R}^\Psi(P)$ and ∞ on $L_\Psi(P) \backslash \mathcal{R}^\Psi(P)$. Then f^* is the biconjugate of $\hat{\alpha}$ in the duality $(L_\Psi(P), H_\Phi(P))$. Since $\hat{\alpha}$ is bounded from below, it follows from Theorem C.18 that f^* is the greatest convex,

$\sigma(L_\Psi(P), H_\Phi(P))$-lower semi-continuous minorant of $\hat{\alpha}$. Since $\rho^{\#}$ is the restriction of f^* to $\mathcal{R}^\Psi(P)$, this completes the proof. □

For every non-empty subset $\mathcal{Q}$ of $\mathcal{R}^\Psi(P)$,

$$\alpha(Q) = \begin{cases} 0 & \text{for } Q \in \mathcal{Q} \\ \infty & \text{for } Q \notin \mathcal{Q} \end{cases}$$

is a penalty function on $\mathcal{R}^\Psi(P)$. It is easily verified, that the corresponding risk measure

$$\rho_{\mathcal{Q}}(x) = \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q[\,-x\,]$$

is coherent and α satisfies the growth condition if and only if $\mathcal{Q}$ is $||\cdot||_\Psi$-bounded. Hence, Theorem 4.3.10 can be written as follows.

Corollary 4.3.12. *Let $\mathcal{Q}$ be a non-empty subset of $\mathcal{R}^\Psi(P)$. Then the following conditions are equivalent:*

(i) *$\mathcal{Q}$ is $||\cdot||_\Psi$-bounded;*

(ii) $\operatorname{core}(\operatorname{dom}(\rho_{\mathcal{Q}})) \neq \emptyset$*;*

(iii) *$\rho_{\mathcal{Q}}$ is real-valued and Lipschitz-continuous with respect to $\mathcal{N}_\Phi(\cdot)$.*

Proof. We once again show (iii) ⇒ (ii) ⇒ (i) ⇒ (iii). The first implication is trivial.

(ii) ⇒ (i): This follows from the fact that α satisfies the growth condition if and only if $\mathcal{Q}$ is $||\cdot||_\Psi$-bounded. This is easily verified.

(i) ⇒ (iii): For all $x, y \in H_\Phi(P)$ and $Q \in \mathcal{Q}$, we have

$$\rho_{\mathcal{Q}}(x) \leq \mathbb{E}_Q[y - x] + \mathbb{E}_Q[-y] \leq \mathcal{N}_\Phi(x - y)||Q||_\Psi + \rho_{\mathcal{Q}}(y)$$

and analogously

$$\rho_{\mathcal{Q}}(y) \leq \mathcal{N}_\Phi(x - y)||Q||_\Psi + \rho_{\mathcal{Q}}(x).$$

So if (i) holds, then $K = \sup_{Q \in \mathcal{Q}} ||Q||_\Psi$ is finite and

$$|\rho_{\mathcal{Q}}(x) - \rho_{\mathcal{Q}}(y)| \leq K\mathcal{N}_\Phi(x - y).$$

□

Thus, for a coherent risk measure ρ, Theorem 4.3.11 reduces to the following.

Corollary 4.3.13. *Let $\rho : H_\Phi(P) \to (-\infty, \infty]$ be a coherent risk measure with acceptance set $\mathcal{A}_\rho$. If* core(dom(ρ)) $\neq \emptyset$*, then ρ is real-valued and can be represented as*

$$\rho(x) = \max_{Q \in \mathcal{Q}} \mathbb{E}_Q[\,-x\,],$$

where $x \in H_\Phi(P)$ and $\mathcal{Q}$ is a $||\cdot||_\Psi$-bounded, convex set given by

$$\mathcal{Q} = \{Q \in \mathcal{R}^\Psi(P) : \mathbb{E}_Q[x] \geq 0 \text{ for all } x \in \mathcal{A}_\rho\}.$$

Proof. Let ρ be a coherent risk measure. Then ρ is a convex risk measure and by Theorem 4.3.11, we have that

$$\rho(x) = \max_{Q \in \mathcal{R}^\Psi(P)} \{\mathbb{E}_Q[-x] - \rho^{\#}(Q)\}.$$

Next, we claim that if ρ is coherent we get

$$\rho^{\#}(Q) = \begin{cases} 0 & \text{for } Q \in \mathcal{Q} \\ \infty & \text{otherwise,} \end{cases}$$

where $\mathcal{Q} = \{Q \in \mathcal{R}^\Psi(P) : \mathbb{E}_Q[X] \geq 0 \text{ for all } X \in \mathcal{A}_\rho\}$. First, note that

$$\begin{aligned} \rho^{\#}(Q) &= \sup_{x \in H_\Phi(P)} \{\mathbb{E}_Q[-x] - \rho(x)\} \\ &= -\inf_{x \in H_\Phi(P)} \{\mathbb{E}_Q[x] + \rho(x)\}. \end{aligned}$$

Secondly, for $\lambda > 0$,

$$\begin{aligned} \rho^{\#}(Q) &= \sup_{x \in H_\Phi(P)} \{\mathbb{E}_Q[-x] - \rho(x)\} \\ &= \sup_{\lambda x \in H_\Phi(P)} \{\mathbb{E}_Q[-\lambda x] - \rho(\lambda x)\} \\ &= \lambda \rho^{\#}(Q), \end{aligned}$$

by the positive homogeneity of ρ. Hence, we must have that $\rho^{\#}(Q) = 0$ if

$$\mathbb{E}_Q[x] + \rho(x) \geq 0 \tag{4.12}$$

for all $x \in H_\Phi(P)$ and $\rho^{\#}(Q) = \infty$ otherwise. Lastly, note that if $\mathbb{E}_Q[x] + \rho(x) \geq 0$, then $\mathbb{E}_Q[x] \geq 0$ if and only if $x \in \mathcal{A}_\rho$. Conversely, if $\mathbb{E}_Q[x] \geq 0$ for $x \in \mathcal{A}_\rho$, then we always have that $\mathbb{E}_Q[x] + \rho(x) \geq 0$. Thus, (4.12) is equivalent to $\mathbb{E}_Q[x] \geq 0$ for $x \in \mathcal{A}_\rho$, proving our claim. □

4.4 Examples of risk measures

For this section, fix some real-valued random variable x on a probability space $(\Omega, \mathcal{F}, P)$. The random variable x represents the profit or loss of some asset or portfolio.

Definition 4.4.1. Given $\alpha \in (0,1)$.

(i) The *lower α-quantile* of x is given by $q_\alpha(x) = \inf\{x_0 \in \mathbb{R} : P(x \leq x_0) \geq \alpha\}$.

(ii) The *upper α-quantile* of x is given by $q^\alpha(x) = \inf\{x_0 \in \mathbb{R} : P(x \leq x_0) > \alpha\}$.

Value at risk (VaR) is a good example of a popular risk measure that is in general not coherent. For a given portfolio, probability and time horizon, VaR is defined as the threshold value such that the probability, that the loss on the portfolio over the given time horizon exceeds this value, is the given probability level.

Definition 4.4.2. Given a confidence level $\alpha \in (0,1)$, VaR_α *at level α of x* is given by

$$VaR_\alpha(x) = q_{1-\alpha}(-x).$$

Note that VaR is not even weakly coherent as, in general, it fails the subadditivity condition. This implies that VaR might discourage diversification. As VaR is dependent on the distribution used, the subadditivity of VaR also depends on which distribution is chosen. If the underlying distribution is elliptic, like the normal or t-distribution, then VaR is coherent (see [55] for more details and a proof). In spite of VaR's drawbacks, it has met the favour of regulatory agencies and has thus become a vital part of financial regulations.

Another risk measure, introduced by Artzner et al. [6], which is easy to compute but once again is, in general, not coherent is the tail conditional expectation (TCE).

Definition 4.4.3. Assume $\mathbb{E}[x^-] < \infty$.

(i) The *lower tail conditional expectation* at level α of x is given by

$$TCE_\alpha(x) = -\mathbb{E}[\, x | x \leq q_\alpha(x)\,].$$

(ii) The *upper tail conditional expectation* at level α of x is given by

$$TCE^\alpha(x) = -\mathbb{E}[\, x | x \leq q^\alpha(x)\,].$$

In general, TCE does not define a subadditive risk measure.

An example of a coherent risk measure is the worst conditional expectation, introduced by Artzner et al. [6].

Definition 4.4.4. Assume $\mathbb{E}[x^-] < \infty$. The *worst conditional expectation* (WCE) at level α of x is defined by

$$WCE_\alpha(x) = -\inf\{\mathbb{E}[x|A] : P(A) > \alpha\}.$$

WCE is coherent but only useful in a theoretical setting as it requires the knowledge of the whole underlying probability space.

The natural coherent alternative to VaR is expected shortfall (ES), sometimes also called *conditional value at risk* or *tail loss*. It was also introduced by Artzner et al. [6] and it answers the question: if things do get bad, what is the expected loss?

Definition 4.4.5. Assume $\mathbb{E}[x^-] < \infty$. The *expected shortfall* at level α of x is defined as

$$ES_\alpha(x) = -\frac{1}{\alpha}\Big(\mathbb{E}[x\mathbb{1}_{\{x \le q_\alpha(x)\}}] + q_\alpha(x)(\alpha - P(x \le q_\alpha(x)))\Big).$$

If the distribution is continuous then ES can be written as

$$ES_\alpha(x) = \mathbb{E}[\, x|x < VaR_\alpha(x)\,].$$

It is easy to show that ES is coherent, continuous with respect to α and monotone in α.

It is also possible to define a risk measure with respect to a utility function. Consider an investor, whose preferences can be represented by a utility function $u : \mathbb{R} \to \mathbb{R}$, which is increasing and concave and satisfies $\lim_{x\to-\infty} u(x) = -\infty$. Without loss of generality, we assume that $u(0) = 0$. Using the Orlicz space associated with the utility function as defined in Section 3.5.2, it is possible to find a risk measure linked to the utility function.

Proposition 4.4.6. *The downside risk* $\Theta : H_{\hat{u}}(P) \to (-\infty, \infty]$ *given by* $\Theta(x) = \mathbb{E}[-u(x)]$, *is a well-defined proper, convex, monotone decreasing and order lower semi-continuous function, which admits the representation*

$$\Theta(x) = \sup_{x^* \in L^+_{\hat{\Phi}}(P)} \{x^*(-x) - \mathbb{E}[\Phi(x^*)]\}.$$

For a proof of this result, see Biagini and Frittelli [16]. The function Θ satisfies all the properties of a convex risk measure, except for translation invariance.

The function $\zeta_u : H_{\hat{u}}(P) \to (-\infty, \infty]$ defined by

$$\zeta_u(x) = \sup_{Q << P, \frac{dQ}{dP} \in L^+_{\hat{\Phi}}(P)} \Big\{\mathbb{E}_Q[-x] - \mathbb{E}\Big[\Phi\Big(\frac{dQ}{dP}\Big)\Big]\Big\}$$

is a well-defined, order lower semi-continuous, convex risk measure. This risk measure is the greatest, order lower semi-continuous, convex risk measure smaller than the downside risk Θ. See [16] for more details.

If u is the quadratic-flat utility function, defined by

$$u(x) = \begin{cases} -\frac{x^2}{2} & \text{for } x \le 0 \\ 0 & \text{for } x \ge 0, \end{cases}$$

then $L_{\hat{u}}(P) = L_2(P)$ and for $x \in L_2(P)$, we have

$$\Theta(x) = \frac{1}{2}\mathbb{E}[(x^-)^2] = \sup_{y \in L_2^+(P)} \Big\{ \mathbb{E}[-yx] - \frac{1}{2}\mathbb{E}[y^2] \Big\}$$

and

$$\zeta_u(x) = \sup_{\frac{dQ}{dP} \in L_2^+(P)} \Big\{ \mathbb{E}_Q[-x] - \frac{1}{2}\mathbb{E}\Big[\Big(\frac{dQ}{dP}\Big)^2\Big] \Big\}.$$

If u is an exponential utility function, defined by $u(x) = -e^{-x} + 1$, then for $x \in H_{\hat{u}}(P)$, we have $\Theta(x) = \mathbb{E}[e^{-x} - 1]$ and

$$\begin{aligned} \zeta_u(x) &= \ln \mathbb{E}[e^{-x}] \\ &= \sup_{Q << P, \frac{dQ}{dP} \in L_{\hat{\Phi}}^+} \Big\{ \mathbb{E}_Q[-x] - \mathbb{E}_Q\Big[\ln\Big(\frac{dQ}{dP}\Big)\Big] \Big\}, \end{aligned}$$

which is known as the *entropic risk measure*.

Another way to define a risk measure with respect to a utility function is as follows. Consider a utility function u, a probability measure $Q \in \mathcal{M}_p$ and fix some threshold $c \in \mathbb{R}$. Define the set

$$\mathcal{A} = \{x \in \mathcal{X} : \mathbb{E}_Q[u(x)] \ge u(c)\}.$$

Clearly, this set satisfies the axioms of an acceptance set and thus, $\rho_{\mathcal{A}}$ is a convex risk measure. One can extend this by defining acceptability in terms of a whole class $\mathcal{Q}$ of probability measures, i.e.

$$\mathcal{A} = \bigcap_{Q \in \mathcal{Q}} \{x \in \mathcal{X} : \mathbb{E}_Q[u(x)] \ge u(c_Q)\},$$

where $\sup_{Q \in \mathcal{Q}} c_Q < \infty$.

This definition of acceptability can be described in terms of a loss function instead of a utility function. Suppose that $l : \mathbb{R} \to \mathbb{R}$ is an increasing convex loss function, which is not identically constant. For a position $x \in L_\infty(P)$, the expected loss is given by $\mathbb{E}_P[\,l(-X)\,]$. Let x_0 be an interior point in the range of l. A position x will be called acceptable if the expected loss is bounded by x_0, i.e. the acceptance set is given by

$$\mathcal{A} = \{x \in L_\infty(P) : \mathbb{E}_P[\,l(-x)\,] \le x_0\}. \tag{4.13}$$

The risk measure $\rho_{\mathcal{A}}$ can be represented in the form

$$\rho_{\mathcal{A}}(x) = \max_{Q \in \mathcal{M}} \{\mathbb{E}_Q[-x] - \alpha_{min}(Q)\}.$$

The minimal penalty function can be calculated for different loss functions. For example, if $l(x) = e^{\beta x}$, then the minimal penalty function can be described in terms of relative entropy. The reader is referred to Föllmer and Schied [61] for more details.

Lastly, consider a set function $c : \mathcal{F} \to [0,1]$ such that $c(\emptyset) = 0$, $c(\Omega) = 1$ and $c(A) \leq c(B)$ if $A \subseteq B$. Define for all $x \in \mathcal{X}$, the integral

$$\int X \; dc = \int_{-\infty}^{0} c(x > x_0) - 1) dx_0 + \int_{0}^{\infty} c(x > x_0) dx_0.$$

This is known as the Choquet integral, which will be defined and described in detail in the next chapter. The Choquet integral of loss

$$\rho(x) = \int (-x) dc$$

is a positively homogeneous monetary risk measure on $\mathcal{X}$. In Chapter 5, we will characterise these risk measures as 'comonotonic risk measures' and prove a representation theorem.

An example of the set function c is the Wang transform, introduced by Wang [159]. This will be discussed in detail in the next chapter.

Chapter 5

Wang transform as a risk measure

In this chapter, we consider a risk measure that has its roots in actuarial science. At the end of Chapter 4, we mentioned the Wang transform [159] as an example of a static risk measure. In this chapter we will develop the theory for this. As the Wang transform is a special case of a Choquet integral, we first define comonotonicity and the Choquet integral and state some representation results for the Choquet integral. We then describe how the Choquet integral can be used to price contingent claims in mathematics of finance. Before we can define the Wang transform and show how it can be used to price financial risk, we define stochastic differential equations and give some of their properties. We also state and prove the necessary condition required on the parameters of the Wang transform to ensure arbitrage-free pricing, as shown by Pelsser [130].

Section 5.4.2 is new and is based on [117]. In this section, we show how the Wang transform can be used to price exotic options. Section 5.5 is based on original work by Labuschagne and Offwood [113] and shows that the Wang transform is in fact a special case of the Esscher-Girsanov transform.

Then, we define comonotonic convex risk measures and state the representation theorem for these risk measures, giving an alternate proof to the original by Song and Yan [148]. Lastly, we show how the Wang transform can be used as a risk measure.

Note that in this chapter, we use the capital X for random variables, as is customary in statistics. In the rest of the thesis, we reserve the capital X for spaces.

5.1 Comonotonicity and the Choquet integral

5.1.1 Comonotonicity

The term ‘comonotonic’ comes from ‘common monotonic’ and the theory surrounding comonotonicity is discussed amongst others by Schmeidler [143] and Yaari [164].

Definition 5.1.1. The set $A \subseteq \mathbb{R}^n$ is comonotonic if for any $\underline{x}$ and $\underline{y}$ in A, either $\underline{x} \leq \underline{y}$ or $\underline{y} \leq \underline{x}$ holds, where $\underline{x} \leq \underline{y}$ means that $x_i \leq y_i$ for all i.

A set $A \subseteq \mathbb{R}^n$ is comonotonic if for any $\underline{x}$ and $\underline{y}$ in A, the inequality $x_i < y_i$ for some i, implies that $\underline{x} \leq \underline{y}$. As a comonotonic set is simultaneously non-decreasing in each component, it is also called a non-decreasing set.

Next we define a comonotonic random vector $X = (X_1, \ldots, X_n)$. The support of a random vector X is a set $A \subseteq \mathbb{R}^n$ for which $\mathbb{P}[X \in A] = 1$.

Definition 5.1.2. A random vector $X = (X_1, \ldots, X_n)$ is comonotonic if it has comonotonic support.

In the following theorem, some equivalent characterisations are given for the comonotonicity of a random vector.

Theorem 5.1.3. *A random vector $X = (X_1, X_2, \ldots, X_n)$ is comonotonic if and only if one of the following equivalent conditions holds:*

(i) X has a comonotonic support.

(ii) X has a comonotonic copula, i.e. for all $x = (x_1, x_2, \ldots, x_n)$, we have

$$F_X(x) = \min\{F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n)\},$$

where F_X denotes the cumulative distribution function (cdf) of X.

(iii) For $U \sim$ Uniform$(0,1)$, we have

$$X =_d (F^{-1}_{X_1}(U), F^{-1}_{X_2}(U), \ldots, F^{-1}_{X_n}(U)).$$

(iv) There exist a random variable Z and non-decreasing functions f_i, $i = 1, \ldots, n$, such that

$$X =_d (f_1(Z), f_2(Z), \ldots, f_n(Z)).$$

Proof. This proof is taken from [44].

(i) $\Rightarrow$ (ii): Assume that X has comonotonic support B. Let $\underline{x} \in \mathbb{R}^n$ and let A_j be defined by

$$A_j = \{\underline{y} \in B : y_j \leq x_j\} \qquad j = 1, \ldots, n.$$

Due to the comonotonicity of B, there exists an i such that $A_i = \cap_{j=1}^n A_j$. Hence, we find

$$\begin{aligned} F_X(\underline{x}) &= \mathbb{P}(X \in \cap_{j=1}^n A_j) = \mathbb{P}(X \in A_i) = F_{X_i}(x_i) \\ &= \min\{F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n)\}. \end{aligned}$$

The last equality follows from $A_i \subseteq A_j$, so that $F_{X_i}(x_i) \leq F_{X_j}(x_j)$ holds for all j.

(ii) $\Rightarrow$ (iii): Now assume that $F_X(x) = \min\{F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n)\}$ for all $\underline{x} = (x_1, \ldots, x_n)$. We have for all $x \in \mathbb{R}$ and $p \in [0,1]$, that

$$F_X^{-1}(p) \leq x \Leftrightarrow p \leq F_X(x).$$

Thus,

$$\begin{aligned} \mathbb{P}(F_{X_1}^{-1}(U) \leq x_1, F_{X_2}^{-1}(U) \leq x_2, \ldots, F_{X_n}^{-1}(U) \leq x_n) &= \mathbb{P}(U \leq F_{X_1}(x_1), \ldots, U \leq F_{X_n}(x_n)) \\ &= \mathbb{P}(U \leq \min_{j=1,\ldots n}\{F_{X_j}(x_j)\}) \\ &= \min_{j=1,\ldots n}\{F_{X_j}(x_j)\}. \end{aligned}$$

(iii) $\Rightarrow$ (iv): Straightforward.

(iv) $\Rightarrow$ (i): Assume that there exists a random variable Z with support B, and non-decreasing functions f_i, $i = 1, \ldots, n$, such that

$$X =_d (f_1(Z), f_2(Z), \ldots, f_n(Z)).$$

The set of possible outcomes of X is $\{f_1(z), f_2(z), \ldots, f_n(z) : z \in B\}$, which is obviously comonotonic. This implies that X is comonotonic. □

Definition 5.1.4. A family $\mathcal{C}$ of subsets of Ω is called a *chain* if for all $C_1, C_2 \in \mathcal{C}$ either $C_1 \subseteq C_2$ or $C_2 \subseteq C_1$.

The following proposition and its proof is taken from Parker [129] and it characterises comonotonicity in terms of chains.

Proposition 5.1.5. *The functions $f : \Omega \to \mathbb{R}$ and $g : \Omega \to \mathbb{R}$ are comonotonic if and only if, for all $\alpha, \beta \in \mathbb{R}$, the subsets $\{f \geq \alpha\}$ and $\{g \geq \beta\}$ form a chain.*

Proof. Suppose that for all $\alpha, \beta \in \mathbb{R}$, the sets $\{f \geq \alpha\}$ and $\{g \geq \beta\}$ form a chain. Let $s, t \in \Omega$ be such that $f(s) > f(t)$ and set $\alpha = f(s)$ and $\beta = g(t)$. Since $t \notin \{f \geq \alpha\}$ and $t \in \{g \geq \beta\}$, it must be that $\{f \geq \alpha\} \subseteq \{g \geq \beta\}$. It follows that $s \in \{g \geq \beta\}$, that is $g(s) \geq g(t)$. Thus, f and g are comonotonic.

Conversely, suppose that there exists $\alpha, \beta \in \mathbb{R}$ such that $\{f \geq \alpha\}$ and $\{g \geq \beta\}$ do not form a chain. Then there exists $s, t \in \Omega$ such that $s \in \{f \geq \alpha\} \backslash \{g \geq \beta\}$ and $t \in \{g \geq \beta\} \backslash \{f \geq \alpha\}$. This means that $f(s) \geq \alpha > f(t)$ and $g(t) \geq \beta \geq g(s)$, i.e. f and g are not comonotonic. □

Denneberg [43] defines comonotonicity of functions in terms of chains.

Definition 5.1.6. A class $\mathcal{C}$ of functions from Ω to $[-\infty, \infty]$ is called comonotonic if

$$\bigcup_{X \in \mathcal{C}} \Lambda_X$$

is a chain, where $\Lambda_X = \{\{X > x\} : x \in [-\infty, \infty]\} \cup \{\{X \geq x\} : x \in [-\infty, \infty]\}$.

Clearly a class of functions $\mathcal{C}$ is comonotonic if and only if each pair of functions in $\mathcal{C}$ is comonotonic. The following proposition gives equivalent conditions for a pair of functions to be comonotonic.

Proposition 5.1.7. *For two functions $X, Y : \Omega \to [-\infty, \infty]$, the following conditions are equivalent.*

(i) X and Y are comonotonic.

(ii) There is no pair $\omega_1, \omega_2 \in \Omega$ such that $X(\omega_1) > X(\omega_2)$ and $Y(\omega_1) < Y(\omega_2)$.

(iii) For all $\omega_1, \omega_2 \in \Omega$, we have

$$(X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2)) \geq 0.$$

For the following two conditions, we suppose that X and Y are real-valued.

(iv) There exists a function $Z : \Omega \to \mathbb{R}$ and increasing functions u and v on $\mathbb{R}$ such that $X = u(Z)$ and $Y = v(Z)$.

(v) There exist continuous and increasing functions u and v on $\mathbb{R}$ such that $u(z) + v(z) = z$ for $z \in \mathbb{R}$, $X = u(X + Y)$ and $Y = v(X + Y)$.

Proof. (i) $\Rightarrow$ (ii): Assume there exists $\omega_1, \omega_2 \in \Omega$ such that $X(\omega_1) < X(\omega_2)$ and $Y(\omega_1) > Y(\omega_2)$. Defining $A = \{X > X(\omega_1)\}$ and $B = \{Y > Y(\omega_2)\}$, we get $\omega_2 \in A \backslash B$ and $\omega_1 \in B \backslash A$, contradicting (i).

(ii) $\Rightarrow$ (i): Assume there exists $A \in \Lambda_X$ and $B \in \Lambda_Y$ such that $A \not\subseteq B$ and $B \not\subseteq A$. Then, choosing $\omega_1 \in A \backslash B$ and $\omega_2 \in A \backslash B$, we get $X(\omega_1) > a \geq X(\omega_2)$ in the case $A = \{X > a\}$ and $X(\omega_1) \geq a > X(\omega_2)$ in the case $A = \{X \geq a\}$. In any case, we have $X(\omega_1) > X(\omega_2)$ and similarly $Y(\omega_1) < Y(\omega_2)$, contradicting (ii).

(ii) $\Leftrightarrow$ (iii): Straightforward.

Now assume X, Y are real-valued. The implications (v) $\Rightarrow$ (iv) $\Rightarrow$ (ii) are trivial.

(ii) $\Rightarrow$ (v): Let $Z = X + Y$. First, we define u, v on $Z(\Omega)$. We show that $z \in Z(\Omega)$ possesses a unique decomposition

$$z = x + y$$

for some $x \in X(\omega)$, $y = Y(\omega)$ and $\omega \in \Omega$. Define $u(z) = x$ and $v(z) = y$. Only the uniqueness of the decomposition has to be checked. Assume there are $\omega_1, \omega_2 \in \Omega$ such that

$$X(\omega_1) + Y(\omega_1) = z = X(\omega_2) + Y(\omega_2).$$

Then $X(\omega_1) - X(\omega_2) = -(Y(\omega_1) - Y(\omega_2))$. From (ii) we get that the last expression has to be zero, i.e. $X(\omega_1) = X(\omega_2)$ and $Y(\omega_1) = Y(\omega_2)$, hence uniqueness.

Next we check if u and v are increasing. Take $z_1, z_2 \in Z(\Omega)$ with $z_1 < z_2$. Then, there are $\omega_1, \omega_2 \in \Omega$ such that

$$X(\omega_1) + Y(\omega_1) = z_1 < z_2 = X(\omega_2) + Y(\omega_2).$$

Then $X(\omega_1) - X(\omega_2) < -(Y(\omega_1) - Y(\omega_2))$. This inequality is compatible with (ii) only if

$$X(\omega_1) - X(\omega_2) \leq 0 \quad \text{ and } Y(\omega_1) - Y(\omega_2) \leq 0,$$

i.e. $u(z_1) \leq u(z_2)$ and $v(z_1) \leq v(z_2)$.

Next, we need to prove that u, v are continuous on $Z(\Omega)$. First notice

$$u(z) \leq u(z + h) \leq u(z) + h$$

for $z, z + h \in Z(\Omega)$, $h > 0$. The first inequality follows from the monotonicity of u, the second follows from

$$\begin{aligned} z + h &= u(z + h) + v(z + h) \\ &\geq u(z + h) + v(z) \\ &= u(z + h) + z - u(z). \end{aligned}$$

Similarly, $u(z) - h \leq u(z-h) \leq u(z)$ for $z, z-h \in Z(\Omega)$, $h > 0$. These inequalities together imply the continuity of u at z. Similarly, we can show that v is continuous.

It remains to show that u, v can be extended continuously from $Z(\Omega)$ to $\mathbb{R}$. We first extend to the closure $\overline{Z(\Omega)}$. If z is only a one-sided boundary point, the continuous extension generates no problem since we deal with increasing functions. If z can be approximated from both sides through points of $Z(\Omega)$, the above inequalities imply that the left sided and right sided continuous extensions coincide. Finally the extension of u and v from $\overline{Z(\Omega)}$ to $\mathbb{R}$ is done linearly on each connected component of $\mathbb{R} \backslash \overline{Z(\Omega)}$ in order to maintain the condition $u(z) + v(z) = z$. □

If X and Y are comonotonic, the outcomes of X and Y always move in the same direction (good or bad), thus there is no hedge or diversifiability when pooling the two risks.

The following lemma is taken from Parker [129].

Lemma 5.1.8.

(i) *Any function f and any constant function are comonotonic.*

(ii) *If f and g are comonotonic then so are αf and g for all $\alpha > 0$.*

(iii) *If f and h are comonotonic and g and h are comonotonic, then $f + g$ and h are comonotonic.*

(iv) *If f, g and h are pairwise comonotonic, then* $\max\{f, g\}$ *and h are comonotonic, as are* $\min\{f, g\}$ *and h.*

(v) *Comonotonicity is not transitive as can be seen by considering the indicator functions of subsets of Ω.*

(vi) *Let $A, B \in \Omega$. Then $\mathbb{1}_A$ and $\mathbb{1}_B$ are comonotonic if and only if $A \subseteq B$ or $B \subseteq A$.*

5.1.2 The Choquet integral

Let $\mu : \mathcal{P}(\Omega) \to \mathbb{R}_+$ be a monotone set function, where $\mathcal{P}(\Omega)$ denotes the power set of Ω, and let $X : \Omega \to [-\infty, \infty]$ be an arbitrary function on Ω. Then the function

$$S_X(x) = S_{\mu,X}(x) = \mu(X > x)$$

is decreasing and is called the *(decreasing) distribution function* of X with respect to μ. The pseudo-inverse function $\tilde{S}_{\mu,X}$ of $S_{\mu,X}$, also called the *quantile function* of X with respect to

μ, is given by

$$\tilde{S}_{\mu,X}(p) = \inf\{x : S_{\mu,X}(x) \geq p\}.$$

Let $\mu : \mathcal{S} \to [0, \infty]$ be a monotone set function and assume $\Omega \in \mathcal{S} \subseteq \mathcal{P}(\Omega)$. Given an upper μ-measurable function $X : \Omega \to [-\infty, \infty]$ with decreasing distribution function $S_{\mu,X}$ and quantile function $\tilde{S}_{\mu,X}$, we define the *asymmetric integral* of X with respect to μ by

$$\int X d\mu = \int_0^{\mu(\Omega)} \tilde{S}_{\mu,X}(t)\; dt.$$

This integral is often called the *Choquet integral* after Gustave Choquet [29]. The asymmetric integral can also be expressed in terms of the distribution function as follows

$$\int X d\mu = \int_0^{\infty} S_{\mu,X}(x)\; dx \quad \text{if } X \geq 0.$$

This is Choquet's [29] original definition. For arbitrary upper μ-measurable X, we have

$$\int X\; d\mu = \int_{-\infty}^{0} (S_{\mu,X}(x) - \mu(\Omega))\; dx + \int_0^{\infty} S_{\mu,X}(x)\; dx \quad \text{if } \mu(\Omega) < \infty.$$

If $\mu(\Omega) = \infty$, then the asymmetric integral can only be defined for $X \geq 0$.

Proposition 5.1.9. *If μ is a monotone set function on $\mathcal{P}(\Omega)$ and $X, Y : \Omega \to [-\infty, \infty]$ are functions, then the Choquet integral has the following properties:*

(i) $\int \mathbb{1}_A\; d\mu = \mu(A)$ *for* $A \in \mathcal{P}(\Omega)$.

(ii) Positive homogeneity, i.e. $\int cX\; d\mu = c \int X\; d\mu$ *if* $c \geq 0$.

(iii) Monotonicity, i.e. $X \leq Y$ *implies* $\int X\; d\mu \leq \int Y\; d\mu$.

(iv) Asymmetry, i.e. if μ is finite, then

$$\int (-X)\; d\mu = -\int X\; d\bar{\mu}, \text{ where } \bar{\mu}(A) = \mu(\Omega) - \mu(A^c).$$

(v) +-Translation invariance, i.e. $\int (X + c)\; d\mu = \int X\; d\mu + c\mu(\Omega)$ *for* $c \in \mathbb{R}$.

(vi) Comonotonic additivity, i.e. if X and Y are comonotonic and real-valued, then

$$\int (X + Y)\; d\mu = \int X\; d\mu + \int Y\; d\mu.$$

Remark 5.1.10. We call property (v) '+-translation invariance' as not to confuse with the translation invariance property in Chapter 4.

Let $\mathcal{S} \subseteq \mathcal{P}(\Omega)$ be a set and $\mu : \mathcal{S} \to [0, \infty]$ a monotone set function. As explained in Appendix A, a function $X : \Omega \to [-\infty, \infty]$ can be decomposed into its positive and negative parts

$$X = X^+ - X^-, \quad X^+ = 0 \vee X, \quad X^- = (-X)^+.$$

If X is μ-measurable, then X^+ and $-X^-$ are upper μ-measurable. Since X^+ and $-X^-$ are comonotonic functions, we have for X real-valued

$$\begin{aligned} \int X \, d\mu &= \int X^+ \, d\mu + \int (-X^-) \, d\mu \\ &= \int X^+ \, d\mu - \int X^- \, d\bar{\mu} \quad \text{if } \mu(\Omega) < \infty. \end{aligned}$$

5.1.3 Integral representation theorems

Let $(\Omega, \mathcal{F})$ be a measurable space. The set of all monotonic set functions $\mu : \mathcal{F} \to [0, 1]$, which are normalised to $\mu(\Omega) = 1$, is denoted by $\mathcal{M}_{1,m}$. Let $\phi : \mathcal{X} \to \mathbb{R}$. We define the following properties:

(i) Comonotonic additivity, i.e. if X, Y are comonotonic, then

$$\phi(X + Y) = \phi(X) + \phi(Y).$$

(ii) Comonotonic convexity, i.e. if X and Y are comonotonic, then

$$\phi(\lambda X + (1 - \lambda)Y) \leq \lambda\phi(X) + (1 - \lambda)\phi(Y).$$

Let $\mu \in \mathcal{M}_{1,m}$. For $X \in \mathcal{X}$, the Choquet integral of X with respect to μ can also be written as

$$\begin{aligned} \mu(X) &= \int_{-\infty}^{0} [\mu(X \geq x) - \mu(\Omega)] \, dx + \int_{0}^{\infty} \mu(X \geq x) \, dx \qquad (5.1) \\ &= \int_{-\infty}^{0} [\mu(X \geq x) - 1] \, dx + \int_{0}^{\infty} \mu(X \geq x) \, dx. \end{aligned}$$

Given a family $\mathcal{G}$ of functions $X : \Omega \to [-\infty, \infty]$ and a functional $\Gamma : \mathcal{G} \to [-\infty, \infty]$, a monotone set function γ on $\mathcal{P}(\Omega)$ is said to *represent* Γ if

$$\Gamma(X) = \int X d\gamma, \quad X \in \mathcal{G}.$$

The following is known as Greco's representation theorem [72].

Theorem 5.1.11. *Given a family $\mathcal{G}$ of functions on Ω with properties* *(i)* $X \geq 0$ *for all* $X \in \mathcal{G}$*, and*

(ii) $aX, X \wedge a, X - X \wedge a \in \mathcal{G}$ *if* $X \in \mathcal{G}$ *and* $a \in \mathbb{R}^+$*,*

and given a monotonic functional $\Gamma : \mathcal{G} \to \mathbb{R}$ *which satisfies comonotonic additivity and the following continuity properties*

(i) $\lim_{a \searrow} \Gamma(X - X \wedge a) = \Gamma(X)$ *for* $X \in \mathcal{G}$*,* $X \geq 0$ *(lower marginal continuity), and*

(ii) $\lim_{b \to \infty} \Gamma(X \wedge b) = \Gamma(X)$ *for* $X \in \mathcal{G}$*,*

then there exists a monotone set function $\gamma : \mathcal{P}(\Omega) \to \mathbb{R}$*, which represents* Γ*.*

Song and Yan [147] use a particular version of Greco's representation theorem, applied to bounded functions.

Theorem 5.1.12. *Let* $\mathcal{H}$ *be a family of bounded functions on* Ω *with properties*

(i) $aX, X \wedge a, X - X \wedge a \in \mathcal{H}$ *if* $X \in \mathcal{H}$ *and* $a \in \mathbb{R}^+$*,*

(ii) $X + 1 \in \mathcal{H}$ *for* $X \in \mathcal{H}$*.*

Then, if Γ *is a real functional on* $\mathcal{H}$ *with* $\Gamma(1) = 1$ *and satisfying monotonicity, positive homogeneity and comonotonic additivity, then there exists* $\gamma \in \mathcal{M}_{1,m}$ *representing* Γ *in the sense that* $\gamma(X) = \Gamma(X)$ *for all* $X \in \mathcal{H}$*.*

5.2 Choquet pricing

Choquet pricing has recently been introduced as an alternative to traditional pricing principles in insurance (see [159]) and in finance (see [24]). The Choquet integral, which is a non-linear generalisation of the Lebesgue integral, has several properties that cause it to be especially suitable for pricing insurance contracts or financial assets. Choquet pricing implies that an insurance contract or a financial asset with payoff X is priced by taking the Choquet integral of X with respect to a concave capacity.

Definition 5.2.1. Let S be any set and H any subset of S. A set function $\nu : H \to [0, \infty)$ is a *capacity* if $\nu(\emptyset) = 0$, $\nu(\Omega) = 1$ and ν satisfies monotonicity with respect to set inclusion, i.e. $A \subseteq B \Rightarrow \nu(A) \leq \nu(B)$. A set function ν is *concave* if for all $A, B \in H$

$$\nu(A \cup B) + \nu(A \cap B) \leq \nu(A) + \nu(B).$$

It is easily verified, that for a non-decreasing function $g : [0,1] \to [0,1]$ with $g(0) = 0$ and $g(1) = 1$, and a probability measure Q, the set function $\nu = g \circ Q$ is a capacity. The function g is called a *distortion function* or a *distortion operator*, as it transforms a probability distribution F_X to a new distribution $g(F_X)$.

One can define a Choquet integral with respect to a capacity by replacing μ by the capacity ν in Equation (5.1).

The intuition behind using a Choquet integral as a pricing method is shown by the following. Let F_X be the cumulative distribution function of a random variable X, i.e. $F_X(x) = P(X \leq x)$. The decumulative distribution function is given by

$$S_X(x) = P(X > x) = 1 - F_X(x).$$

Then S_X is a non-increasing function from $\mathbb{R}^+$ to $[0,1]$.

Proposition 5.2.2. *The expected value of a random variable X can be written as*

$$\mathbb{E}[X] = \int_{-\infty}^{0} (S_X(x) - 1)\, dx + \int_{0}^{\infty} S_X(x)\, dx.$$

Proof. Firstly, let $X \geq 0$. Then X can be written as $X = \int_0^\infty \mathbb{1}_{\{X>u\}}\, du$. Using Fubini's theorem to change the order of integration, we have

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^\infty \mathbb{E}[\mathbb{1}_{\{X>u\}}]\, du \\
&= \int_0^\infty P(X > u)\, du \\
&= \int_0^\infty S_X(u)\, du.
\end{aligned}$$

Next, let $X \leq 0$. As above, X has the representation $X = \int_{-\infty}^0 -\mathbb{1}_{\{X<u\}}\, du$. Then

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^0 -\mathbb{E}[\mathbb{1}_{\{X<u\}}]\, du \\
&= -\int_{-\infty}^0 P(X < u)\, du \\
&= \int_{-\infty}^0 \Big(P(X > u) - 1\Big)\, du \\
&= \int_{-\infty}^0 \Big(S_X(u) - 1\Big)\, du.
\end{aligned}$$

As $X = X^+ - X^-$, the result follows. □

Alternatively, the above proposition can be restated in terms of the cdf.

Proposition 5.2.3. *The expected value of a random variable X can be written as*

$$\mathbb{E}[X] = -\int_{-\infty}^{0} F_X(x)\,dx + \int_{0}^{\infty} [1 - F_X(x)]\,dx.$$

This expansion of the expected value leads to a different method of calculating the risk adjusted premium for a risk.

Definition 5.2.4. The risk adjusted premium $H_g[\cdot]$ with respect to a distortion operator g is given by the Choquet integral

$$H_g[X] = \int_{-\infty}^{0} (g(S_X(x)) - 1)\,dx + \int_{0}^{\infty} g(S_X(x))\,dx \tag{5.2}$$

for any risk X with decumulative distribution function S_X.

Another method of characterising the transformation of a probability distribution is using the Radon-Nikodým derivative. The following proposition shows how a continuous distortion function and the Radon-Nikodým derivative are related.

Proposition 5.2.5. *Given any continuous and differentiable distortion function*

$$S_X^*(x) = g(S_X(x)),$$

its Radon-Nikodým derivative is given by the derivative of g with respect to x evaluated at $S_X(x)$.

Proof. Firstly, $S_X^*(x) = 1 - F_X^*(x)$, thus $\frac{d}{dx}S_X^*(x) = -f_X^*(x)$. As $S_X^*(x) = g(S_X(x))$, we get

$$\frac{d}{dx}S_X^*(x) = -f_X^*(x) = \frac{d}{dx}g(S_X(x)).$$

Using the chain rule,

$$-f_X^*(x) = \frac{dg}{dx}(S_X(x)) \cdot \frac{dS_X}{dx}(x).$$

Since, $\frac{dS_X}{dx}(x) = \frac{d}{dx}(1 - F_X(x)) = -f_X(x)$, we have

$$\frac{f_X^*(x)}{f_X(x)} = \frac{dg}{dx}(S_X(x)),$$

which proves the proposition. □

Note that the Radon-Nikodým derivative of a distortion function can also be given by the derivative of g with respect to x evaluated at $F_X(x)$. The proof follows similarly.

This theorem provides a way to compute the distorted probability density function (pdf),

$$f_X^*(x) = \frac{d}{dx}F_X^*(x) = \frac{d}{dx}g(F_X(x)) = f(x)g'(F_X(x)),$$

where f is the original probability density function of X.

Expressing an asset as a negative loss, it can be shown that

$$H_g[-A] = -H_{g^*}[A],$$

where $g^*(u) = 1 - g(1-u)$ is called the *dual distortion operator*.

The dual distortion operator links the distortion of the decumulative distribution function to the distortion of the cumulative distribution function, i.e. $S_X^*(x) = g(S_X(x))$ if and only if $F_X^*(x) = g^*(F_X(x))$.

5.3 Stochastic differential equations

Assume that the evolution of traded assets in the economy can be described by stochastic differential equations. Suppose we have a traded asset X_t, whose price follows the stochastic differential equation

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \tag{5.3}$$

where W_t is a Brownian motion under the 'real world' measure P, $\mu(t, X_t)$ is referred to as the drift and $\sigma(t, X_t)$ the volatility. The dW_t-term can be interpreted as a noise term with $\mathbb{E}[dW_t] = 0$ and $\mathbb{E}[(dW_t)^2] = dt$. Note that the paths of Brownian motions are continuous functions of time that are nowhere differentiable. Thus, the usual rules of integration and differentiation cannot be applied to Brownian motion.

The most common process, that a traded asset follows, is called *geometric Brownian motion*, which is given by

$$dX_t = \mu X_t dt + \sigma X_t dW_t.$$

Here, X_t is lognormally distributed with parameters μ and σ.

A stochastic differential equation with a zero dt-term is a martingale.

A financial derivative is an asset in the market, whose value at some time T in the future is given by a function $f(X_T)$. The expected value of the payoff $f(X_T)$ at time $t \leq T$ is given by $u(t, x) = \mathbb{E}[f(X_T)|X_t = x]$, where u satisfies the Kolmogorov backward equation, defined next.

Definition 5.3.1. Let the stochastic differential equation for the process $x(t)$ be given by Equation (5.3). The *Kolmogorov backward equation* for a function $u(t,x)$ is given by

$$\frac{\partial}{\partial t}u(t,x) + \mu(t,x)\frac{\partial}{\partial x}u(t,x) + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2}{\partial x^2}u(t,x) = 0,$$

for $t \leq T$ with boundary condition $u(T,x) = f(x)$.

Definition 5.3.2 (Itô's lemma)**.** Let the stochastic differential equation for the process $x(t)$ be given by Equation (5.3). Let the function $g(t,x)$ be continuous and twice differentiable. Define $Y_t = g(t,x)$, then the stochastic differential equation for Y_t is given by

$$\begin{aligned} dY_t = &\Big[\frac{\partial}{\partial t}g(t,x) + \mu(t,x)\frac{\partial}{\partial x}g(t,x) + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2}{\partial x^2}g(t,x)\Big]dt \\ &+ \sigma(t,x)\frac{\partial}{\partial x}g(t,x)dW_t. \end{aligned}$$

Lastly, we state a fundamental result of stochastic analysis, called Girsanov's theorem. Girsanov's theorem states that if we change the probability distribution of a stochastic process, then we only change the *dt*-coefficient of the stochastic differential equation.

Theorem 5.3.3 (Girsanov)**.** *For any stochastic process K_t, such that $\int_0^t K_s^2 ds < \infty$ a.s., consider the stochastic process*

$$R_t = e^{\int_0^t K_s dW_s - \frac{1}{2}\int_0^t K_s^2 ds},$$

where W_t is a Brownian motion under the probability measure Q. Define the probability measure Q^* as $dQ^* = R_t dQ$. Then

$$W_t^* = W_t - \int_0^t K_s ds$$

is a Brownian motion under Q^*.

The process R_t is often called the Radon-Nikodým derivative and K_t the *Girsanov kernel* or *Girsanov exponent*. To see what effect a change in probability has on a stochastic process, we proceed as follows. The change in probability measure results in a Radon-Nikodým derivative R_t. Applying Itô's lemma to R_t, we obtain its stochastic differential equation and can infer the Girsanov kernel K_t. The stochastic differential equation for X_t under the new probability measure Q^* can now be obtained by substituting $dW_t = dW_t^* + K_t dt$. For example, the stochastic process X_t, which follows (5.3) under Q, follows

$$\begin{aligned} dX_t &= \mu(t,X_t)dt + \sigma(t,X_t)(dW_t^* + K_t dt) \\ &= \Big(\mu(t,X_t) + \sigma(t,X_t)K_t\Big)dt + \sigma(t,X_t)dW_t^*, \end{aligned} \tag{5.4}$$

under Q^*. The change in measure only changed the coefficient of the dt-term.

For arbitrage-free pricing, we need the discounted price process to be a martingale under the risk-neutral measure. Hence, we require Equation (5.4) to have zero drift. The only possible choice for the Girsanov kernel is thus

$$K_t = -\frac{\mu(t, X_t)}{\sigma(t, X_t)}. \tag{5.5}$$

This expression is known as the *market price of risk*.

5.4 The Wang transform

The Wang transform was defined by Wang [159], as a universal pricing method for pricing both financial and actuarial risk. This transform connects the traditional actuarial standard deviation loading principle, Yaari's economic theory of risk [164], CAPM and the option-pricing theory (see [159] for more details).

Definition 5.4.1. Let Φ denote the standard normal cumulative distribution function, i.e. $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds$, and let $\alpha \in \mathbb{R}$. Then $g_\alpha : [0,1] \to [0,1]$, defined by

$$g_\alpha(p) = \Phi[\Phi^{-1}(p) + \alpha],$$

defines a distortion operator, called the *Wang transform*.

Assuming that X follows a normal distribution, this operator shifts the p^{th} quantile of X, i.e. the value y such that $P(X \leq y) = p$, by a positive or negative value α, and then re-evaluates the normal cumulative probability for the shifted quantile. The parameter α can be viewed as the market price of risk. By shifting the quantiles to the left, i.e. assigning higher probabilities to lower values, the Wang transform allows for pessimistic behaviour and by shifting the quantiles to the right, it allows for more optimistic attitudes.

The first derivative of the Wang transform is given by

$$\frac{dg_\alpha(p)}{dp} = e^{-\alpha p - \frac{\alpha^2}{2}} > 0$$

and the second derivative by

$$\frac{d^2 g_\alpha(p)}{dp^2} = -\alpha\sqrt{2\pi}\, e^{\frac{p}{2} - \alpha p - \frac{\alpha^2}{2}}.$$

Thus g_α is continuous and increasing for all α, convex for $\alpha < 0$ and concave for $\alpha > 0$.

Wang [159] uses the following Choquet integral with respect to g_α to define the risk adjusted premium $H[\cdot, \alpha]$ of a risk X:

$$H[X, \alpha] = \int_{-\infty}^{0} [g_\alpha(S_X(x)) - 1]\, dx + \int_{0}^{\infty} g_\alpha(S_X(x))\, dx. \tag{5.6}$$

It is easy to verify that the Choquet integral with respect to the Wang transform $H[X, \alpha]$ has the following properties:

(i) $\min(X) \leq H[X, \alpha] \leq \max(X)$.

(ii) $H[X, \alpha]$ is an increasing function of α.

(iii) $H[c, \alpha] = c$ and $H[X + c, \alpha] = H[X, \alpha] + c$.

(iv) If $b > 0$, then $H[bX, \alpha] = bH[X, \alpha]$ and if $b < 0$, then $H[bX, \alpha] = bH[X, -\alpha]$.

(v) If $\alpha < 0$, then $H[X, \alpha] < \mathbb{E}[X]$, and if $\alpha > 0$, then $H[X, \alpha] > \mathbb{E}[X]$.

(vi) If X_1 and X_2 are comonotonic, then $H[X_1 + X_2, \alpha] = H[X_1, \alpha] + H[X_2, \alpha]$.

(vii) If $\alpha > 0$, then $H[X_1 + X_2, \alpha] \leq H[X_1, \alpha] + H[X_2, \alpha]$, and if $\alpha < 0$, then $H[X_1 + X_2, \alpha] \geq H[X_1, \alpha] + H[X_2, \alpha]$.

Remark 5.4.2. Note that $1 - g_\alpha(1 - u) = g_{-\alpha}(u)$. Thus, $1 - g_{-\alpha}(F_X(x)) = 1 - g_{-\alpha}(1 - S_X(x)) = g_\alpha(S_X(x))$. The dual distortion operator of g_α is, hence, given by

$$g^*_\alpha(p) = g_{-\alpha}(p).$$

This implies that when valuing an asset, we need to work with $-\alpha$. Hence, the risk adjusted premium defined in (5.6), can also be written as

$$\pi_X^{WT}(\alpha) = -\int_{-\infty}^{0} g_{-\alpha}(F_X(x))\, dx + \int_{0}^{\infty} (1 - g_{-\alpha}(F_X(x)))\, dx. \tag{5.7}$$

Wang's asset pricing approach involves applying $H[X, \alpha]$ to the present value of an asset to get the risk-neutral price. Underlying this is the assumption that the certainty equivalent of a risky cash flow is the market price and $H[X, \alpha]$ can be seen as the certainty equivalent. In order to derive pricing results using $H[X, \alpha]$, the α is calibrated in such a way that the certainty equivalent is consistent with the market price of the asset. More formally, α is determined in such a way that the martingale condition is fitted for all T, i.e. solves the following equation

$$\mathbb{E}^{WT}[X_T | \mathcal{F}_t] = X_t,$$

where $\mathbb{E}^{WT}[\cdot]$ refers to the conditional expectation under the Wang transform distribution defined above.

Proposition 5.4.3.

(i) If $X \sim Normal(\mu, \sigma^2)$, then $g_\alpha(S_X) \sim Normal(\mu + \alpha\sigma, \sigma^2)$.

(ii) If $X \sim Lognormal(\mu, \sigma^2)$, then $g_\alpha(S_X) \sim Lognormal(\mu + \alpha\sigma, \sigma^2)$.

Proof. (i): For $X \sim \text{Normal}(\mu, \sigma^2)$, we have

$$\begin{aligned} S_X(t) = P(X > t) &= 1 - P\Big(\frac{X-\mu}{\sigma} \leq \frac{t-\mu}{\sigma}\Big) \\ &= 1 - \Phi[\frac{t-\mu}{\sigma}] \\ &= \Phi[-\frac{t-\mu}{\sigma}] \end{aligned}$$

by symmetry of the standard normal. Then

$$\begin{aligned} g_\alpha(S_X(t)) &= \Phi[\Phi^{-1}[S_X(t)] + \alpha] \\ &= \Phi\Big[\Phi^{-1}\Big(\Phi[-\frac{t-\mu}{\sigma}]\Big) + \alpha\Big] \\ &= \Phi[-\frac{t-\mu-\alpha\sigma}{\sigma}] = S_Z(t), \end{aligned}$$

where $Z \sim \text{Normal}(\mu + \alpha\sigma, \sigma^2)$. Also

$$\begin{aligned} H[X, \alpha] &= \int_{-\infty}^{0} [g_\alpha(S_X(t)) - 1]\, dt + \int_{0}^{\infty} g_\alpha(S_X(t))\, dt \\ &= \int_{-\infty}^{0} [S_Z(t) - 1]\, dt + \int_{0}^{\infty} S_Z(t)\, dt \\ &= \mathbb{E}[Z] = \mu + \alpha\sigma. \end{aligned}$$

(ii): Follows similarly to (i), replacing t by $\ln t$. □

Proposition 5.4.4. *The Radon-Nikodým derivative corresponding to the Wang transform is given by*

$$RND_g(x) = e^{-\alpha\Phi^{-1}(F_X(x)) - \frac{1}{2}\alpha^2}.$$

Proof. Let

$$f(x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$$

and $h(x) = \Phi^{-1}(x)$ such that $f(h(x)) = h(f(x)) = x$. Note that $\Phi(x)$ can be written as

$$\Phi(x) = \frac{1}{2}\Big[\frac{2}{\sqrt{\pi}} \int_{0}^{\frac{x}{\sqrt{2}}} e^{-t^2} dt + 1\Big].$$

Also, note that $\frac{d}{dx}(f(h(x))) = 1$. But

$$\begin{aligned}\frac{d}{dx}(f(h(x))) &= \frac{df}{dx}(h(x)) \cdot \frac{dh}{dx} \\ &= \frac{d}{dx}\Big(\frac{1}{2}\Big[\frac{2}{\sqrt{\pi}}\int_0^{\frac{x}{\sqrt{2}}} e^{-t^2}dt + 1\Big]\Big)\Big|_{h(x)} \cdot \frac{dh}{dx} \\ &= \frac{1}{\sqrt{\pi}} e^{-\frac{h^2(x)}{2}} \cdot \frac{1}{\sqrt{2}} \cdot \frac{dh}{dx} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{h^2(x)}{2}} \cdot \frac{dh}{dx}.\end{aligned}$$

Thus,

$$\frac{dh}{dx} = \sqrt{2\pi} e^{\frac{1}{2}h^2(x)},$$

where $h(x) = \Phi^{-1}(x)$.

The Wang transform is given by

$$g(u) = \Phi(\Phi^{-1}(u) + \alpha) = \int_{-\infty}^{\Phi^{-1}(u)+\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$$

and hence, its Radon-Nikodým derivative is given by

$$\begin{aligned}g'(u) &= \frac{d}{du}\int_{-\infty}^{\Phi^{-1}(u)+\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\Phi^{-1}(u)+\alpha)^2} \cdot \frac{d}{du}(\Phi^{-1}(u) + \alpha) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\Phi^{-1}(u))^2 - \alpha\Phi^{-1}(u) - \frac{1}{2}\alpha^2} \cdot \sqrt{2\pi} e^{\frac{1}{2}(\Phi^{-1}(u))^2} \\ &= e^{-\alpha\Phi^{-1}(u) - \frac{1}{2}\alpha^2}.\end{aligned}$$

Finally, substituting $u = F_X(x)$, we get our result. □

An important question that arises, is how did Wang decide on this particular distortion? What follows next, hopefully clears up this question.

Assume the stock price follows geometric Brownian motion, i.e.

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where W_t is a Brownian motion under the real world measure P. Then the solution to this stochastic differential equation is

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}.$$

To be able to price, we need to transform our probability measure into the risk-neutral measure under which the discounted price process is a martingale. Thus, we apply Girsanov's theorem. The risk-neutral probability Q is determined by the Radon-Nikodým derivative

$$R_t = e^{-(\frac{\mu-r}{\sigma})W_t - \frac{1}{2}(\frac{\mu-r}{\sigma})^2 t}$$

and therefore, $W_t^* = W_t + (\frac{\mu-r}{\sigma})t$ is a Brownian motion under Q. The stock dynamics then become

$$dS_t = rS_t dt + \sigma S_t dW_t^*.$$

Now

$$\begin{aligned} P(S_t \le x) &= P(\ln S_0 - (\mu - \frac{\sigma^2}{2})t + \sigma W_t \le \ln x) \\ &= \Phi\Big[\frac{\ln(\frac{S_0}{x}) - (\mu - \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\Big]. \end{aligned}$$

Under the risk-neutral measure,

$$\begin{aligned} Q(S_t \le x) &= \Phi\Big[\frac{\ln(\frac{S_0}{x}) - (r - \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}\Big] \\ &= \Phi\Big[\frac{\ln(\frac{S_0}{x}) - (\mu - \frac{\sigma^2}{2})t}{\sigma\sqrt{t}} + \frac{\mu - r}{\sigma}\sqrt{t}\Big] \\ &= \Phi\Big[\Phi^{-1}[P(S_t \le x)] + \frac{\mu - r}{\sigma}\sqrt{t}\Big]. \end{aligned}$$

This derivation works for any attainable contingent claim of the form $X_T = h(S_T)$, where h is a positive, increasing function. Hence,

$$\mathbb{E}_Q[X_T] = \int_0^\infty Q(X_T > x)dx = H[X_T, -\frac{\mu - r}{\sigma}\sqrt{T}].$$

Next, let us consider the case, where the underlying security price has a lognormal distribution with time-varying drift and volatility. Let X_t be a variable following

$$dX_t = \mu(t)dt + \sigma(t)dW_t,$$

where W_t is a Brownian motion under the real-world probability P. Written more formally, we have

$$X_t = X_0 + \int_0^t \mu(s)\, ds + \int_0^t \sigma(s)\, dW_s.$$

Assume that the underlying security price is given by $S_t = e^{X_t}$, then by Itô's lemma, we have

$$dS_t = \Big(\mu(t) + \frac{1}{2}\sigma^2(t)\Big)S_t dt + \sigma(t)S_t dW_t.$$

Since $X_t \sim \text{Normal}(X_0 + \int_0^t \mu(s)\,ds, \int_0^t \sigma^2(s)\,ds)$, we obtain

$$\begin{aligned} P[S_t \le x] &= P[X_t \le \ln x] \\ &= \Phi\left[\frac{\ln x - (X_0 + \int_0^t \mu(s)\,ds)}{\sqrt{\int_0^t \sigma^2(s)\,ds}}\right]. \end{aligned}$$

Let γ_t denote the market price of risk for the security S_t. The market price of risk is determined in such a way that the security price discounted to the present value at the risk-free rate will be a martingale under the risk-neutral probability measure. Thus, γ_t will be equal to

$$\gamma_t = \frac{\mu(t) + \frac{1}{2}\sigma^2(t) - r}{\sigma(t)}.$$

By Girsanov's theorem, we have that $W_t^* = W_t + \int_0^t \gamma_s ds$ is a Brownian motion under the risk-neutral measure Q. Hence, X_t follows

$$dX_t = \Big(\mu(t) - \sigma(t)\gamma_t\Big)dt + \sigma(t)dW_t^*$$

under Q. The probability distribution under Q becomes

$$\begin{aligned} Q[S_t \le x] &= Q[X_t \le \ln x] \\ &= \Phi\left[\frac{\ln x - (X_0 + \int_0^t(\mu(s) - \sigma(s)\gamma_s)ds)}{\sqrt{\int_0^t \sigma^2(s)ds}}\right] \\ &= \Phi\left[\frac{\ln x - (X_0 + \int_0^t \mu(s)\,ds)}{\sqrt{\int_0^t \sigma^2(s)ds}} + \frac{\int_0^t \sigma(s)\gamma_s ds}{\sqrt{\int_0^t \sigma^2(s)ds}}\right] \\ &= \Phi\left[\Phi^{-1}[P[S_t \le x]] + \frac{\int_0^t \sigma(s)\gamma_s ds}{\sqrt{\int_0^t \sigma^2(s)ds}}\right]. \end{aligned} \tag{5.8}$$

Define

$$g_\gamma(u) = \Phi\left[\Phi^{-1}[u] + \frac{\int_0^t \sigma(s)\gamma_s ds}{\sqrt{\int_0^t \sigma^2(s)ds}}\right],$$

then the risk adjusted premium becomes

$$H[Y,\gamma] = \int_0^\infty g_\gamma\big(P(Y > u)\big)du$$

for any nonnegative random variable Y. From Equation (5.8), we have

$$Q(S_T > x) = g_{-\gamma}\big(P(S_t > x)\big)$$

and since $\mathbb{E}_Q[S_t] = \int_0^\infty Q(S_t > x)dx$, it follows that $\mathbb{E}_Q[S_t] = H[S_t, -\gamma]$.

5.4.1 Pricing financial risk

The following proposition simplifies matters when we are dealing with functions of standard normal random variables.

Proposition 5.4.5. *Let Z be a standard normal variable and h a continuous increasing function with range in $[0,\infty)$. If $X = h(Z)$ and $\alpha \in \mathbb{R}$, then*

$$H[X,\alpha] = \mathbb{E}[h(Z+\alpha)].$$

Proof. We have

$$H[X,\alpha] = \int_0^\infty g_\alpha(S_X(t))\,dt.$$

Now

$$S_X(t) = P(X > t) = P(h(Z) > t) = P(Z > h^{-1}(t)) = \Phi(-h^{-1}(t))$$

and

$$\begin{aligned} g_\alpha(S_X(t)) &= \Phi[\Phi^{-1}(\Phi(-h^{-1}(t))) + \alpha] \\ &= \Phi[-h^{-1}(t) + \alpha] \\ &= P(Z > h^{-1}(t) - \alpha) \\ &= P(h(Z+\alpha) > t). \end{aligned}$$

Therefore

$$H[X,\alpha] = \int_0^\infty g_\alpha(S_X(t))\,dt = \mathbb{E}[h(Z+\alpha)].$$

□

Corollary 5.4.6. *Let Z be a standard normal variable and h be of the form*

$$h(x) = \begin{cases} 0 & \text{for } x \in [0,a] \\ h_2(x) & \text{for } x \in (a,\infty), \end{cases}$$

where h_2 is a continuous increasing function with range in $[0,\infty)$ and $a \in \mathbb{R}$. If $X = h(Z)$ and $\alpha \in \mathbb{R}$, then

$$H[X,\alpha] = \mathbb{E}[h_2(Z+\alpha)].$$

Proof. Since

$$P(h(Z) > t) = P(h_2(Z) > t),$$

the corollary follows easily by suitably adapting the proof of Proposition 5.4.5. □

This corollary will be required to price binary options.

Note that if we replace the standard normal cumulative distribution function in the Wang transform by the cumulative distribution function F of a symmetric distribution, i.e.

$$g_\alpha(x) = F(F^{-1}(x) + \alpha),$$

then the above proposition and corollary still hold.

Proposition 5.4.5 allows for the easy computation of prices of claims that can be written as functions of standard normal random variables.

Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ and let S_t be the price of a security at time t. We assume that S_t follows geometric Brownian motion, i.e.

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where W_t is a Brownian motion under the real-world measure P. Since S_t follows geometric Brownian motion, it is well known that $S_T = h(Z)$, where

$$h(Z) = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma\sqrt{t}Z}.$$

Hence, using Proposition 5.4.5,

$$\begin{aligned} H[S_T, -\alpha] &= \mathbb{E}[h(Z - \alpha)] \\ &= \mathbb{E}[S_0 e^{(\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}Z - \sigma\sqrt{T}\alpha}] \\ &= S_0 e^{(\mu - \frac{\sigma^2}{2})T - \sigma\sqrt{T}\alpha + \frac{\sigma^2}{2}T}. \end{aligned}$$

For the condition of no arbitrage to hold, we need the price of S_T at time 0 to be the future value of the current security price $e^{rT}S_0$. Thus $\alpha = \frac{\mu - r}{\sigma}\sqrt{T}$ and

$$S_0 = e^{-rT} H[S_T, -\alpha].$$

In other words, we used α to calibrate the discounted certainty equivalent of the security price on a future date to the initial price of the security.

Next we shall look at pricing a European call option using the Wang transform. The payoff of a European call option with strike K and maturity T is given by

$$C(T, K) = (S_T - K)^+,$$

where $(x)^+ = \max(x, 0)$. Once again we can write this payoff as a function of a standard normal random variable $C(T, K) = f(Z)$, where $f(Z) = (S_0 e^{(\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}Z} - K)^+$. Then using Proposition 5.4.5, we get

$$
\begin{aligned}
H[C(T,K),-\alpha] &= \mathbb{E}[f(Z-\alpha)] \\
&= \int_{-\infty}^{\infty} \left(S_0 e^{(\mu-\frac{\sigma^2}{2})T-\sigma\sqrt{T}\alpha+\sigma\sqrt{T}z} - K\right)^{+} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.
\end{aligned}
$$

The region of integration is determined by the values of z for which

$$
S_0 e^{(\mu-\frac{\sigma^2}{2})T-\sigma\sqrt{T}\alpha+\sigma\sqrt{T}z} \geq K.
$$

Solving for z results in

$$
z \geq \frac{\ln(\frac{K}{S_0}) - (\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}\alpha}{\sigma\sqrt{T}}.
$$

Let z_{min} denote the minimum value for which the latter inequality holds. Then

$$
\begin{aligned}
H[C(T,K),-\alpha] &= \int_{z_{min}}^{\infty} \left(S_0 e^{(\mu-\frac{\sigma^2}{2})T-\sigma\sqrt{T}\alpha+\sigma\sqrt{T}z} - K\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= S_0 e^{\mu T-\sigma\sqrt{T}\alpha} \int_{z_{min}}^{\infty} e^{-\frac{\sigma^2}{2}T+\sigma\sqrt{T}z} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz - K \int_{z_{min}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= S_0 e^{\mu T-\sigma\sqrt{T}\alpha} \int_{z_{min}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\sigma\sqrt{T})^2} dz - K \int_{z_{min}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= S_0 e^{\mu T-\sigma\sqrt{T}\alpha} \Phi(-z_{\min} + \sigma\sqrt{T}) - K\Phi(-z_{\min}).
\end{aligned}
$$

Using $\alpha = \frac{\mu-r}{\sigma}\sqrt{T}$ as calculated above, i.e. calibrating the call option price to the underlying security price, results in the Black-Scholes price of a call option at time 0. It makes sense that we should use the same α as when pricing the underlying security.

The above shows that the Wang transform is consistent with arbitrage-free pricing in the setting of geometric Brownian motion with constant coefficients.

What happens if you let the traded asset be described by a general stochastic process? Let X_t be a traded asset that follows (5.3). The question Pelsser [130] asked himself, is which probability measure, i.e. Girsanov transformation, is implied by the Wang transform?

To investigate this, we have to find the stochastic differential equation, which gives rise to the Wang measure, which we will denote by F^{WT} and, as above, is given by

$$
F^{WT}(p) = \Phi(\Phi^{-1}(F(p)) + \alpha(t,T)). \tag{5.9}
$$

Since the Girsanov transformation only affects the dt-coefficient of the stochastic process, the Girsanov transformation will give rise to the following stochastic differential equation for the process X_t:

$$
dX_t = \left(\mu(t,X_t) + \sigma(t,X_t)K_t^{WT}\right)dt + \sigma(t,X_t)dW_t^{WT},
$$

where W_t^{WT} is a Brownian motion under the 'Wang probability measure' and K_t^{WT} denotes the Girsanov kernel associated with the Wang transform.

Proposition 5.4.7. *The Wang transform is consistent with arbitrage-free pricing if and only if the following conditions on $\mu(t, X_t)$ and $\sigma(t, X_t)$ are satisfied*

$$\frac{\partial}{\partial x}\Big(\mu(t,X_t)\frac{\frac{\partial F}{\partial x}}{\phi(\Phi^{-1}(F))}\Big) = 0$$

and

$$\frac{\partial}{\partial x}\Big(\sigma(t,X_t)\frac{\frac{\partial F}{\partial x}}{\phi(\Phi^{-1}(F))}\Big) = 0,$$

where ϕ denotes the derivative of Φ, i.e. the standard normal probability density function.

Proof. The function F^{WT} has to satisfy the Kolmogorov backward equation, i.e.

$$\frac{\partial}{\partial t}F^{WT} + (\mu(t,X_t) + \sigma(t,X_t)K_t^{WT})\frac{\partial}{\partial x}F^{WT} + \frac{1}{2}\sigma^2(X_t,t)\frac{\partial^2}{\partial x^2}F^{WT} = 0. \tag{5.10}$$

This implies that

$$(\mu(t,X_t) + \sigma(t,X_t)K_t^{WT}) = \frac{\frac{\partial}{\partial t}F^{WT} + \frac{1}{2}\sigma^2(t,X_t)\frac{\partial^2}{\partial x^2}F^{WT}}{-\frac{\partial}{\partial x}F^{WT}}.$$

Using Equation (5.9), we get

$$\begin{aligned}
\frac{\partial}{\partial t}F^{WT} &= \frac{\phi(\Phi^{-1}(F)+\alpha)}{\phi(\Phi^{-1}(F))}\Big[\frac{\partial}{\partial t}F + \phi(\Phi^{-1}(F))\frac{\partial}{\partial t}\alpha\Big],\\
\frac{\partial}{\partial x}F^{WT} &= \frac{\phi(\Phi^{-1}(F)+\alpha)}{\phi(\Phi^{-1}(F))}\frac{\partial}{\partial x}F \quad \text{and}\\
\frac{\partial^2}{\partial x^2}F^{WT} &= \frac{\phi(\Phi^{-1}(F)+\alpha)}{\phi(\Phi^{-1}(F))}\Big[\frac{\alpha}{\phi(\Phi^{-1}(F))}(\frac{\partial}{\partial x}F)^2 + \frac{\partial^2}{\partial x^2}F\Big].
\end{aligned}$$

Substituting these equations into (5.10) and using the fact that F satisfies the Kolmogorov backward equation, yields

$$\sigma(t,X_t)K_t^{WT} = \Big(\frac{\phi(\Phi^{-1}(F))}{\frac{\partial F}{\partial x}}\Big)\frac{\partial\alpha}{\partial t} - \frac{1}{2}\sigma^2(t,X_t)\Big(\frac{\frac{\partial F}{\partial X}}{\phi(\Phi^{-1}(F))}\Big)\alpha.$$

To find the conditions under which the Wang transform is consistent with arbitrage-free pricing, we must substitute the market price of risk $\sigma(t,x)K_t^{WT} = -\mu(t,x)$ into the above equation. This results in the following differential equation in α

$$\frac{\partial}{\partial t}\alpha(t,T) = -\Big(\mu(t,x)\frac{\frac{\partial F}{\partial x}}{\phi(\Phi^{-1}(F))}\Big) + \frac{1}{2}\Big(\sigma(t,x)\frac{\frac{\partial F}{\partial x}}{\phi(\Phi^{-1}(F))}\Big)^2\alpha(t,T).$$

□

Note that a necessary condition for this proposition to hold, is that the ratio $\frac{\mu(t,X_t)}{\sigma(t,X_t)}$ is a function of time only. Pelsser [130] thus concludes that, since the Wang transform is in general not consistent with no arbitrage, it should not be seen as a universal framework for pricing.

5.4.2 Pricing exotic options

The Wang transform is not only consistent with the Black-Scholes option pricing formula for vanilla options, as shown by Hamada and Sherris [76], but also for more exotic options, as is shown in this section. This section is new and due to the author.

Proposition 5.4.8. *Consider an option with underlying f, maturity T, strike K and payoff $(f_T - K)^+$. If the underlying f follows geometric Brownian motion, i.e.*

$$df_t = \hat{\mu} f_t dt + \hat{\sigma} f_t dW_t,$$

where W is a Brownian motion, $\hat{\mu}$ is the drift and $\hat{\sigma}$ is the volatility, then $e^{-r(T-t)}H[f, -\alpha]$ with $\alpha = \frac{\hat{\mu}-r}{\hat{\sigma}}\sqrt{T-t}$ reduces to the equivalent Black-Scholes price.

Proof. If f follows the process $df_t = \hat{\mu} f_t dt + \hat{\sigma} f_t dW_t$, then the value of f_T at time t can be written as

$$f_T = f_t e^{(\hat{\mu}-\frac{1}{2}\hat{\sigma}^2)(T-t)+\hat{\sigma}\sqrt{T-t}z},$$

where z is a standard normal random variable.

Thus the payoff of the option can be written as

$$(f_T - K, 0)^+ = (f_t e^{(\hat{\mu}-\frac{1}{2}\hat{\sigma}^2)(T-t)+\hat{\sigma}\sqrt{T-t}z} - K, 0)^+.$$

Let

$$h(z) = (f_t e^{(\hat{\mu}-\frac{1}{2}\hat{\sigma}^2)(T-t)+\hat{\sigma}\sqrt{T-t}z} - K)^+,$$

then h is increasing and continuous. Apply Proposition 5.4.5 to obtain

$$\begin{aligned} H[f_T, -\alpha] &= \mathbb{E}[h(z-\alpha)] \\ &= \int_{-\infty}^{\infty} (f_t e^{(\hat{\mu}-\frac{1}{2}\hat{\sigma}^2)(T-t)+\hat{\sigma}\sqrt{T-t}(z-\alpha)} - K)^+ f_Z(z) dz. \end{aligned}$$

Let z_* be the value of z such that $f_T = K$, i.e

$$z_* = \frac{\ln(\frac{K}{f_t}) - (\hat{\mu} - \frac{1}{2}\hat{\sigma}^2)(T-t) + \hat{\sigma}\sqrt{T-t}\alpha}{\hat{\sigma}\sqrt{T-t}}.$$

Thus

$$\begin{aligned}
H[f_T, -\alpha] &= \mathbb{E}[h(z-\alpha)] \\
&= \int_{z_*}^{\infty} (f_t e^{(\hat{\mu}-\frac{1}{2}\hat{\sigma}^2)(T-t)+\hat{\sigma}\sqrt{T-t}(z-\alpha)} - K)^+ f_Z(z)dz \\
&= f_t e^{(\hat{\mu}-\frac{1}{2}\hat{\sigma}^2)(T-t)-\hat{\sigma}\sqrt{T-t}\alpha+\frac{1}{2}\hat{\sigma}^2(T-t)} \int_{z_*}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\hat{\sigma}\sqrt{T-t})^2} dz \\
&\quad - K \int_{z_*}^{\infty} f_Z(z)dz \\
&= f_t e^{\hat{\mu}(T-t)-\hat{\sigma}\sqrt{T-t}\alpha} \Phi[-(z_* - \hat{\sigma}\sqrt{T-t})] - K\Phi[-z_*].
\end{aligned}$$

Setting $\alpha = \frac{\hat{\mu}-r}{\hat{\sigma}}\sqrt{T-t}$, results in

$$-(z_* - \hat{\sigma}\sqrt{T-t}) = \frac{\ln(\frac{f_t}{K}) + (r + \frac{1}{2}\hat{\sigma}^2)(T-t)}{\hat{\sigma}\sqrt{T-t}},$$

which is commonly referred to as d_1 and

$$-z_* = \frac{\ln(\frac{f_t}{K}) + (r - \frac{1}{2}\hat{\sigma}^2)(T-t)}{\hat{\sigma}\sqrt{T-t}},$$

which is referred to as d_2. Then

$$H[f_T, -\alpha] = f_t e^{r(T-t)} \Phi[d_1] - K\Phi[d_2].$$

Discounting yields the Black-Scholes price. □

The following example illustrates various applications of Proposition 5.4.8.

Example 5.4.9.

(i) **Standard European option**

First, we need to ensure that the proposition holds for vanilla European options. The underlying of a standard normal European option is a stock which follows

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

This is just a straightforward application of Proposition 5.4.8.

(ii) **Margrabe option**

A Margrabe option, named after William Margrabe [122], gives the holder the right but not the obligation to exchange one asset for another. The Margrabe option can be seen as a generalisation of the Black-Scholes model. The insights into the derivation of the Margrabe option price is useful in other applications of option pricing. For example, a vanilla European option can be viewed as a Margrabe option, where the asset exchanged is cash.

Assume we have two assets S_1 and S_2, which follow

$$dS^{(i)} = \mu_i S^{(i)} dt + \sigma_i S^{(i)} dW_i, \text{ for } i = 1, 2,$$

where the W_i's are correlated Brownian motions with correlation ρ_{12}. Then the payoff of the Margrabe option at expiry T is given by $(S_T^{(1)} - S_T^{(2)}, 0)^+$. Using $S^{(2)}$ as the numéraire, we can rewrite the payoff as

$$(\frac{S_T^{(1)}}{S_T^{(2)}} - 1, 0)^+.$$

The underlying in this case is $\frac{S^{(1)}}{S^{(2)}}$ and the strike is 1.

We first need to find the process that $\frac{S_T^{(1)}}{S_T^{(2)}}$ follows. Let $f(S^{(1)}, S^{(2)}) = \frac{S^{(1)}}{S^{(2)}}$. By the multivariate version of Itô's lemma, we get

$$df = \frac{S_t^{(1)}}{S_t^{(2)}} \Big([\mu_1 - \mu_2 + \sigma_2^2 - \sigma_1 \sigma_2 \rho_{12}] dt + \sigma_1 dW_1 - \sigma_2 dW_2 \Big).$$

Let $\hat{\sigma} dW^* = \sigma_1 dW_1 - \sigma_2 dW_2$, where

$$\hat{\sigma}^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_1 \sigma_2 \rho_{12}$$

and $dW^* \sim \text{Normal}(0, \hat{\sigma}^2 dt)$. Thus,

$$df = \hat{\mu} f dt + \hat{\sigma} f dW^*,$$

where $\hat{\mu} = \mu_1 - \mu_2 + \sigma_2^2 - \sigma_1 \sigma_2 \rho_{12}$.

Note that since we used $S^{(2)}$ as the numéraire, we have that $r = 0$ and hence

$$\alpha = \frac{\hat{\mu}}{\hat{\sigma}} \sqrt{T - t}$$

will yield the Black-Scholes price.

(iii) **Geometric basket option**

A basket option is an option whose underlying is a sum or average of different assets. To price an arithmetic basket option, which uses an arithmetic average as the underlying, Monte Carlo techniques are often used. To reduce the variance in the computational price of the arithmetic basket option, the geometric basket option, which uses a geometric average as the underlying and can be priced using Proposition 5.4.8, can be used as a control variate.

Thus, we will take a closer look at a geometric basket option. The payoff of this option takes the form $(\sqrt{S^{(1)}S^{(2)}} - K, 0)^+$. The processes $S^{(1)}$ and $S^{(2)}$ follow geometric Brownian motion, i.e.

$$dS^{(i)} = \mu_i S^{(i)}dt + \sigma_i S^{(i)}dW_i, \text{ for } i = 1, 2,$$

where the W_i's are correlated Brownian motions with correlation ρ_{12}.

The underlying in this case is $\sqrt{S^{(1)}S^{(2)}}$ and the strike is K. We first need to find the process that $\sqrt{S^{(1)}S^{(2)}}$ follows. Let $f(S^{(1)}, S^{(2)}) = \sqrt{S^{(1)}S^{(2)}}$. By the multivariate version of Itô's lemma, we get

$$df_t = f_t\left(\left[\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 + \frac{1}{8}\sigma_1^2 + \frac{1}{8}\sigma_2^2 + \frac{1}{8}\sigma_1\sigma_2\rho_{12}\right]dt + \frac{1}{2}\Big(\sigma_1 dW_1 + \sigma_2 dW_2\Big)\right).$$

Let $\hat{\sigma}dW^* = \frac{1}{2}(\sigma_1 dW_1 + \sigma_2 dW_2)$, where

$$\hat{\sigma}^2 = \frac{1}{4}(\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\rho_{12})$$

and $dW^* \sim \text{Normal}(0, \hat{\sigma}^2 dt)$. Thus,

$$df_t = \hat{\mu} f_t dt + \hat{\sigma} f_t dW^*,$$

where

$$\hat{\mu} = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 + \frac{1}{8}\sigma_1^2 + \frac{1}{8}\sigma_2^2 + \frac{1}{8}\sigma_1\sigma_2\rho_{12}.$$

Setting

$$\alpha = \frac{\hat{\mu}}{\hat{\sigma}}\sqrt{T - t},$$

reduces $e^{-r(T-t)}H[f, -\alpha]$ to the Black-Scholes price.

(iv) **Asset-or-nothing**

Next we take a look at the two main types of binary options: the asset-or-nothing and the cash-or-nothing option. The binary options market can reveal the market's estimate of the current skewness in the market.

Consider an asset-or-nothing call option with strike K and maturity T. The payoff of this option is given by

$$AoN(K, T) = S_T \mathbb{1}_{\{S_T \geq K\}},$$

where S follows geometric Brownian motion with drift μ and volatility σ.

Let

$$h(z) = S_0 e^{(\mu - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}z} \mathbb{1}_{\{S_0 e^{(\mu - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}z} \geq K\}}.$$

Then h is an increasing, continuous function. Letting

$$z_* = \frac{\ln(\frac{K}{S_0}) - (\mu - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}\alpha}{\sigma\sqrt{T}},$$

we have by Proposition 5.4.5 that

$$\begin{aligned} H[AoN(K,T), -\alpha] &= \mathbb{E}[h(z-\alpha)] \\ &= \int_{z_*}^{\infty} S_0 e^{(\mu - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}(z-\alpha)} f_Z(z) dz \\ &= S_0 e^{\mu T - \sigma\sqrt{T}\alpha} \Phi[-(z_* - \sigma\sqrt{T})]. \end{aligned}$$

If $\alpha = \frac{\mu - r}{\sigma}\sqrt{T}$, then $e^{-rT} H[AoN(K,T), -\alpha]$ is equal to the Black-Scholes price.

(v) **Cash-or-nothing**

Consider a cash-or-nothing call option with strike K and maturity T, whose payoff is give by

$$CoN(K,T) = \mathbb{1}_{\{S_T \geq K\}}.$$

If we let

$$\begin{aligned} h(z) &= \mathbb{1}_{\{S_T \geq K\}}(z) \\ &= \begin{cases} 0 & \text{for } z \in [0, K] \\ h_2(z) & \text{for } z \in (K, \infty), \end{cases} \end{aligned}$$

where $h_2(z) = 1$ is continuous and non-decreasing. Hence, by Corollary 5.4.6 we have

$$\begin{aligned} H[CoN(K,T), -\alpha] &= \mathbb{E}[h_2(z-\alpha)] \\ &= \int_{z_*}^{\infty} f_Z(z) dz \\ &= \Phi[-z_*], \end{aligned}$$

where z_* is as in (iv). Then, letting α be as above and discounting, yields the Black-Scholes price.

The next example is a practical application of using the Wang transform to price exotic options.

Example 5.4.10. This example is purely an illustration of the use of the Wang transform compared to the Black-Scholes price and the parameters were chosen arbitrarily. To apply the methods described in the paper to real world data, one would need to look at for example stochastic interest rates and volatilities, which is left for further research.

As in Hamada and Sherris [76], we implemented Wang's approach using simulation. We simulated lognormal security prices and use these to estimate the relevant α. Consider a Margrabe option to exchange $S^{(1)}$ for $S^{(2)}$ in half a years time. We first show that the Wang price for a Margrabe option with the above-derived value for α converges to the Black-Scholes price. The following parameters were used:

- $r = 8\%$,
- $S_0^{(1)} = 20$, $\mu_1 = 16\%$ and $\sigma_1 = 20\%$,
- $S_0^{(2)} = 25$, $\mu_2 = 10\%$ and $\sigma_2 = 15\%$, and
- $\rho = 0.2$.

Figure 5.1 shows the convergence of both the Wang price to the Black-Scholes price and of α to the true value of 0.2407. The convergence is relatively quick, i.e. not many paths are required to get a fairly accurate price.

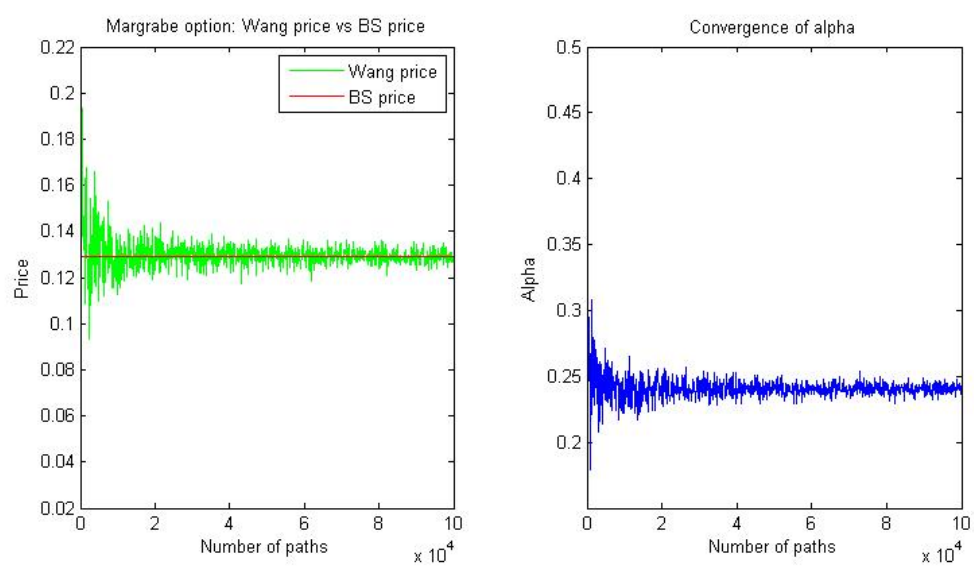

Figure 5.1: Convergence of Wang price and of alpha

One of the disadvantages of the Black-Scholes model is the lack of flexibility in the derived prices. Thus, we would like to see what prices the Wang transform could yield. To do this, we compare the Wang price to the Black Scholes price for varying α in the case of the above-described Margrabe option. This is shown in Figure 5.2.

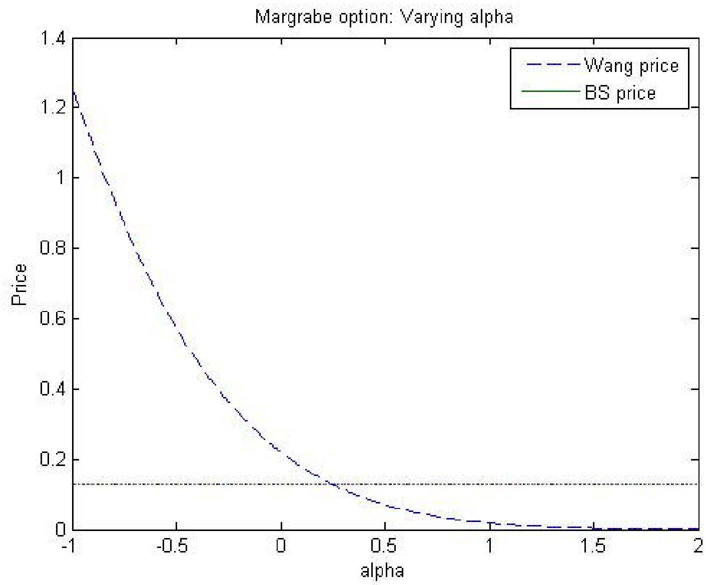

Figure 5.2: Wang price vs alpha

As can be seen in Figure 5.2, it is possible to obtain a wide variety of prices using the Wang transform, while the Black Scholes price remains constant. It should be noted that some of the prices obtained using the Wang transform could lead to arbitrage opportunities (see Pelsser [130]).

We obtained similar results for the geometric basket option and the asset-or-nothing option [117].

5.5 The Esscher-Girsanov transform vs. the Wang transform

The price of a financial security that satisfies the condition of no arbitrage, is found by using a risk-neutral probability measure, which is generally a transformation of the real-world probability measure. This change of measure is referred to as 'risk-neutralising' the statistical distribution. In complete markets, there is a unique risk-neutral probability measure and thus a unique price. In incomplete markets, however, there can be infinitely many risk-neutral probability measures and as of yet no consensus on how to choose one.

The use of exponential tilting to risk-neutralise the real-world distribution has been considered by numerous authors including Bühlmann [21], Gerber and Shiu [67], Wang [159, 161, 162] and Goovaerts and Laeven [70]. Bühlmann [21] developed an economic premium principle based on exponential tilting. A special case of exponential tilting and of Bühlmann's economic premium principle is the Esscher transform, named after the Swedish actuary Fredrik Esscher [59]. Gerber and Shiu [67] used the Esscher transform to price options whose underlying follows a Lévy process. Goovaerts and Laeven [70] introduced a variation of the Esscher transform, which they called the Esscher-Girsanov transform.

Independently, whilst trying to find a unifying approach for pricing insurance and financial risk, Wang introduced a distortion function, which is called the Wang transform.

Kijima and Muromachi pointed out in [97] that the Esscher-Girsanov transform and the Wang transform are technically the same, claiming that Wang [161] contains a proof thereof.

This section explicitly gives the proof of that fact and shows that the Esscher-Girsanov price is equivalent to Wang's risk adjusted premium. This section is based on Labuschagne and Offwood [113].

5.5.1 The Esscher transform

The Esscher transform was developed to approximate the aggregate claim amount distribution around a point of interest x_0, by applying an analytic approximation, the Edgeworth series, to the transformed distribution with the parameter θ chosen such that the new mean is equal to x_0. The reason the transform was required, is the fact that the Edgeworth approximation performs well in the vicinity of the mean, but not so well in the tails.

The Esscher transform was originally a transformation of distribution functions.

Definition 5.5.1. Given a cumulative distribution function $F(x)$ and a parameter θ, the *Esscher transform* is defined by

$$dF^\theta(x) = \frac{e^{\theta x} dF(x)}{\int e^{\theta y} dF(y)}.$$

If $F(x)$ admits a density $f(x)$, then $F^\theta(x)$ has the density

$$f^\theta(x) = \frac{e^{\theta x} f(x)}{\int e^{\theta y} f(y) dy}.$$

In the statistical literature this is a special case of *exponential tilting*.

Definition 5.5.2. The *exponential tilting* of X with respect to Y is given by

$$f_X^*(x) = \frac{\mathbb{E}[e^{\lambda Y}|X = x]}{\mathbb{E}[e^{\lambda Y}]} f_X(x).$$

Note that F_X and its Esscher transform F_X^θ have the same null sets and are thus *equivalent distributions*. For a normal cdf with mean μ and variance σ^2, its Esscher transform is a normal cdf with mean $\mu + \theta\sigma^2$ and variance σ^2.

The real-valued function, given by

$$\pi_X^{Ess}(\theta) = \int_{-\infty}^{\infty} x dF_X^\theta = \frac{\mathbb{E}[Xe^{\theta X}]}{\mathbb{E}[e^{\theta X}]},$$

is known as the *Esscher premium* or the *Esscher price* with parameter θ.

The Esscher transform for probability measures is defined analogously.

Definition 5.5.3. Given a probability space $(\Omega, \mathcal{F}, P)$, a random variable X and a parameter θ, the Esscher transform, sometimes called the Esscher measure, is defined by

$$dP^\theta = \frac{e^{\theta X} dP}{\mathbb{E}[e^{\theta X}]},$$

provided the expectation exists.

The Esscher transform has a sound economic meaning. Bühlmann [21] argues that in the real world, premiums do not only depend on the risk to be covered but also on the surrounding market conditions. Thus, he considered a risk-exchange model, where all individual agents act to maximise their own expected utility.

Consider risk exchanges among agents $j = 1, 2, \ldots, n$. Each investor is characterised by his utility function $u_j(x)$, where $u_j'(x) > 0$ and $u_j''(x) \leq 0$, and his initial wealth W_j. The investor faces the risk of a potential loss of $X_j(\omega)$ and buys a risk-exchange $Y_j(\omega)$. While the original risk X_j belongs to the investor j, the risk-exchange Y_j can be bought/sold in the market. Bühlmann pointed out in [21] that the pricing density $\eta(\omega)$, defined by

$$\text{Price}(Y_j) = \int_\omega Y_j(\omega)\eta(\omega)dP(\omega),$$

can be seen as a distortion of the real-world probabilities.

Definition 5.5.4. The pair $\{Y_{e,j}, \eta_e\}$ are in *equilibrium* if

(i) for all j, $\mathbb{E}[u_j(W_j - X_j + Y_{e,j} - \text{Price}(Y_j))]$ is a maximum among all possible choices of the exchange variables Y_j, and

(ii) $\sum_{j=1}^{n} Y_{e,j}(\omega) = 0$ for all $\omega \in \Omega$.

In the equilibrium, $Y_{e,j}$ is called the *equilibrium exchange* and η_e the *equilibrium price density.*

Bühlmann [21] then proved the following theorem.

Theorem 5.5.5. *Assume that each investor has an exponential utility function*

$$u_j(x) = 1 - e^{-\theta_j x}.$$

Then the equilibrium price density satisfies

$$\eta_e(\omega) = \frac{e^{\theta Z(\omega)}}{\mathbb{E}[e^{\theta Z}]},$$

where

$$Z = \sum_{j=1}^{n} X_j(\omega)$$

is the aggregate risk and θ satisfies

$$\theta^{-1} = \sum_{j=1}^{n} \theta_j^{-1}.$$

The parameter θ_j can be seen as the *risk aversion index* of the jth agent. Using Theorem 5.5.5, the equilibrium price for any risk X is given by

$$\pi_X^{Buhl}(\theta) = \frac{\mathbb{E}[Xe^{\theta Z}]}{\mathbb{E}[e^{\theta Z}]}.$$

Note that if Z is replaced by X in the above, this results in the Esscher transform, i.e. the Esscher transform is a special case of Bühlmann's economic principle.

5.5.2 The Esscher-Girsanov transform

Instead of considering the random variable X, Goovaerts and Laeven [70] consider the extended real-valued function $\Phi^{-1}(F_X(x))$, where Φ^{-1} denotes the inverse distribution function of the standard normal distribution. It is known that if F_X is continuous, then $\Phi^{-1}(F_X(x))$ is normally distributed with mean 0 and variance 1.

From here onwards we will assume that each of the random variables we are working with has a continuous cdf.

Definition 5.5.6. Given a cdf $F(x)$ and parameters $h, v \in \mathbb{R}$, the *Esscher-Girsanov transform* is defined by

$$dF_X^{(h,v)}(x) = \frac{e^{hv\Phi^{-1}(F_X(x))}}{\mathbb{E}[e^{hv\Phi^{-1}(F_X(X))}]} dF_X(x) = e^{hv\Phi^{-1}(F_X(x)) - \frac{1}{2}h^2v^2} dF_X(x). \tag{5.11}$$

The parameter h can be interpreted as the absolute risk aversion and v as the penalty parameter. Goovaerts and Laeven attached the name of Girsanov to the probability measure transform defined above, to emphasize the close resemblance between the Radon-Nikodým derivative used in (5.11) and that used in Girsanov's theorem. It is easy to verify that for a normal cdf with mean μ and variance σ^2, its Esscher-Girsanov transformation is normal with mean $\mu + hv\sigma$ and variance σ^2. If we let $v = \sigma$, then the Esscher-Girsanov transform reduces to the Esscher transform.

Goovaerts and Laeven [70] axiomatically characterise a pricing mechanism involving the Esscher-Girsanov transform. Their pricing mechanism can generate arbitrage-free prices for financial derivatives with an underlying asset driven by a general diffusion process. Similar to the paper by Goovaerts et al. [69], the pricing mechanism of Goovaerts and Laeven [70] allows for a mixture function, weighting the different values of h. We assume the mixture function to be degenerate.

Definition 5.5.7. The *Esscher-Girsanov price* of the random variable X, with parameters $h \leq 0$ and $v > 0$, is given by

$$\pi_X^v(h) = \mathbb{E}[Xe^{hv\Phi^{-1}(F_X(x)) - \frac{1}{2}h^2v^2}].$$

5.5.3 The Esscher-Girsanov transform vs. the Wang transform

Goovaerts and Laeven [70], Kijima and Muromachi [97] and Wang [162] pointed out that a connection exists between the Esscher-Girsanov transform and the Wang transform. In this section, we prove this fact.

Proposition 5.5.8. *Let $\alpha = hv$ and let X be a random variable with pdf $f_X(x)$. The pdf generated by the Wang transform with parameter $-\alpha$ is equal to the pdf generated by the Esscher-Girsanov transform with parameters h and v.*

Proof. By Proposition 5.4.4, the pdf generated by the Wang transform with parameter $-\alpha$

is given by

$$f_X^{WT}(x) = e^{\alpha\Phi^{-1}(F_X(x))-\frac{1}{2}\alpha^2} f_X(x) = e^{hv\Phi^{-1}(F_X(x))-\frac{1}{2}h^2v^2} f_X(x),$$

which by Definition 5.5.6 is the Esscher-Girsanov transformed pdf of X. □

Alternatively, we can also show that the cdf of the Esscher-Girsanov transform is equal to the Wang transform, as is done in the next proposition.

Proposition 5.5.9. *Let $\alpha = hv$ and let X be a random variable with cdf $F_X(x)$. The cdf generated by the Esscher-Girsanov transform with parameters h and v is equal to the cdf generated by the Wang transform with parameter $-\alpha$.*

Proof. The Esscher-Girsanov transform is given by

$$dF_X^* = e^{hv\Phi^{-1}(F_X(x))-\frac{1}{2}h^2v^2} dF_X.$$

Thus,

$$F_X^*(x) = \int_{-\infty}^{x} e^{hv\Phi^{-1}(F_X(t))-\frac{1}{2}h^2v^2} dF_X(t).$$

Let $y = \Phi^{-1}(F_X(t))$, then $dy = \sqrt{2\pi}e^{\frac{1}{2}(\Phi^{-1}(F_X(t)))^2} dF_X$ and substituting this into the integral results in

$$\begin{aligned}
F_X^*(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}h^2v^2} \int_{-\infty}^{\Phi^{-1}(F_X(x))} e^{hvy} \cdot e^{-\frac{1}{2}y^2} dy \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}h^2v^2} \int_{-\infty}^{\Phi^{-1}(F_X(x))} e^{-\frac{1}{2}(y-hv)^2+\frac{1}{2}h^2v^2} dy \\
&= \int_{-\infty}^{\Phi^{-1}(F_X(x))} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-hv)^2} dy \\
&= \Phi(\Phi^{-1}(F_X(x)) - hv).
\end{aligned}$$

This is equal to the Wang Transform with parameter $-hv = -\alpha$. □

Proposition 5.5.10. *Let $hv = \alpha$ and let X be a random variable with pdf f_X and continuous cdf F_X. The Esscher-Girsanov price for a risk X with parameters h and v is equal to Wang's risk adjusted premium for X with parameter α.*

Proof. Wang's risk adjusted premium for X with parameter α, using (5.7), is given by

$$\pi_X^{WT}(\alpha) = -\int_{-\infty}^{0} \Phi(\Phi^{-1}(F_X(x)) - \alpha)dx + \int_{0}^{\infty} [1 - \Phi(\Phi^{-1}(F_X(x)) - \alpha)]\, dx.$$

Let $c(x) = \Phi^{-1}(F_X(x)) - \alpha$, then

$$\pi_X^{WT}(\alpha) = -\int_{-\infty}^{0}\int_{-\infty}^{c(x)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt\, dx + \int_{0}^{\infty}\int_{c(x)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt\, dx.$$

Next, we need to change the order of integration. Letting $k(t) = F_X^{-1}(\Phi(t+\alpha))$ and $\gamma = \Phi^{-1}(F_X(0)) - \alpha$, we get

$$\begin{aligned}
\pi_X^{WT}(\alpha) &= -\int_{-\infty}^{\gamma}\int_{k(t)}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\, dx\, dt + \int_{\gamma}^{\infty}\int_{0}^{k(t)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\, dx\, dt \\
&= -\int_{-\infty}^{\gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} (0 - k(t))\, dt + \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} (k(t) - 0)\, dt \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} k(t)\, dt \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} F_X^{-1}(\Phi(t+\alpha))\, dt.
\end{aligned}$$

Substituting $y = t + \alpha$, we get

$$\pi_X^{WT}(\alpha) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\alpha)^2} F_X^{-1}(\Phi(y)))\, dy. \tag{5.12}$$

On the other hand, the Esscher-Girsanov price for X is given by

$$\begin{aligned}
\pi_X^{v}(h) &= \mathbb{E}[X e^{hv\Phi^{-1}(F_X(x)) - \frac{1}{2}h^2v^2}] \\
&= \int_{-\infty}^{\infty} x e^{hv\Phi^{-1}(F_X(x)) - \frac{1}{2}h^2v^2} f_X(x) dx.
\end{aligned} \tag{5.13}$$

Let $y = \Phi^{-1}(F_X(x))$, then $dy = \sqrt{2\pi} e^{\frac{1}{2}y^2} dF_X$. Substituting this into (5.13), we get

$$\pi_X^{v}(h) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-hv)^2} F_X^{-1}(\Phi(y)) dy. \tag{5.14}$$

As (5.12) is the same as (5.14), we have that $\pi_X^{WT}(\alpha) = \pi_X^{v}(h)$. □

Note that the relationship between the parameters is different in Proposition 5.5.10 compared to Propositions 5.5.8 and 5.5.9. This is due to the way the risk adjusted premium is defined. Traditionally, one applies the distortion function to the tail probabilities (decumulative distribution function) rather than the regular probabilities (cumulative distribution function). We followed this tradition, however, the price we paid is that the sign of the parameter α in the propositions does not correspond. If the risk adjusted premium is defined in the following way

$$\begin{aligned}
\pi_X^{WT}(\alpha) &= -\int_{-\infty}^{0} g_\alpha(F_X(x))\, dx + \int_{0}^{\infty} (1 - g_\alpha(F_X(x)))\, dx \\
&= \int_{-\infty}^{0} (g_{-\alpha}(S_X(x)) - 1)\, dx + \int_{0}^{\infty} g_{-\alpha}(S_X(x))\, dx,
\end{aligned}$$

then Proposition 5.5.10 would be stated as follows and the proof follows similarly.

Proposition 5.5.11. *Let $hv = \alpha$ and let X be a random variable with pdf f_X and continuous cdf F_X. The Esscher-Girsanov price for a risk X with parameters h and v is equal to Wang's risk adjusted premium for X with parameter $-\alpha$.*

5.5.4 Discussion

In the previous section, we showed that in a static setting, the Esscher-Girsanov transform and the Wang transform coincide. However, the two transforms are not equivalent. The Esscher-Girsanov transform is a two-parameter transform and can therefore never be fully equivalent to the one-parameter Wang transform. This becomes apparent in the dynamic setting, where the two parameters in the Esscher-Girsanov transform start to play a distinct role. The two-parameter Esscher-Girsanov transform can generate arbitrage-free prices for financial derivatives governed by general diffusion processes as shown by Goovaerts and Laeven [70] and emphasised by Badescu et al. in [10]. This, however, is not true for the one-parameter Wang transform, as was explicitly shown by Pelsser [130]. It proves that the Esscher-Girsanov transform is not equivalent to the Wang transform in a dynamic setting.

5.6 Comonotonic convex risk measures

Let $\alpha : \mathcal{M}_{1,m} \to \mathbb{R} \cup \{\infty\}$ be any functional such that $\inf_{\mu \in \mathcal{M}_{1,m}} \alpha(\mu)$ is finite. If we define $\rho : \mathcal{X} \to \mathbb{R}$ by

$$\rho(X) = \sup_{\mu \in \mathcal{M}_{1,m}} (\mu(X) - \alpha(\mu)),$$

then ρ is monotonic, +-translation invariant and comonotonic convex. The function ρ defined above is called a *comonotonic convex risk measure.*

The following theorem proves the representation theorem for comonotonic convex risk measures in terms of Choquet integrals.

Theorem 5.6.1. *If $\rho : L_\infty(P) \to \mathbb{R}$ satisfies monotonicity, +-translation invariance and comonotonic convexity, then ρ is of the form*

$$\rho(X) = \max_{\mu \in \mathcal{M}_{1,m}} (\mu(X) - \alpha(\mu)) \quad \text{for } X \in L_\infty(P), \tag{5.15}$$

where

$$\alpha(\mu) = \sup_{\rho(X) \leq 0} \mu(X) \quad \text{for } \mu \in \mathcal{M}_{1,m}.$$

Proof. We may assume without loss of generality that $\rho(0) = 0$. For any $X \in L_\infty(P)$, let $X_1 = X - \rho(X)$, then $\rho(X_1) = \rho(X - \rho(X)) = 0$ by +-translation invariance. Hence,

$$\alpha(\mu) \geq \mu(X_1) = \mu(X) - \rho(X)$$

for all $\mu \in \mathcal{M}_{1,m}$. Thus, for $X \in L_\infty(P)$, we have

$$\rho(X) \geq \sup_{\mu \in \mathcal{M}_{1,m}} (\mu(X) - \alpha(\mu)).$$

For a given X, we will now construct some $\mu_X \in \mathcal{M}_{1,m}$ such that

$$\rho(X) \leq \mu_X(X) - \alpha(\mu_X), \tag{5.16}$$

which, in view of the previous step, will prove our representation (5.15). If $\rho(X) \neq 0$, then there exists $\epsilon \in \mathbb{R}$ such that $\rho(X) + \epsilon = 0$. By the +-translation invariance of ρ, we have $\rho(X + \epsilon) = 0$. Then letting $Y = X + \epsilon$, we have $\rho(Y) = 0$. Thus, it suffices to prove (5.16) for $\rho(X) = 0$.

The subset $\mathcal{B}$ of $\mathcal{A}_\rho$, defined by $\mathcal{B} = \{Y \in L_\infty(P) : \rho(Y) < 0\}$, is not convex. However, using the following notation

$$Y \sim Z \quad \Leftrightarrow \quad Y, Z \text{ comonotonic},$$

we define the set $[X] = \{Z \in \mathcal{B} : Z \sim X\}$, i.e. $Z \in [X]$ if and only if there exists non-decreasing functions u and v such that $X = u(W)$ and $Z = v(W)$ for some W. This set is, by definition, convex.

As $\rho(X) = 0$, we have that $X \notin \mathcal{B}$ and therefore $X \notin [X]$. By the Hahn-Banach theorem, there exists a non-trivial (linear) $\theta_X \in (L_\infty(P))^*$, such that

$$\sup_{Y \in [X]} \theta_X(Y) \leq \theta_X(X).$$

Define $\rho^* : L_\infty(P) \to \mathbb{R}$ by

$$\rho^*(Y) = \sup\{\theta_X(Z) : Z \leq Y, Z \sim X\}.$$

Then ρ^* is monotonic increasing and positive homogeneous.

Next we show that ρ^* is comonotonic additive. Let $Y_1, Y_2 \in L_\infty(P)$ be comonotonic. Then there exists continuous and increasing functions u, v on $\mathbb{R}$ such that $u(z) + v(z) = z$ for $z \in \mathbb{R}$ and

$$Y_1 = u(Y_1 + Y_2) \text{ and } Y_2 = v(Y_1 + Y_2).$$

For any $Z \sim X$ with $Z \leq Y_1 + Y_2$, we have $u(Z) \leq u(Y_1 + Y_2) = Y_1$ and $v(Z) \leq v(Y_1 + Y_2) = Y_2$. Since $Z \sim X$, there exists a random variable W and non-decreasing functions h and g such that $Z = h(W)$ and $X = g(W)$. Therefore, we have that $u(Z) = u(h(W))$ and $v(Z) = v(h(W))$, which implies $u(Z) \sim X$ and $v(Z) \sim X$. Hence,

$$\rho^*(Y_1) + \rho^*(Y_2) \geq \theta_X(u(Z)) + \theta_X(v(Z)) = \theta_X(u(Z) + v(Z)) = \theta_X(Z).$$

We have shown that $\rho^*(Y_1) + \rho^*(Y_2) \geq \theta_X(Z)$ for all $Z \sim X$ with $Z \leq Y_1 + Y_2$, thus

$$\rho^*(Y_1) + \rho^*(Y_2) \geq \sup\{\theta_X(Z) : Z \leq Y_1 + Y_2, Z \sim X\} = \rho^*(Y_1 + Y_2).$$

Conversely, let $Z_1 \sim X$ and $Z_2 \sim X$ such that $Z_1 \leq Y_1$ and $Z_2 \leq Y_2$. Define $Z = Z_1 + Z_2$. Then, by Lemma 5.1.8, $Z_1 + Z_2 \sim X$ and $Z \leq Y_1 + Y_2$. By the definition of ρ^*, we have

$$\rho^*(Y_1 + Y_2) \geq \theta_X(Z) = \theta_X(Z_1) + \theta_X(Z_2).$$

Taking the supremum over all $Z_1 \sim Y_1$ such that $Z_1 \leq Y_1$ and all $Z_2 \sim Y_2$ such that $Z_2 \leq Y_2$, we get

$$\rho^*(Y_1 + Y_2) \geq \rho^*(Y_1) + \rho^*(Y_2),$$

proving that ρ^* is comonotonic additive.

Note that $\rho^*(X) = \theta_X(X)$.

We also have that $L_\infty(P)$ satisfies all the conditions required in Theorem 5.1.12. Therefore, by Greco's representation theorem for bounded functions, Theorem 5.1.12, there exists $\mu_X \in \mathcal{M}_{1,m}$ representing ρ^*.

If $\rho(Y) \leq 0$, then $\rho(Y - \epsilon) < 0$ for any $\epsilon > 0$. This, along with the fact that $\rho(X) = 0$, implies that $Y - \epsilon < Y \leq X$ and hence, $\rho^*(Y - \epsilon) \leq \rho^*(X) = \theta_X(X)$.

Thus, we have that

$$\mu_X(Y) - \epsilon = \mu_X(Y - \epsilon) = \rho^*(Y - \epsilon) \leq \theta_X(X).$$

Finally, we have

$$\alpha(\mu_X) = \sup_{\rho(Y) \leq 0} \mu_X(Y) \leq \theta_X(X),$$

$$\mu_X(X) - \alpha(\mu_X) \geq \rho^*(X) - \theta_X(X) = \theta_X(X) - \theta_X(X) = 0 = \rho(X).$$

□

5.7 The Wang transform as a risk measure

Recently, risk measures based on distortion probabilities have been developed in actuarial science and applied to insurance rate making. As mentioned at the end of Chapter 4, the Wang transform can be used to measure risk. Assume that X is a loss random variable (i.e. $X < 0$ for a gain and $X > 0$ for a loss) and has distribution function $F_X(x)$.

Definition 5.7.1. The family of *distortion risk measures* is defined as the mean under the distortion probability $F_X^*(x) = g(F_X(x))$, where g is a distortion function. In other words,

$$\rho(X) = \mathbb{E}^*[X] := -\int_{-\infty}^{0} g(F_X(x))\,dx + \int_0^{\infty} (1 - g(F_X(x)))\,dx.$$

Note that the distortion risk measure is just a Choquet integral.

Proposition 5.7.2. *A distortion risk measure is coherent if and only if the associated distortion function is concave.*

For the proof of this proposition see the properties of the Choquet integral in Section 5.1.

Distortion risk measures can be seen as a generalisation of some of the known risk measures. VaR_α corresponds to the distortion

$$g(u) = \begin{cases} 0 & \text{for } u < \alpha \\ 1 & \text{for } u \geq \alpha. \end{cases}$$

This function is not necessarily concave, and hence is not coherent.

TVaR corresponds to the distortion

$$g(u) = \begin{cases} 0 & \text{for } u < \alpha \\ \frac{u-\alpha}{1-\alpha} & \text{for } u \geq \alpha, \end{cases}$$

which is continuous everywhere but not differentiable at $u = \alpha$.

ES, however, cannot be written as a distortion risk measure.

Wang [159] suggested using the Wang transform as a risk measure as follows.

Definition 5.7.3. For a random variable x with distribution F_x, define a risk measure for capital requirement as follows:

1. For a preselected security level α, let $\lambda = \Phi^{-1}(\alpha)$.

2. Apply the Wang transform: $F_X^*(x) = \Phi[\Phi^{-1}(F_X(x)) - \lambda]$.

3. Set the capital requirement to be the expected value under F_X^*: $WT_\alpha(X) = \mathbb{E}^*[X]$.

Example 5.7.4. Figure 5.7.4 shows an example of how the Wang transform measure differs to VaR, TVaR and ES.

Portfolio A

x	f(x)	F(x)	F*(x)
0	0.6	1	0.08204
-1	0.375	0.4	0.54163
-5	0.025	0.025	0.37634

Portfolio B

x	f(x)	F(x)	F*(x)
0	0.6	1	0.08204
-1	0.39	0.4	0.67018
-11	0.01	0.01	0.24778

Portfolio	$VaR_{0.95}$	$TVaR_{0.95}$	$ES_{0.95}$	$WT_{0.95}$
A	-1	-1.25	-3.00	-2.42
B	-1	-1.25	-3.00	-3.40

Figure 5.3: Wang transform vs. other risk measures.

Chapter 6

Vector-valued Orlicz spaces

The importance of Orlicz spaces in the study of mathematics of finance came to the fore in the 2000's when Frittelli and his collaborators connected the theory of utility functions to Orlicz spaces (see [15, 16, 17, 65]). This was explained in Chapter 3. Orlicz spaces now also play an important role in the theory of risk measures (see [16, 26]), as was already discussed in Chapter 4 and will be expanded upon in the next chapter. In some applications of mathematics of finance (as in the case of systemic risk), the theory of the real-valued case has to be extended to the multi-valued case (see [91]). In particular, there have been recent developements in the theory of set-valued risk measures. A neat way to work with these set-valued risk measures is, in our opinion, via tensor products. Thus, to define a set-valued risk measure on a vector-valued Orlicz space, we need to describe these Orlicz spaces as suitable tensor products.

The aim of this chapter is to give descriptions of Banach space-valued Orlicz spaces and of their duals. We use the former to derive the latter.

The latter is motivated by the important role that the dual of the real-valued Orlicz heart plays in utility maximisation problems and in the risk measure representation theorems, as in [15, 16, 17, 26, 65]. We use the above mentioned descriptions to characterise martingale convergence in Banach space-valued Orlicz spaces and also to describe the Radon-Nikodým property in such spaces.

We refer the reader to Appendix B for the preliminaries on tensor products and to Chapter 2 for Orlicz spaces.

This chapter is organised in the following way. In Section 6.1 we describe the Y-valued Orlicz heart $H_\Phi(P,Y)$; more precisely, we show that the Y-valued Orlicz heart $H_\Phi(P,Y)$ is isometrically isomorphic to the l-completed tensor product $H_\Phi(P)\widetilde{\otimes}_l Y$ of the scalar-valued

Orlicz heart $H_\Phi(P)$ and Y.

The main result of Section 6.2 is the characterisation of the equality of $(H_\Phi(P)\widetilde{\otimes}_l Y)^*$ and $(H_\Phi(P))^*\widetilde{\otimes}_l Y^*$ in terms of the Radon-Nikodým property on Y^*.

We show that the l-norm is associative. As an application thereof, we give an alternative proof of a result noted by Popa, which states that for any separable Banach lattice E and any Banach space Y, E^* and Y have the Radon-Nikodým property if and only if $E^*\widetilde{\otimes}_l Y$ has the Radon-Nikodým property. Via a deep result of Talagrand, E^* may be replaced by E in the above mentioned result. This, together with the results of Section 6.1, enables us to describe the Radon-Nikodým property in $H_\Phi(P,Y)$ in terms of the Radon-Nikodým property on $H_\Phi(P)$ and Y. The latter extends a result of Sundaresan (see [151]) and of Turret and Uhl (see [154]).

Section 6.5 deals with the convergence of norm bounded martingales in $H_\Phi(P,Y)$, characterised in terms of the Radon-Nikodým property on Y. Note that the contents of this chapter are new and are based on [114].

6.1 Connecting $H_\Phi(P,Y)$ to $H_\Phi(P)\widetilde{\otimes}_l Y$

Let $(\Omega,\mathcal{F},P)$ be a probability space, Φ a finite Young function and Y a Banach space. The purpose of this section is to show that $H_\Phi(P,Y)$ is isometrically isomorphic to $H_\Phi(P)\widetilde{\otimes}_l Y$.

First we recap the following definitions. A function $s:\Omega\to Y$ is *simple* if there exist $y_1,y_2,\ldots,y_n\in Y$ and sets $A_1,A_2,\ldots,A_n\in\mathcal{F}$ such that $s=\sum_{i=1}^n y_i\chi_{A_i}$. Here, χ_{A_i} denotes the characteristic function of A_i, given by $\chi_{A_i}(\omega)=1$ when $\omega\in A_i$ and $\chi_{A_i}(\omega)=0$ when $\omega\notin A_i$. A function $f:\Omega\to Y$ is called *P-measurable* if there exists a sequence of simple functions (s_n) with $\lim_{n\to\infty}\|s_n-f\|=0$ P-a.s..

Moreover, $H_\Phi(P,Y)$ is $\mathcal{N}_\Phi$-closed in $L_\Phi(P,Y)$. If Φ is finite, then the set $S(P,Y)$ of step-functions, defined by

$$\left\{\sum_{i=1}^n y_i\chi_{A_i} : \chi_{A_i} \text{ is } P\text{-integrable}, y_i\in Y, n\in\mathbb{N}\right\},$$

is $\|\cdot\|_\Phi$-dense in $H_\Phi(P,Y)$.

If $Y=\mathbb{R}$, then we write $L_\Phi(P)=L_\Phi(P,\mathbb{R})$ and $H_\Phi(P)=H_\Phi(P,\mathbb{R})$.

Define γ by

$$\gamma(f,y)=f(t)y \text{ for all } (f,y)\in H_\Phi(P)\times Y \text{ and } t\in\Omega.$$

Then $\gamma: H_\Phi(P)\times Y\to H_\Phi(P,Y)$ is a bilinear map. It is well-known that the unique linear map $\kappa : H_\Phi(P) \otimes_l Y \to H_\Phi(P, Y)$ for which $\kappa \circ \otimes = \gamma$, is given by

$$\left(\kappa\left(\sum_{i=1}^n f_i \otimes y_i\right)\right)(t) = \sum_{i=1}^n f_i(t) y_i \text{ for all } t \in \Omega.$$

Regarding the continuity of κ, we have the following result.

Lemma 6.1.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, Φ a finite Young function and Y a Banach space. The canonical map $\kappa : H_\Phi(P) \otimes_l Y \to H_\Phi(P, Y)$, defined by*

$$\left(\kappa\left(\sum_{i=1}^n f_i \otimes y_i\right)\right)(t) = \sum_{i=1}^n f_i(t) y_i \text{ for all } t \in \Omega,$$

is a continuous linear map with $\|\kappa\| \leq 1$.

Proof. We identify $H_\Phi(P)\otimes Y$ with its image $\kappa(H_\Phi(P)\otimes Y)$ in $H_\Phi(P, Y)$. The norm induced on (the image of) $H_\Phi(P) \otimes Y$ is given by

$$\mathcal{N}_\Phi(u) = \inf\left\{a > 0 : \int_\Omega \Phi\left(\frac{1}{a}\left\|\sum_{i=1}^n f_i(t) y_i\right\|\right) dP(t) \leq 1\right\}$$

for all $u = \sum_{i=1}^n f_i \otimes y_i \in H_\Phi(P) \otimes Y$ and is independent of the representation of u.

We show that $\mathcal{N}_\Phi(u) \leq \|u\|_l$ for all $u \in H_\Phi(P) \otimes Y$. Let $u = \sum_{i=1}^n f_i \otimes y_i \in H_\Phi(P) \otimes Y$. Then, for any $a > 0$,

$$\frac{1}{a}\left\|\sum_{i=1}^n f_i(t) y_i\right\| \leq \frac{1}{a}\sum_{i=1}^n \|y_i\|\, |f_i|(t),$$

and since Φ is non-decreasing,

$$\Phi\left(\frac{1}{a}\left\|\sum_{i=1}^n f_i(t) y_i\right\|\right) \leq \Phi\left(\frac{1}{a}\sum_{i=1}^n \|y_i\|\, |f_i|(t)\right).$$

Hence,

$$\begin{aligned}
\mathcal{N}_\Phi(u) &= \inf\left\{a > 0 : \int_\Omega \Phi\left(\frac{1}{a}\left\|\sum_{i=1}^n f_i(t) y_i\right\|\right) dP(t) \leq 1\right\} \\
&\leq \inf\left\{a > 0 : \int_\Omega \Phi\left(\frac{1}{a}\sum_{i=1}^n \|y_i\|\, |f_i|(t)\right) dP(t) \leq 1\right\} \\
&= \mathcal{N}_\Phi\left(\sum_{i=1}^n \|y_i\|\, |f_i|\right),
\end{aligned}$$

which implies that

$$\mathcal{N}_\Phi(u) \leq \inf\left\{\mathcal{N}_\Phi\left(\sum_{i=1}^n \|y_i\|\, |f_i|\right) : u = \sum_{i=1}^n f_i \otimes y_i\right\} = \|u\|_l.$$

□

It is well-known that the set $S(P)$ of step-functions, defined by

$$\left\{\sum_{i=1}^{n} \lambda_i \chi_{A_i} : \chi_{A_i} \text{ is } P\text{-integrable}, \lambda_i \in \mathbb{R}, n \in \mathbb{N}\right\},$$

is dense in $H_\Phi(P)$ (see [54]).

Lemma 6.1.2. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, Φ a finite Young function and Y a Banach space. Then*

$$S(P) \otimes Y := \left\{\sum_{i=1}^{n} \chi_{A_i} \otimes y_i : \chi_{A_i} \textit{ is } P\textit{-integrable}, y_i \in Y, n \in \mathbb{N}\right\}$$

is dense in $H_\Phi(P)\widetilde{\otimes}_l Y$.

Proof. Suppose that $\sum_{i=1}^{n} f_i \otimes y_i \in H_\Phi(P) \otimes Y$. Let $\epsilon > 0$ be given. As $S(P)$ is dense in $H_\Phi(P)$, there exist $s_1, s_2, \ldots, s_n \in S(P)$ such that

$$\mathcal{N}_\Phi(f_i - s_i) < \frac{\epsilon}{\sum_{i=1}^{n} \|y_i\|} \quad \text{for } i = 1, 2, \ldots, n.$$

Hence,

$$\mathcal{N}_\Phi\left(\sum_{i=1}^{n} \|y_i\| f_i - \sum_{i=1}^{n} \|y_i\| s_i\right) < \epsilon,$$

from which we get that

$$\left\|\sum_{i=1}^{n} f_i \otimes y_i - \sum_{i=1}^{n} s_i \otimes y_i\right\|_l < \epsilon.$$

But, for each i, we have that $s_i = \sum_{j=1}^{k} \lambda_j^i \chi_{A_j^i}$, where each $\chi_{A_j^i}$ is P-integrable and $\lambda_i \in \mathbb{R}$. Consequently, $S(P) \otimes Y$ is dense in $H_\Phi(P)\widetilde{\otimes}_l Y$. □

The following is the main result of this section.

Theorem 6.1.3. *Let $(\Omega, \mathcal{F}, P)$ denote a probability space, Φ a finite Young function and Y a Banach space. Then the canonical map $\kappa : H_\Phi(P)\widetilde{\otimes}_l Y \to H_\Phi(P, Y)$, defined by*

$$\left(\kappa\left(\sum_{i=1}^{n} f_i \otimes y_i\right)\right)(t) = \sum_{i=1}^{n} f_i(t) y_i \textit{ for all } t \in \Omega,$$

is a surjective isometry.

Proof. We first show that $\|u\|_l \leq \mathcal{N}_\Phi(u)$ for all $u \in S(P) \otimes Y$.

Consider $u = \sum_{i=1}^n \chi_{A_i} \otimes y_i \in S(P) \otimes Y$ for which $\{A_i : 1 \leq i \leq n\}$ is a mutually disjoint set and $\bigcup_{i=1}^n A_i = \Omega$. It follows from

$$\sum_{i=1}^n \|y_i\| \, \chi_{A_i}(t) = \left\| \sum_{i=1}^n \chi_{A_i}(t) y_i \right\|$$

that

$$\begin{aligned}
\mathcal{N}_\Phi(u) &= \inf\left\{a > 0 : \int_\Omega \Phi\left(\frac{1}{a}\left\|\sum_{i=1}^n \chi_{A_j}(t) y_i\right\|\right) dP(t) \leq 1\right\} \\
&= \inf\left\{a > 0 : \int_\Omega \Phi\left(\frac{1}{a}\sum_{i=1}^n \|y_i\| \, \chi_{A_j}(t)\right) dP(t) \leq 1\right\} \\
&= \mathcal{N}_\Phi\left(\sum_{i=1}^n \|y_i\| \, \chi_{A_j}\right) \\
&\geq \inf\left\{\mathcal{N}_\Phi\left(\sum_{i=1}^n \|y_i\| \, \chi_{A_j}\right) : u = \sum_{i=1}^n \chi_i \otimes y_i\right\} \\
&= \|u\|_l.
\end{aligned}$$

By Lemma 6.1.1, we get that $\kappa : S(P) \otimes_l Y \to S(P, Y)$ is an isometry.

By the definition of κ, we also have that $\kappa\left(S(P) \otimes_l Y\right) = S(P, Y)$. As $S(P, Y)$ is dense in $H_\Phi(P, Y)$, $\kappa : S(P) \otimes_l Y \to S(P, Y)$ is a surjective isometry and since $S(P) \otimes Y$ is dense in $H_\Phi(P)\widetilde{\otimes}_l Y$, it follows that κ has an extension (again denoted by) $\kappa : H_\Phi(P)\widetilde{\otimes}_l Y \to H_\Phi(P, Y)$ which is a surjective isometry. □

Theorem 6.1.3 provides a connection between the theory of $H_\Phi(P, Y)$-spaces and the theory of l-tensor products. We exhibit some applications of this connection in later sections and in the next chapter.

6.2 $(H_\Phi(P, Y))^*$

Let Φ be a finite Young function and Y a Banach space. The aim of this section is to describe $(H_\Phi(P, Y))^*$. We first do some preparation.

Let E be a Banach lattice. We recall the following definition from [142, Chapter IV, Section 3].

Definition 6.2.1. A linear map $T : E \to Y$ is called *cone absolutely summing* if for every positive summable sequence (x_n) in E, the sequence (Tx_n) is absolutely summable in Y.

The space

$$\mathcal{L}^{\text{cas}}(E,Y) = \{T : E \to Y : T \text{ is cone absolutely summing}\}$$

is a Banach space with respect to the norm defined by

$$\|T\|_{\text{cas}} = \sup\left\{\sum_{i=1}^n \|Tx_i\| : x_1,\dots,x_n \in E_+, \left\|\sum_{i=1}^n x_i\right\| = 1, n \in \mathbb{N}\right\}$$

for all $T \in \mathcal{L}^{\text{cas}}(E,Y)$ (see also [45]).

Cone absolutely summing maps extend the Chaney-Schaefer l-tensor product in the following sense: The canonical map $W_u : E^* \otimes_l Y \to \mathcal{L}^{\text{cas}}(E,Y)$, given by

$$W_u x = \sum_{i=1}^n x_i^*(x) y_i \text{ for all } x \in E,$$

for $u = \sum_{i=1}^n x_i^* \otimes y_i$ is an isometry (see [142, Chapter IV, Section 7] and [23, 90, 111]).

Chaney [23] proved that Y has the Radon-Nikodým property if and only if $L_p(P)\widetilde{\otimes}_l Y = \mathcal{L}^{\text{cas}}(L_q(P),Y)$ for any probability measure space $(\Omega,\mathcal{F},P)$, $1 < p < \infty$ and $\frac{1}{p}+\frac{1}{q}=1$.

The following was shown in [31].

Theorem 6.2.2. *A Banach space Y has the Radon-Nikodým property if and only if $E^*\widetilde{\otimes}_l Y = \mathcal{L}^{\text{cas}}(E,Y)$ for any separable Banach lattice E for which E^* has order continuous norm.*

The following was proved in [111].

Lemma 6.2.3. *If E is a Banach lattice and Y a Banach space, then $u \in E\widetilde{\otimes}_l Y$ if and only if $u = \sum_{i=1}^\infty x_i \otimes y_i$, where*

$$\left\|\sum_{i=1}^\infty |x_i|\right\|_E < \infty \quad \textit{and} \quad \lim_{i\to\infty} \|y_i\|_Y = 0. \tag{6.1}$$

Moreover,

$$\|u\|_l = \inf\left\{\left\|\sum_{i=1}^\infty |x_i|\right\| \sup\|y_i\| : u = \sum_i x_i \otimes y_i, \left\|\sum_{i=1}^\infty |x_i|\right\| < \infty, \lim_{i\to\infty}\|y_i\| = 0\right\}.$$

The following lemma is required to prove the main result of this section.

Let $(\Omega,\mathcal{F},P)$ be a probability space, $1<p<\infty$, $\frac{1}{p}+\frac{1}{q}=1$, and define the function Φ by $\Phi(u) = \frac{u^p}{p}$ for all $u \in [0,\infty)$. Then Φ is a finite Young function and the Orlicz heart associated with Φ is $L_p(P)$, endowed with the norm $p^{-\frac{1}{p}}\|\cdot\|_{L_p(P)}$.

Lemma 6.2.4. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $1 < p < \infty$. Then the Banach lattice $(L_p(P), p^{-\frac{1}{p}}\|\cdot\|_p)$ is reflexive.*

Proof. Let $\frac{1}{p} + \frac{1}{q} = 1$. We claim that $(L_p(P), p^{-\frac{1}{p}}\|\cdot\|_p)^* = (L_q(P), p^{\frac{1}{p}}\|\cdot\|_q)$. Indeed, if $f \in (L_p(P), p^{-\frac{1}{p}}\|\cdot\|_p)^*$, then $f : L_p(P) \to \mathbb{R}$ is a continuous linear function and

$$\begin{aligned}\|f\| &= \sup\{|f(x)| : p^{-\frac{1}{p}}\|x\|_p \leq 1\} \\ &= p^{\frac{1}{p}} \sup\{|f(x)| : \|x\|_p \leq 1\} \\ &= p^{\frac{1}{p}}\|f\|_q,\end{aligned}$$

proving our claim. Consequently,

$$(L_p(P), p^{-\frac{1}{p}}\|\cdot\|_p)^{**} = (L_q(P), p^{\frac{1}{p}}\|\cdot\|_q)^* = (L_p(P), p^{-\frac{1}{p}}\|\cdot\|_p).$$

□

Theorem 6.2.5. *Let Y be a Banach space. Then Y has the Radon-Nikodým property if and only if $(H_\Phi(P))^*\widetilde{\otimes}_l Y = \mathcal{L}^{\text{cas}}(H_\Phi(P), Y)$ for any probability space $(\Omega, \mathcal{F}, P)$, and all finite Young functions Φ for which $H_\Phi(P)$ is separable and $(H_\Phi(P))^*$ has order continuous norm.*

Proof. Suppose Y has the Radon-Nikodým property. Let $(\Omega, \mathcal{F}, P)$ be a probability space and Φ a finite Young function for which $H_\Phi(P)$ is separable and $(H_\Phi(P))^*$ has order continuous norm. It follows from Theorem 6.2.2 that $(H_\Phi(P))^*\widetilde{\otimes}_l Y = \mathcal{L}^{\text{cas}}(H_\Phi(P), Y)$.

Conversely, suppose that $(H_\Phi(P))^*\widetilde{\otimes}_l Y = \mathcal{L}^{\text{cas}}(H_\Phi(P), Y)$ for any probability space $(\Omega, \mathcal{F}, P)$, and all Young functions Φ for which $H_\Phi(P)$ is separable and $(H_\Phi(P))^*$ has order continuous norm. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, and define Φ by $\Phi(u) = \frac{u^p}{p}$ for all $u \in [0, \infty)$. Then Φ is a finite Young function which yields $L_p(P)$, endowed with the norm $p^{-\frac{1}{p}}\|\cdot\|_{L_p(P)}$, as Orlicz heart. Let $T \in \mathcal{L}^{\text{cas}}(L_p(P), Y)$. Then

$$\|T\|_{\text{cas}} = p^{-\frac{1}{p}}\|T\|_{\mathcal{L}^{\text{cas}}((L_p(P), p^{\frac{1}{p}}\|\cdot\|_{L_p(P)}), Y)}. \tag{6.2}$$

By assumption, we have that

$$(L_q(P), p^{-\frac{1}{p}}\|\cdot\|_{L_q(P)})^*\widetilde{\otimes}_l Y = \mathcal{L}^{\text{cas}}((L_p(P), p^{\frac{1}{p}}\|\cdot\|_{L_p(P)}), Y).$$

An application of Lemma 6.2.3 yields $T = \sum_{i=1}^{\infty} x_i \otimes y_i$, where

$$p^{\frac{1}{p}}\left\|\sum_{i=1}^{\infty}|x_i|\right\|_{L_p(P)} = \left\|\sum_{i=1}^{\infty}|x_i|\right\|_{(L_q(P), p^{-\frac{1}{p}}\|\cdot\|_{L_q(P)})^*} < \infty \text{ and } \lim_{i\to\infty}\|y_i\|_Y = 0.$$

By (6.2) and Lemma 6.2.3, we get $T \in L_p(P)\widetilde{\otimes}_l Y$, which completes the proof. □

The following description of $(H_\Phi(P,Y))^*$ is the main result of this section.

Theorem 6.2.6. *Let Y be a Banach space. Then Y^* has the Radon-Nikodým property if and only if $(H_\Phi(P)\widetilde{\otimes}_l Y)^* = (H_\Phi(P))^*\widetilde{\otimes}_l Y^*$ for any probability space $(\Omega,\mathcal{F},P)$, and all finite Young functions Φ for which $H_\Phi(P)$ is separable and $(H_\Phi(P))^*$ has order continuous norm.*

Proof. It is well known that $(H_\Phi(P)\widetilde{\otimes}_l Y)^* = \mathcal{L}^{\mathrm{cas}}(H_\Phi(P),Y^*)$ (see [142]). An application of Theorem 6.2.5 yields the desired result. □

6.3 Associativity of the l-norm

In this section, we consider the associativity of the l-norm. For this purpose, we recall some known facts about the l-norm from [23, 90, 111, 112, 142].

If E and F are Banach lattices, then $E\widetilde{\otimes}_l F$ is a Banach lattice, with positive cone $(E\widetilde{\otimes}_l F)_+$ given by the l-closure of the projective cone

$$E_+\otimes F_+ := \left\{\sum_{i=1}^n x_i\otimes y_i : n\in\mathbb{N},\ x_1,\cdots,x_n\in E_+,\ y_1,\cdots,y_n\in F_+\right\}$$

of E and F,

$$|x\otimes y| = |x|\otimes|y| \text{ for all } x\in E \text{ and } y\in F,$$

and $\|\cdot\|_l$ is a Riesz norm on $E\widetilde{\otimes}_l F$.

If Y is a Banach space, the m-norm on $Y\otimes E$ is given by

$$\|u\|_m = \inf\left\{\left\|\sum_{i=1}^n \|x_i\|\,|y_i|\right\| \ :\ u=\sum_{i=1}^n x_i\otimes y_i\right\}.$$

It is well known that the linear bijection $t: E\otimes_l Y \hookrightarrow Y\otimes_m E$, given by

$$t(x\otimes y) = y\otimes x,$$

is an isometric isomorphism (see [23, 111, 112, 142]). Moreover, it follows from [111, Theorem 3.2] that

$$\|u\|_l = \inf\left\{\left\|\sum_{i=1}^n |x_i|\right\| \sup_{1\le i\le n}\|y_i\| : u=\sum_{i=1}^n x_i\otimes y_i\right\} \text{ for all } u\in E\otimes_l Y$$

and

$$\|u\|_m = \inf\left\{\left\|\sum_{i=1}^n |y_i|\right\| \sup_{1\le i\le n}\|x_i\| : u=\sum_{i=1}^n x_i\otimes y_i\right\} \text{ for all } u\in Y\otimes_m E.$$

It is also well known that if E, F and Y are vector spaces, then there exists a unique linear bijection $\gamma : E \otimes (F \otimes Y) \to (E \otimes F) \otimes Y$ such that $\gamma(x \otimes (y \otimes z)) = (x \otimes y) \otimes z$ (see [73]).

Theorem 6.3.1. *Let E and F be Banach lattices and Y a Banach space. The unique linear bijection $\gamma : E \otimes (F \otimes Y) \to (E \otimes F) \otimes Y$ such that $\gamma(x \otimes (y \otimes z)) = (x \otimes y) \otimes z$ for all $(x, y, z) \in E \times F \times Y$, has a unique extension $\Gamma : E\widetilde{\otimes}_l(F\widetilde{\otimes}_l Y) \to (E\widetilde{\otimes}_l F)\widetilde{\otimes}_l Y$, which is an isometric isomorphism.*

Proof. We first show that $\gamma : E \otimes_l (F \otimes_l Y) \to (E \otimes_l F) \otimes_l Y$ is continuous and $\|\gamma\| \leq 1$. Let $w \in E \otimes_l (F \otimes_l Y)$. Then

$$w = \sum_{i=1}^{n} x_i \otimes u_i \text{ for } x_1, \ldots, x_n \in E \text{ and } u_1, \ldots, u_n \in F \otimes_l Y,$$

and for each i, where $1 \leq i \leq n$,

$$u_i = \sum_{j=1}^{m} y_j^i \otimes z_j^i \text{ for } y_1^i, \ldots, y_m^i \in F \text{ and } z_1^i, \ldots, z_m^i \in Y.$$

Hence,

$$w = \sum_{i,j} x_i \otimes (y_j^i \otimes z_j^i) \text{ and } \gamma(w) = \sum_{i,j} (x_i \otimes y_j^i) \otimes z_j^i.$$

Since

$$\begin{aligned}
\|\gamma(w)\|_{(E\otimes_l F)\otimes_l Y} &\leq \Big\| \sum_{i,j} |x_i \otimes y_j^i| \Big\|_{E\otimes_l F} \sup_{i,j} \|z_j^i\| \\
&= \Big\| \sum_{i,j} |x_i| \otimes |y_j^i| \Big\|_{E\otimes_l F} \sup_{i,j} \|z_j^i\| \\
&\leq \Big(\Big\| \sum_{i} |x_i| \Big\| \sup_{i,j} \|y_j^i\| \Big) \sup_{i,j} \|z_j^i\| \\
&\leq \Big\| \sum_{i} |x_i| \Big\| \Big(\Big\| \sum_{i,j} |y_j^i| \Big\| \sup_{i,j} \|z_j^i\| \Big),
\end{aligned}$$

it follows that

$$\|\gamma(w)\|_{(E\otimes_l F)\otimes_l Y} \leq \Big\| \sum_{i} |x_i| \Big\| \sup_{i} \|u_i\|_{F\otimes_l Y},$$

and consequently, $\|\gamma(w)\|_{(E\otimes_l F)\otimes_l Y} \leq \|w\|_{E\otimes_l (F\otimes_l Y)}$.

Next, we show that $\gamma^{-1} : (E \otimes_l F) \otimes_l Y \to E \otimes_l (F \otimes_l Y)$ is continuous and $\|\gamma^{-1}\| \leq 1$. Consider the linear bijection $\delta : Y \otimes_m (F \otimes_m E) \to (Y \otimes_m F) \otimes_m E$, defined by

$$\delta(y \otimes (f \otimes e)) = (y \otimes f) \otimes e.$$

Let $w \in Y \otimes_m (F \otimes_m E)$. Then

$$w = \sum_{i=1}^{n} x_i \otimes u_i \text{ for } x_1, \ldots, x_n \in Y \text{ and } u_1, \ldots, u_n \in F \otimes_m E,$$

and for each i, where $1 \le i \le n$,

$$u_i = \sum_{j=1}^{m} y^i_j \otimes z^i_j \text{ for } y^i_1, \ldots, y^i_m \in F \text{ and } z^i_1, \ldots, z^i_m \in E.$$

Hence,

$$w = \sum_{i,j} x_i \otimes (y^i_j \otimes z^i_j) \text{ and } \delta(w) = \sum_{i,j} (x_i \otimes y^i_j) \otimes z^i_j.$$

Since

$$\begin{aligned} \|\delta(w)\|_{(Y\otimes_m F)\otimes_m E} &\le \sup_{i,j} \|x_i \otimes y^i_j\|_{Y\otimes_m F} \left\| \sum_{i,j} |z^i_j| \right\| \\ &= \left(\sup_i \|x_i\| \sup_{i,j} \|y^i_j\| \right) \left\| \sum_{i,j} |z^i_j| \right\| \\ &= \sup_i \|x_i\| \left(\sup_{i,j} \|y^i_j\| \left\| \sum_{i,j} |z^i_j| \right\| \right), \end{aligned}$$

it follows that

$$\|\delta(w)\|_{(Y\otimes_m F)\otimes_m E} \le \sup_i \|x_i\| \left\| \sum_i |u_i| \right\|_{F\otimes_m E}.$$

Consequently,

$$\|\delta(w)\|_{(Y\otimes_m F)\otimes_m E} \le \|w\|_{Y\otimes_m (F\otimes_m E)}.$$

But, as $(E \widetilde{\otimes}_l F)\widetilde{\otimes}_l Y$ and $Y \widetilde{\otimes}_m (F\widetilde{\otimes}_m E)$ are isometrically isomorphic, the map

$$\delta : Y\widetilde{\otimes}_m(F\widetilde{\otimes}_m E) \hookrightarrow (Y\widetilde{\otimes}_m F)\widetilde{\otimes}_m E$$

is continuous with $\|\delta\| \le 1$ and as $(Y\widetilde{\otimes}_m F)\widetilde{\otimes}_m E$ and $E\widetilde{\otimes}_l(F\widetilde{\otimes}_l Y)$ are isometrically isomorphic, we get that $\gamma^{-1} : (E \otimes_l F) \otimes_l Y \to E \otimes_l (F \otimes_l Y)$ is continuous and $\|\gamma^{-1}\| \le 1$.

Since $\|\gamma\| \le 1$ and $\|\gamma^{-1}\| \le 1$, we get that $\|\gamma\| = 1$. Hence, $\gamma : E \otimes_l (F \otimes_l Y) \to (E \otimes_l F) \otimes_l Y$ is an isometry. By a standard density argument γ has a unique extension $\Gamma : E\widetilde{\otimes}_l(F\widetilde{\otimes}_l Y) \to (E\widetilde{\otimes}_l F)\widetilde{\otimes}_l Y$, which is an isometric isomorphism. □

As an application of the associativity of the l-norm, we derive a result noted by Popa in [132] (see Theorem 6.3.7 below).

First, we recall some necessary terminology from [30, 31]. Let $(\Omega, \mathcal{F}, P)$ be a finite measure space, $1 \le p < \infty$ and $(\mathcal{F}_i)$ an increasing sequence of sub σ-algebras of $\mathcal{F}$. Let E be a Banach lattice and Y a Banach space.

Definition 6.3.2.

(i) If (T_i) is a commuting sequence (i.e. $T_iT_j = T_i = T_jT_i$ for all $i \leq j$) of contractive projections on Y, then (T_i) is called a *BS-filtration* on Y.

(ii) If (T_i) is a BS-filtration on E such that each $T_i \geq 0$ and the range $\mathcal{R}(T_i)$ of T_i, for each $i \in \mathbb{N}$, is a (closed) Riesz subspace of E, then (T_i) is called a *BL-filtration* on E.

It is well known that the sequence $(\mathbb{E}[\,\cdot\,|\,\mathcal{F}_i])$ of conditional expectations on $L_p(P)$ is a BS-filtration on $L_p(P)$.

In [142, p.214] it is shown that, if $T : E \to E$ is a projection which is strictly positive (i.e. $\{f \in E : T(|f|) = 0\} = \{0\}$) on a Banach lattice E, then $\mathcal{R}(T)$ is a Banach sublattice of E. Thus, the sequence $(\mathbb{E}[\,\cdot\,|\,\mathcal{F}_i])$ of conditional expectations on $L_p(P)$ is a BL-filtration on the Banach lattice $L_p(P)$.

Definition 6.3.3.

(i) If (T_i) is a BS-filtration on Y and $(f_i) \subseteq Y$, then the pair (f_i, T_i) is a *martingale* in Y, if $T_if_j = f_i$ for all $i \leq j$.

(ii) If (f_i, T_i) is a martingale in Y, then (f_i, T_i) is *fixed* if there exists $f \in Y$ such that $f_i = T_if$ for all $i \in \mathbb{N}$.

Let

$$\begin{aligned}
\mathcal{M}(Y, T_i) &= \{(f_i, T_i) \text{ is a martingale in } Y : \sup_i \|f_i\| < \infty\},\\
\mathcal{M}_{\mathrm{nc}}(Y, T_i) &= \{(f_i, T_i) \in \mathcal{M}(Y, T_i) : (f_i) \text{ is norm convergent in } Y\}\,, \text{ and}\\
\mathcal{M}_{\mathrm{f}}(Y, T_i) &= \{(f_i, T_i) \in \mathcal{M}(Y, T_i) : (f_i, T_i) \text{ is fixed}\}\,.
\end{aligned}$$

It is easily verified that if (T_i) is a BS-filtration on E, then the sequence of adjoint maps (T_i^*) is a BS-filtration on E^*. It is known that if Y is a Banach space and (T_i) is a BL-filtration on E, then $(T_i^* \otimes_l id_Y)$ is a BS-filtration on $E^*\widetilde{\otimes}_l Y$ (see [31]), and $(T_i \otimes_l id_Y)$ is a BL-filtration on $E\widetilde{\otimes}_l Y$ (see [30]).

We require the following two theorems taken from [31].

Theorem 6.3.4. ([31, Theorem 3.8]) *Let Y be a Banach space. Then the following are equivalent.*

(i) Y has the Radon-Nikodým property.

(ii) $E^*\widetilde{\otimes}_l Y = \mathcal{L}^{\text{cas}}(E, Y)$ *for all separable Banach lattices* E *with order continuous dual.*

(iii) $\mathcal{M}(E^*\widetilde{\otimes}_l Y, T_i^* \otimes_l \text{id}_Y) = \mathcal{M}_{\text{f}}(E^*\widetilde{\otimes}_l Y, T_i^* \otimes_l \text{id}_Y)$ *for all separable Banach lattices* E *with order continuous dual and all BL-filtrations* (T_i) *on* E.

(iv) $\mathcal{M}(E\widetilde{\otimes}_l Y, T_i \otimes_l \text{id}_Y) = \mathcal{M}_{\text{nc}}(E\widetilde{\otimes}_l Y, T_i \otimes_l \text{id}_Y)$ *for all separable reflexive Banach lattices* E *with order continuous dual and all complemented quasi-interior preserving BL-filtrations* (T_i) *on* E.

(v) $\mathcal{M}(E\widetilde{\otimes}_l Y, T_i \otimes_l \text{id}_Y) = \mathcal{M}(E, T_i)\widetilde{\otimes}_l Y$ *for all separable reflexive Banach lattices* E *with order continuous dual and all complemented quasi-interior preserving BL-filtrations* (T_i) *on* E.

Theorem 6.3.5. ([31, Corollary 3.9]) *Let* Y *be a Banach space. Then* Y^* *has the Radon-Nikodým property if and only if* $E^*\widetilde{\otimes}_l Y^* = (E\widetilde{\otimes}_l Y)^*$ *for all separable Banach lattices* E *with order continuous dual.*

The following result was noted by Popa (see [132]) and an alternative proof may be found in [109].

Theorem 6.3.6. *Let* E *and* F *be Banach lattices each with order continuous norm. Then* $E\widetilde{\otimes}_l F$ *has order continuous norm.*

Based on a martingale approach and the associativity of the l-tensor product, we give an alternative proof for Popa's Radon-Nikodým theorem [132].

Theorem 6.3.7. (Popa) *Let* Y *be a Banach space and let* E *be a separable Banach lattice. Then* E^* *and* Y *have the Radon-Nikodým property if and only if* $E^*\widetilde{\otimes}_l Y$ *has the Radon-Nikodým property.*

Proof. Suppose E^* and Y have the Radon-Nikodým property. We verify that $E^*\widetilde{\otimes}_l Y$ has the Radon-Nikodým property, by using Theorem 6.3.4.

Consider a separable Banach lattice F for which F^* has order continuous norm. Let (T_i) be a BL-filtration on F and let $(f_i, T_i^* \otimes_l id_{E^*\widetilde{\otimes}_l Y})$ be a norm bounded martingale in $F^*\widetilde{\otimes}_l(E^*\widetilde{\otimes}_l Y)$. As F is separable and E^* has the Radon-Nikodým property, Theorem 6.3.5 implies that $F^*\widetilde{\otimes}_l E^* = (F\widetilde{\otimes}_l E)^*$. Since E and F are separable, it readily follows that $F\widetilde{\otimes}_l E$ is separable. The assumption that E^* has the Radon-Nikodým property implies that E^* has order continuous norm ([125]). Theorem 6.3.6 then yields the order continuity of the norm of $F^*\widetilde{\otimes}_l E^* = (F\widetilde{\otimes}_l E)^*$. Hence, $\big(f_i, (T_i \otimes_l id_E)^* \otimes_l id_Y\big)$ is a norm bounded martingale in $(F\widetilde{\otimes}_l E)^*\widetilde{\otimes}_l Y$ and $(T_i \otimes_l id_E)$ is a BL-filtration on $F\widetilde{\otimes}_l E$. As Y has the Radon-Nikodým property by assumption, an application of Theorem 6.3.4 yields that $\big(f_i, (T_i \otimes_l id_E)^* \otimes_l id_Y\big)$ is fixed. By Theorem 6.3.1 the l-norm is associative, so under identification, the martingale $(f_i, T_i^* \otimes_l id_{E^*\widetilde{\otimes}_l Y})$ in $F^*\widetilde{\otimes}_l(E^*\widetilde{\otimes}_l Y)$ is fixed. Thus, by Theorem 6.3.4, $E^*\widetilde{\otimes}_l Y$ has the Radon-Nikodým property.

Conversely, if $E^*\widetilde{\otimes}_l Y$ has the Radon-Nikodým property, then E^* and Y have the Radon-Nikodým property, as both spaces are closed subspaces of $E^*\widetilde{\otimes}_l Y$ (see [52, p.217]). □

6.4 The Radon-Nikodým property in $H_\Phi(P, Y)$

The aim of this section is to consider the Radon-Nikodým property in $H_\Phi(P, Y)$. This requires a stronger version of Popa's Radon-Nikodým Theorem: E^* needs to be replaced by E in Theorem 6.3.7. This requires an application of the following highly non-trivial result noted by Talagrand in [152, 153] (see also [125]).

Theorem 6.4.1. *Let E be a separable Banach lattice. Then E has the Radon-Nikodým property if and only if E is the dual of a separable Banach lattice.*

The main result needed for describing the Radon-Nikodým property in $H_\Phi(P, Y)$ is the following.

Theorem 6.4.2. *Let Y be a Banach space and let E be a separable Banach lattice. Then E and Y have the Radon-Nikodým property if and only if $E\widetilde{\otimes}_l Y$ has the Radon-Nikodým property.*

Proof. Suppose that E and Y have the Radon-Nikodým property. As E is separable, there exists a separable Banach lattice E_0 such that $E = E_0^*$, by Talagrand's theorem. By Theorem 6.3.7, $E\widetilde{\otimes}_l Y$ has the Radon-Nikodým property.

Conversely, if $E\widetilde{\otimes}_l Y$ has the Radon-Nikodým property, then E and Y have the Radon-Nikodým property, as both spaces are closed subspaces of $E\widetilde{\otimes}_l Y$ (see [52, p.217]). □

As a consequence of Theorems 6.1.3 and 6.4.2, we obtain the following result.

Theorem 6.4.3. *Let $(\Omega, \mathcal{F}, P)$ denote a probability space, Φ a finite Young function and Y a Banach space. Then $H_\Phi(P)$ and Y have the Radon-Nikodým property if and only if $H_\Phi(P, Y)$ has the Radon-Nikodým property.*

Corollary 6.4.4. *Let $(\Omega, \mathcal{F}, P)$ denote a probability space, Φ a finite Young function and Y a Banach space. If the Δ_2-condition holds for large u, then $L_\Phi(P)$ and Y have the Radon-Nikodým property if and only if $L_\Phi(P, Y)$ has the Radon-Nikodým property.*

Proof. If the Δ_2-condition holds for large u, then $H_\Phi(P, Y) = L_\Phi(P, Y)$ for any Banach space Y. Thus, the result follows from Theorem 6.4.3. □

Sundaresan [151] and Turret and Uhl [154] obtained results weaker than Corollary 6.4.4 under different assumptions on $L_\Phi(P)$.

6.5 Martingale convergence in $H_\Phi(P, Y)$

Let $\mathcal{F}_1$ be a sub σ-algebra of $\mathcal{F}$ and $1 \leq p < \infty$. The *conditional expectation* of $f \in L_p(P, Y)$ relative to $\mathcal{F}_1$, denoted by $\mathbb{E}[f \mid \mathcal{F}_1]$, is the unique $\mathcal{F}_1$-measurable element of $L_p(P, Y)$ which is given by

$$\int_A \mathbb{E}[f \mid \mathcal{F}_1] dP = \int_A f dP \quad \text{for all } A \in \mathcal{F}_1.$$

The map $\mathbb{E}[\cdot \mid \mathcal{F}_1] : L_p(P, Y) \to L_p(P, Y)$ is a contractive linear projection (see [30, 52]). Furthermore, if we identify $S(P, Y)$ and $S(P) \otimes Y$, then

$$\mathbb{E}\left[\sum_{i=1}^{n} \chi_{A_i}(\cdot) y_i \mid \mathcal{F}_1\right] = (\mathbb{E}[\cdot \mid \mathcal{F}_1] \otimes id_Y)\left(\sum_{i=1}^{n} \chi_{A_i} \otimes y_i\right), \tag{6.3}$$

where $\mathbb{E}[\chi_{A_i} | \mathcal{F}_1]$ denotes the conditional expectation of $\chi_{A_i} \in L_p(P)$ (see [30, 52]). Since $S(P, Y)$ is dense in $L_p(P, Y)$, and since $S(P) \otimes Y$ is dense in $L_p(P) \widetilde{\otimes}_l Y$, it follows that the conditional expectation operator $\mathbb{E}[\cdot \mid \mathcal{F}_1]$ on $L_p(P, Y)$ is the continuous extension of $\mathbb{E}[\cdot \mid \mathcal{F}_1] \otimes id_Y$ to $L_p(P) \widetilde{\otimes}_l Y$.

Next, we consider the situation in Banach space-valued Orlicz spaces.

Lemma 6.5.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{F}_1$ a sub σ-algebra of $\mathcal{F}$ and Φ a finite Young function.*

(i) The conditional expectation operator $\mathbb{E}[\cdot \mid \mathcal{F}_1]$ on $L_1(P, Y)$ restricted to $L_\Phi(P, Y)$, again denoted by $\mathbb{E}[\cdot \mid \mathcal{F}_1]$, is a contractive projection from $L_\Phi(P, Y)$ to $L_\Phi(P, Y)$.

(ii) The conditional expectation operator $\mathbb{E}[\cdot \mid \mathcal{F}_1]$ on $L_1(P, Y)$ restricted to $H_\Phi(P, Y)$, again denoted by $\mathbb{E}[\cdot \mid \mathcal{F}_1]$, is a contractive projection from $H_\Phi(P, Y)$ to $H_\Phi(P, Y)$.

(iii) The continuous extension to $H_\Phi(P)\widetilde{\otimes}_l Y$ of $\mathbb{E}[\cdot \mid \mathcal{F}_1] \otimes id_Y : S(P) \otimes Y \to S(P) \otimes Y$ is the conditional expectation operator $\mathbb{E}[\cdot \mid \mathcal{F}_1] : H_\Phi(P,Y) \to H_\Phi(P,Y)$.

Proof. (i): Let $f \in L_\Phi(P,Y)$. Then there exists $a > 0$ such that $\Phi(\frac{1}{a}\|f\|) \in L_1(P)$. Consider $\phi = \Phi \circ \|\cdot\|$, which is a convex function. By Jensen's inequality, we get that

$$\Phi\left(\frac{1}{a}\| \,\mathbb{E}[f|\mathcal{F}_1]\, \|\right) = \phi\left(\mathbb{E}\left[\frac{1}{a} f|\mathcal{F}_1\right]\right) \leq \mathbb{E}\left[\phi\left(\frac{1}{a} f\right)|\mathcal{F}_1\right] = \mathbb{E}\left[\Phi\left(\frac{1}{a}\|f\|\right)|\mathcal{F}_1\right] \text{ a.s..}$$

Hence, $\mathbb{E}[f|\mathcal{F}_1] \in L_\Phi(P,X)$ and $\mathcal{N}_\Phi(\mathbb{E}[f|\mathcal{F}_1]) \leq \mathcal{N}_\Phi(f)$. But, as $\mathbb{E}[\cdot|\mathcal{F}_1]$ is a projection, we get that $\mathcal{N}_\Phi(\mathbb{E}[f|\mathcal{F}_1]) = \mathcal{N}_\Phi(f)$. Thus, $\mathbb{E}[\cdot \mid \mathcal{F}_1] : L_\Phi(P,X) \to L_\Phi(P,X)$ is a contractive linear projection.

(ii): Let $f \in H_\Phi(P,Y)$. Then, $\Phi(\frac{1}{a}\|f\|) \in L_1(P)$ for every $a > 0$. As in (i), it follows that $\mathbb{E}[\cdot|\mathcal{F}_1] : H_\Phi(P,Y) \to H_\Phi(P,Y)$ is a contractive linear projection.

(iii): Suppose that Φ is a finite Young function. If $\chi_{A_i} \in L_1(P)$, then $\chi_{A_i} \in H_\Phi(P)$ and $\mathbb{E}[\chi_{A_i}|\mathcal{F}_1] \in H_\Phi(P)$. As $S(P,Y)$ is dense in $H_\Phi(P,Y)$ and $S(P) \otimes Y$ is dense in $H_\Phi(P)\widetilde{\otimes}_l Y$, it follows from (6.3) that the continuous extension of $\mathbb{E}[\cdot \mid \mathcal{F}_1] \otimes id_Y$ to $H_\Phi(P)\widetilde{\otimes}_l Y$ is the conditional expectation $\mathbb{E}[\cdot \mid \mathcal{F}_1]$ on $H_\Phi(P,Y)$. □

A martingale (f_n) in $H_\Phi(P,Y)$ is *norm-convergent* if there exists $f \in H_\Phi(P,Y)$ such that $\mathcal{N}_\Phi(f - f_n) \to 0$ as $n \to \infty$. From this point on we shall simply refer to a norm-convergent martingale as *convergent.*

Theorem 6.5.2. *Let Y be a Banach space. Then Y has the Radon Nikodým property if and only if for all probability spaces $(\Omega, \mathcal{F}, P)$, all filtrations $(\mathcal{F}_i)$, and all finite Young functions Φ for which $H_\Phi(P)$ is separable and reflexive, every $\mathcal{N}_\Phi$-bounded martingale $(f_i, \mathcal{F}_i)$ in $H_\Phi(P,Y)$ is $\mathcal{N}_\Phi$-convergent.*

Proof. Assume that Y has the Radon Nikodým property. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $(\mathcal{F}_i)$ a filtration, and Φ a finite Young function for which $H_\Phi(P)$ is separable and reflexive. Suppose that $(f_i, \mathcal{F}_i)$ is a martingale in $H_\Phi(P,Y)$ such that $\sup_i \|f_i\|_l = \sup_i \mathcal{N}_\Phi(f_i) < \infty$.

Let $\bigvee_{i=1}^\infty \mathcal{F}_i$ denote the σ-algebra generated by $\bigcup_{i=1}^\infty \mathcal{F}_i$. It is readily verified that the conditions of (d) in Theorem 6.3.4 are satisfied by the sequence $\big(\mathbb{E}[\cdot \mid \mathcal{F}_1], \mathbb{E}[\cdot \mid \mathcal{F}_2], \ldots, \mathbb{E}[\cdot \mid \bigvee_{i=1}^\infty \mathcal{F}_i]\big)$.

It follows from (a) of Theorem 6.3.4 that the martingale $(f_i, \mathcal{F}_i)$ is $\mathcal{N}_\Phi$-convergent.

Conversely, assume that for all probability spaces $(\Omega, \mathcal{F}, P)$, all filtrations $(\mathcal{F}_i)$, and all finite Young functions Φ for which $H_\Phi(P)$ is separable and reflexive, every $\mathcal{N}_\Phi$-bounded martingale $(f_i, \mathcal{F}_i)$ in $H_\Phi(P, Y)$ is $\mathcal{N}_\Phi$-convergent.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $1 < p < \infty$, and $(\mathcal{F}_i)$ a filtration, and suppose that $(f_i, \mathcal{F}_i)$ is a martingale in $L_p(P, Y)$ such that $\sup_i \|f_i\|_l = \sup_i \Delta_p(f_i) < \infty$. We show that the martingale $(f_i, \mathcal{F}_i)$ in $L_p(P, Y)$ is norm convergent.

Define Φ by $\Phi(u) = \frac{u^p}{p}$ for all $u \in [0, \infty)$. Then Φ is a finite Young function which yields $L_p(P)$, endowed with the norm $p^{-\frac{1}{p}} \Delta_p$, as Orlicz heart. By Lemma 6.2.4, $(L_p(P), p^{-\frac{1}{p}} \Delta_p)$ is reflexive. As $\sup_i p^{-\frac{1}{p}} \|f_i\|_l = \sup_i p^{-\frac{1}{p}} \Delta_p(f_i) < \infty$, the assumption implies that $(f_i, \mathcal{F}_i)$ is $p^{-\frac{1}{p}} \Delta_p(\cdot)$-convergent; hence, $(f_i, \mathcal{F}_i)$ is also $\Delta_p(\cdot)$-convergent.

Thus, for all probability spaces $(\Omega, \mathcal{F}, P)$, $1 < p < \infty$ and all filtrations $(\mathcal{F}_i)$ of $\mathcal{F}$, every norm bounded martingale $(f_i, \mathcal{F}_i)$ in $L_p(P, Y)$ is norm convergent. It is well-known that the latter is equivalent to Y having the Radon-Nikodým property (see [52]). □

Chapter 7

Set-valued risk measures on Orlicz hearts

The concept of coherent risk measures together with its axiomatic characterisation was introduced by Artzner et al. [6] in 1999 in a finite probability space setting and generalised to a general probability space setting by Delbaen [42] in 2002. The following approach was used in defining the risk measure. Among the set of all possible financial positions, an investor chooses a subset $\mathcal{A}$ of acceptable positions, which he regards as risk-free. The risk measure $\rho(x)$ then corresponds to the extra capital required at the beginning of the investment in some 'secure' instrument, usually a money market account, so that the resulting position is acceptable, i.e. $x + \rho(x) \in \mathcal{A}$. The axioms are there to guarantee the economic coherence of the risk measure.

The concept of risk measures has been studied and extended by many authors. Föllmer and Schied [60] and independently Frittelli and Rosazza Gianin [62] generalised the definition of a risk measure to a convex risk measure. Cheridito et al. [25] extended the dual representation of both coherent and convex risk measures to the space of càdlàg processes. Cheridito and Li [26, 27] looked at convex risk measures in an Orlicz space setting.

In all the above-mentioned methods, the risky portfolio under consideration is a given real-valued random variable and the risk measure is a map into $\mathbb{R}$. In other words, these risk measures do not consider portfolio aggregation. In reality, however, investors have access to different markets and form multi-asset portfolios. It is not always possible or desirable to transform a multi-dimensional portfolio into a position in one financial market, i.e. the position cannot be described by one real-valued number. The reason for this could be transaction costs, liquidity bounds, fluctuating exchange rates, etc.

Thus, we could require a risk measure that takes values in $\mathbb{R}^d$ and gives us a value in $\mathbb{R}^m$, where $m \leq d$. These m markets could, for example, be money market accounts in different currencies. In other words, it is necessary to look at risk measures in a set-valued setting.

Set-valued risk measures have gained in popularity over the past few years. Jouini et al. were among the first to introduce the set-valued coherent risk measure (see [91]). Since then, amongst others, Hamel et al. [77, 79] extended the approach of Jouini et al. to define set-valued convex risk measures and Konstantinides and Kountzakis [101] used the ideas from Stoica [149] and Jaschke and Küchler [89] to define risk measures on partially ordered normed linear spaces.

Hamel et al. [78, 79] defined convex set valued risk measures on the space $L_p(P, \mathbb{R}^d)$ of Bochner p-integrable functions with values in $\mathbb{R}^d$. Their method for the case $1 \leq p < \infty$ can be generalised to include spaces $H_\Phi(P, \mathbb{R}^d)$ of $\mathbb{R}^d$-valued Orlicz hearts. In view of the connection between utility functions and real-valued Orlicz spaces, as noted by Frittelli and his co-workers [15], this extension to Orlicz spaces may be of interest. We use tensor products of Banach lattices and Banach spaces to achieve our goal.

The reader is advised to first proceed to the appendix for an outline of the basic relevant preliminaries on tensor products, Bochner spaces and the l-norm of Chaney and Schaefer.

Using the result from Chaney and Schaefer [23, 142] that states that $L_p(P, \mathbb{R}^d)$ is isometrically isomorphic to $L_p(P)\widetilde{\otimes}_l\mathbb{R}^d$, we show that the results of Hamel et al. [79] can be obtained via a tensor-product approach. In addition, by the tensor product approach and using a result from Labuschagne and Offwood [114], we get a representation of set-valued convex risk measures on vector-valued Orlicz hearts.

This chapter is based on [116].

7.1 Set-valued setting

Hamel et al. [79] replaced the range $(-\infty, \infty]$ of the risk measure defined in Chapter 4, by an appropriate space to generalise the notion of risk measures to a set-valued setting. For this purpose, we need to fix some terminology and notation.

Let Y be a Banach lattice and let $G \subseteq Y$ be a Banach subspace of Y. The set of all subsets of G will be denoted by $\mathcal{P}(G)$.

For all $M, L \in \mathcal{P}(G)$, we denote the Minkowski sum $M + L$ of M and L by

$$M + L = \{m + l : m \in M \text{ and } l \in L\},$$

with the convention $\emptyset + M = M + \emptyset = \emptyset$ for all $M \in \mathcal{P}(G)$, where the empty set is considered to be closed and convex.

A *cone* C is a subset of a vector space for which $\lambda C \subseteq C$ for all $\lambda > 0$. If, in addition, C is convex, then C is called a *convex cone*.

Define $K_Y \subseteq Y$ to be a closed convex cone such that $Y_+ \subseteq K_Y$. This cone generates a reflexive, translative relation in Y, given by

$$x \leq_{K_Y} y \quad \Longleftrightarrow \quad y - x \in K_Y.$$

We do not assume that $\leq_{K_Y}$ is antisymmetric (i.e. we do not assume that $K_Y \cap (-K_Y) = \{0\}$). We do assume that $K_Y \cap (-\operatorname{int} Y_+) = \emptyset$.

We consider the cone K on G induced by the cone K_Y. As $K = K_Y \cap G$, we have that K is closed and convex, $G_+ \subseteq K$ and the ordering on Y induced by K_Y and the ordering on G induced by K is the same.

The order relation $\leq_K$ can be canonically extended to $\mathcal{P}(G)$ by $A \leq_K B$ if and only if $B \subseteq A + K$ or equivalently $A \subseteq B - K$. Hence, for $A, B \in \mathcal{P}(G)$, $A \leq_K B$ if and only if $B + K \subseteq A + K$. We say that two sets $A, B \in \mathcal{P}(G)$ are *equivalent* if $A \leq_K B$ and $B \leq_K A$, i.e. $A + K = B + K$. Thus, we define the set

$$\mathcal{P}_K = \{M \in \mathcal{P}(G) : M = M + K\}$$

and can identify it with the set of equivalence classes with respect to the above-mentioned equivalence relation. For $A, B \in \mathcal{P}_K$, we have $A \leq_K B$ if and only if $B \subseteq A$.

Next we define some concepts in this set-valued setting with respect to the partial ordering $\leq_K$.

Definition 7.1.1.

(i) A function $f : \mathcal{X} \to \mathcal{P}(G)$ is *convex* if for all $\lambda \in (0,1)$ and $x, y \in \mathcal{X}$

$$f(\lambda x + (1-\lambda) y) \leq_K \lambda f(x) + (1-\lambda) f(y).$$

(ii) The *convex hull* of a function $f : \mathcal{X} \to \mathcal{P}(G)$ is the (uniquely determined) function $\operatorname{co} f : \mathcal{X} \to \mathcal{P}(G)$ which satisfies $\operatorname{epi}(\operatorname{co} f) = \operatorname{co}(\operatorname{epi} f)$.

(iii) A function $f : \mathcal{X} \to \mathcal{P}_K$ is *subadditive* if for all $x, y \in \mathcal{X}$

$$f(x+y) \leq_K f(x) + f(y).$$

(iv) A function $f : \mathcal{X} \to \mathcal{P}_K$ is *positive homogeneous* if for all $t > 0$ and $x \in \mathcal{X}$

$$f(tx) \leq_K tf(x).$$

Proposition 7.1.2. *A function $f : \mathcal{X} \to \mathcal{P}(G)$ is convex if and only if the epigraph, defined by*

$$\text{epi}f = \{(x, g) \in \mathcal{X} \times G : g \in f(x) + K\},$$

is convex.

The image space of a set-valued convex function $f : \mathcal{X} \to \mathcal{P}(G)$ is the collection of *upper convex* subsets of G defined by

$$\mathcal{P}_K^c = \{M \in \mathcal{P}(G) : M = \text{co}(M + K)\}.$$

Definition 7.1.3.

(i) A function $f : \mathcal{X} \to \mathcal{P}(G)$ is *closed* if epi$f \subseteq \mathcal{X} \times G$ is a closed set with respect to the product topology on $\mathcal{X} \times G$.

(ii) The *closed hull* of a function $f : \mathcal{X} \to \mathcal{P}(G)$ is the (uniquely determined) function cl$f : \mathcal{X} \to \mathcal{P}(G)$ which satisfies epi(cl$f$) = cl(epi$f$).

A closed function automatically maps into the collection of *upper closed* subsets of G defined by

$$\mathcal{K} = \{M \subseteq G : M = \text{cl}(M + K)\}.$$

The Minkowski sum of two closed sets in G is not closed in general, but the addition $\oplus$, defined by

$$M_1 \oplus M_2 = \text{cl}(M_1 + M_2),$$

has the property that $\oplus : \mathcal{K} \times \mathcal{K} \to \mathcal{K}$. Moreover, $\oplus$ is commutative, associative and for all $M \in \mathcal{K}$, we have

$$M \oplus K = K \oplus M = M.$$

We also have that $0M = K$ for $M \in \mathcal{K}$. In other words, the convex cone K serves as zero element in the space $(\mathcal{K}, \oplus, \subseteq)$.

Finally, a closed convex function $f : \mathcal{X} \to \mathcal{P}(G)$ maps into the collection of *upper closed convex* subsets of G, defined by

$$\mathcal{K}_c = \{M \subseteq G : M = \text{cl}\,\text{co}(M + K)\}.$$

Note that $\mathcal{P}_K \subseteq \mathcal{K}_c \subseteq \mathcal{K}$ and $\mathcal{P}_K \subseteq \mathcal{K}_c \subseteq \mathcal{P}_K^c$. Also, note that properties like convexity depend both on the function and on the cone.

Consider a function $F : \mathcal{X} \to \mathcal{K}$. The *graph* of F is given by

$$\text{graph}F = \{(x, g) \in \mathcal{X} \times G : g \in F(x)\},$$

the *epigraph* of F by

$$\text{epi}F = \{(x, g) \in X \times G : g \in F(x) + K\}$$

and its *effective domain* by

$$\text{dom}F = \{x \in \mathcal{X} : F(x) \neq \emptyset\}.$$

If $F : \mathcal{X} \to \mathcal{P}_K$, then graph$F$ = epi$F$. Note that F is convex if and only if its graph is convex, and F is closed if and only if its graph is a closed subset of $\mathcal{X} \times G$.

Definition 7.1.4.

(i) The function $F : \mathcal{X} \to \mathcal{P}_K$ is called *proper* if and only if dom$F \neq \emptyset$ and $F(x) \neq G$ for all $x \in \mathcal{X}$.

(ii) The function F is called K*-proper* if dom$F \neq \emptyset$ and $(F(x) - K)\backslash F(x) \neq \emptyset$ for all $x \in \mathcal{X}$.

These two definitions coincide if K is generating, i.e. $K - K = G$.

Definition 7.1.5. A subset Y of $\mathcal{X}$ is *sequentially closed* if, whenever (x_n) is a sequence in Y converging to x, then x must also be in Y.

Proposition 7.1.6. *Let $K \subseteq G$ be a cone and $F : \mathcal{X} \to \mathcal{K}_c$ be convex and sequentially closed (i.e.* graph$F$ *is sequentially closed in the product topology on $\mathcal{X} \times G$). If there exists $x_0 \in$* dom$F$ *such that $F(x_0) + K \subseteq F(x_0)$, then $F(x) + K \subseteq F(x)$ for all $x \in$* dom$F$.

Proof. Assume there exists $x_0 \in$ dom$F$ such that $F(x_0)+K \subseteq F(x_0)$. This implies that for $g_0 \in F(x_0)$ and $k \in K$, we have $g_0 + nk \in F(x_0)$, i.e. $(x_0, g_0 + nk) \in$ graph$F$ for all $n \in \mathbb{N}$. Consider $x \in$ dom$F$ and $g \in F(x)$, i.e. $(x, g) \in$ graph$F$. The convexity of F implies that

$$\frac{1}{n}(x_0, g_0 + nk) + \frac{n-1}{n}(x, g) = (\frac{1}{n}x_0 + \frac{n-1}{n}x, \frac{1}{n}g_0 + \frac{n-1}{n}g + k) \in \text{graph}F$$

for all $n \in \mathbb{N}$. Taking the limit as $n \to \infty$ and using the fact that epi$F$ is sequentially closed, we get $(x, g + k) \in$ graph$F$, which completes the proof. $\square$

Let $\mathcal{X}$ and G be separated locally convex spaces with topological duals $\mathcal{X}^*$ and G^* respectively. The positive polar cone of K is given by

$$K^\circ = \{g^* \in G^* : g^*(g) \geq 0 \text{ for all } g \in K\}$$

and the negative polar cone by $K^\circ_- = -K^\circ$.

In the (extended) real-valued case, where $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$, the Fenchel conjugate of a function $f : \mathcal{X} \to \overline{\mathbb{R}}$ is $f^* : \mathcal{X}^* \to \overline{\mathbb{R}}$, the definition of which is repeated here for ease of reading: for all $x^* \in \mathcal{X}^*$,

$$f^*(x^*) = \sup_{x \in \mathcal{X}} \{x^*(x) - f(x)\}. \tag{7.1}$$

To extend this to the set-valued case, $x^*(x)$ in (7.1) has to be replaced by a set-valued function that has appropriate properties.

In the case, where G is a real linear space containing at least two elements, Hamel [77] showed that $S_{(x^*,g^*)}$, defined for $x^* \in \mathcal{X}^*$ and $g^* \in G^* \backslash \{0\}$ by

$$S_{(x^*,g^*)}(x) = \{g \in G : x^*(x) + g^*(g) \leq 0\} \text{ for all } x \in \mathcal{X},$$

provides a suitable replacement. Then $S_{(x^*,g^*)} : \mathcal{X} \to \mathcal{P}(G)$.

Consider the special case where $G = \mathbb{R}^m$. Note that the dual of $\mathbb{R}^m$ is isometrically isomorphic to $\mathbb{R}^m$; the map $\theta : (\mathbb{R}^m)^* \to \mathbb{R}^m$, defined by

$$\theta(v)(u) = v^T u,$$

is such an isomorphism. Then, by replacing the $g^*(g)$ by $v^T u$ in the definition of $S_{(x^*,g^*)}$, you arrive at the version in Hamel et al. [79] and by replacing the $x^*(x)$ with $\mathbb{E}[X^T Y]$, you arrive at the version in Hamel and Heyde [77].

Lemma 7.1.7. *For each $x^* \in \mathcal{X}^*$ and $g^* \in G^* \backslash \{0\}$, the function $S_{(x^*,g^*)}$ has the following properties:*

(i) $S_{(x^*,g^*)}(x + y) = S_{(x^*,g^*)}(x) + S_{(x^*,g^*)}(y)$ *for all* $x, y \in \mathcal{X}$,

(ii) $S_{(x^*,g^*)}(tx) = tS_{(x^*,g^*)}(x)$ *for all* $x \in \mathcal{X}$ *and* $t \neq 0$, *and*

(iii) $S_{(y^*+tx^*,g^*)}(x) = S_{(y^*,g^*)}(x) + tS_{(x^*,g^*)}(x)$ *for* $t > 0$ *and* $x \in \mathcal{X}$.

Proof. (i) Consider $u_1 \in S_{(x^*,g^*)}(x_1)$ and $u_2 \in S_{(x^*,g^*)}(x_2)$. Then $x^*(x_1) + g^*(u_1) \leq 0$ and $x^*(x_2) + g^*(u_2) \leq 0$. Hence,

$$x^*(x_1 + x_2) + g^*(u_1 + u_2) = x^*(x_1) + g^*(u_1) + x^*(x_2) + g^*(u_2) \leq 0,$$

which implies $u_1 + u_2 \in S_{(x^*,g^*)}(x_1 + x_2)$.

For the converse inclusion, consider $u \in S_{(x^*,g^*)}(x_1+x_2)$ and $u_1 \in S_{(x^*,g^*)}(x_1)$ such that $x^*(x_1) + g^*(u_1) = 0$. Such a u_1 always exists because $g^* \neq 0$. Set $u_2 = u - u_1$. Then

$$\begin{aligned} x^*(x_2) + g^*(u_2) &= x^*(x_1 + x_2) + g^*(u - u_1) - x^*(x_1) \\ &= x^*(x_1 + x_2) + g^*(u) \leq 0. \end{aligned}$$

Thus, $u_2 \in S_{(x^*,g^*)}(x_2)$, which proves the desired inclusion.

(ii) Follows easily, using (i) repetitively.

(iii) First, note that

$$\begin{aligned} &S_{(y^*,g^*)}(x) + tS_{(x^*,g^*)}(x) \\ &\quad = \{g \in G : y^*(x) + g^*(g) \leq 0\} + t\{g \in G : x^*(x) + g^*(g) \leq 0\} \\ &\quad = \{g \in G : y^*(x) + g^*(g) \leq 0\} + \{g \in G : x^*(x) + \frac{1}{t}g^*(g) \leq 0\} \\ &\quad = \{g \in G : y^*(x) + g^*(g) \leq 0\} + \{g \in G : tx^*(x) + g^*(g) \leq 0\}. \end{aligned}$$

Let $v \in S_{(y^*,g^*)}(x) + tS_{(x^*,g^*)}(x)$, then $v = v_1 + v_2$, where

$$y^*(x) + g^*(v_1) \leq 0$$

and

$$tx^*(x) + g^*(v_2) \leq 0.$$

Thus,

$$y^*(x) + g^*(v_1) + tx^*(x) + g^*(v_2) = y^*(x) + tx^*(x) + g^*(v) \leq 0,$$

i.e. $S_{(y^*,g^*)}(x) + tS_{(x^*,g^*)}(x) \subseteq S_{(y^*+tx^*,g^*)}(x)$.

Conversely, let $u \in S_{(y^*+tx^*,g^*)}(x)$. Then, $y^*(x) + tx^*(x) + g^*(u) \leq 0$. Take u_1 such that $y^*(x) + g^*(u_1) = 0$. Let $u_2 = u - u_1$. Hence,

$$\begin{aligned} tx^*(x) + g^*(u_2) &= y^*(x) + tx^*(x) + g^*(u - u_1) - y^*(x) \\ &= y^*(x) + tx^*(x) + g^*(u) - (y^*(x) + g^*(u_1)) \\ &= y^*(x) + tx^*(x) + g^*(u) \\ &\leq 0. \end{aligned}$$

Hence, $u_1 \in S_{(y^*,g^*)}(x)$ and $u_2 \in tS_{(x^*,g^*)}(x)$ and thus

$$u = u_1 + u_2 \in S_{(y^*,g^*)}(x) + tS_{(x^*,g^*)}(x),$$

completing the proof. □

A classical starting point of convex analysis is to prove that a proper closed convex function is the pointwise supremum of its affine minorants.

A function $h : \mathcal{X} \to \mathcal{P}(G)$ of the form $h(x) = S_{(x^*,g^*)}(x) + \{g\}$ for some $x^* \in \mathcal{X}^*$, $g^* \in K^\circ_- \backslash \{0\}$ and $g \in G$ is called an *affine function*. If an affine function h satisfies $h(x) \leq_K F(x)$ for all $x \in \mathcal{X}$, then h is called an *affine minorant* of F. In other words, if h is an affine minorant of F, then $F(x) \subseteq h(x) + K$ for all $x \in \mathcal{X}$. The case $g^* = 0$ is excluded, to avoid improper minorants.

Lemma 7.1.8. *The epigraph of an affine function is convex.*

Proof. Let $x^* \in \mathcal{X}^*$, $g^* \in K^\circ_- \backslash \{0\}$ and $g \in G$. Consider the affine function $h : \mathcal{X} \to \mathcal{P}(G)$ given by

$$h(x) = S_{(x^*,g^*)}(x) + \{g\}.$$

It is easy to show that h is convex by using the properties of $S_{(x^*,g^*)}$. Thus, the epigraph of an affine function is convex. □

The following theorem and proof are taken from [77].

Theorem 7.1.9. *The following properties are equivalent for a function $F : \mathcal{X} \to \mathcal{P}_K$.*

(i) The function F is the pointwise supremum of its K-proper affine minorants.

(ii) The function F is closed and convex into $\mathcal{K}_c$ and K-proper, or $F \equiv G$ or $F \equiv \emptyset$.

The equivalence remains true if 'K-proper' is replaced with 'proper but not K-proper'.

Proof. (i) ⇒ (ii): Assume that F is the pointwise supremum of its K-proper affine minorants. If F has no K-proper affine minorants, then F will be the pointwise supremum of the empty set, i.e. it will be identically G.

Assume that the family of affine minorants of F is nonempty. As F is the pointwise supremum of its K-proper affine minorants, its epigraph is the intersection of the closed convex epigraphs of the affine functions. Hence, F is closed and convex.

As F is K-proper, F cannot attain the value G. To see this, let $F(x) = G$. Hence, there exists $g \in (G - K)\backslash G$, i.e. there exists $g_1 \in G$ and $k \in K$ with $k \neq 0$ such that $g = g_1 - k$ and $g \notin G$. But $K \subseteq G$ and since G is a vector space, we must have that $g_1 - k \in G$, a contradiction. Thus, $F(x) \neq G$.

If one of the affine minorants is K-proper, i.e. $(h(x) - K)\backslash h(x)$ for some affine minorant h, then since $\sup h \subseteq h$, we must have that the supremum is also K-proper or it is $\emptyset$.

(ii) $\Rightarrow$ (i): The converse will be proved by showing that for all $(x_0, g_0) \notin \text{graph}F$, there is a K-proper (proper but not K-proper) affine minorant h such that $\text{graph}F \subseteq \text{graph } h$ and $(x_0, g_0) \notin \text{graph } h$. In other words, we show that h is an affine minorant but all points that are not in F are also not in h.

Assume that F is closed and convex into $\mathcal{K}_c$. Consider $(x_0, g_0) \notin \text{graph}F$. Since $\text{graph}F$ is closed and convex, we can separate it from (x_0, g_0), getting $x^* \in \mathcal{X}^*$, $g^* \in G^*$ and $\alpha \in \mathbb{R}$ such that for all $(x, g) \in \text{graph}F$

$$x^*(x) + g^*(g) \leq \alpha < x^*(x_0) + g^*(g_0). \tag{7.2}$$

As $\text{dom}F \neq \emptyset$, there exists $y \in \text{dom}F$, such that $F(y) \neq \emptyset$. Hence, there exists $h \in F(y)$, i.e. $(y, h) \in \text{graph}F$. Since $F : \mathcal{X} \to \mathcal{P}_K$, we have that $F(y) \subseteq \mathcal{P}_K$, i.e. $F(y) = F(y) + K$ (and as K is a cone, we have $tK \subseteq K$). Thus, if $h \in F(y)$, then there exists $h_1 \in F(y)$ such that $h = h_1 + tk \in F(y)$, i.e. $(y, h_1 + tk) \in \text{graph}F$, for all $t \geq 0$ and $k \in K$.

Thus,

$$\begin{aligned} x^*(y) + g^*(h + tk) &\leq x^*(x_0) + g^*(g_0) \\ g^*(h) + tg^*(k) &\leq x^*(x_0 - y) + g^*(g_0) \\ g^*(k) &\leq \frac{1}{t}[x^*(x_0 - y) + g^*(g_0) - g^*(h)]. \end{aligned}$$

By the Archimedean property of the reals, we must have that $g^*(k) \leq 0$ for all $k \in K$, i.e. $g^* \in K^\circ_-$.

First, we assume that F is K-proper. Let $g^*(k) < 0$ for all $k \in K$. As g^* is onto, there exists $\bar{g} \in G$ such that $g^*(\bar{g}) = \alpha$. Consider $g \in F(x)$, then using Inequality (7.2), we get

$$x^*(x) + g^*(g - \bar{g}) \leq 0,$$

i.e. $g - \bar{g} \in S_{(x^*, g^*)}(x)$. This implies, that $g \in S_{(x^*, g^*)}(x) + \{\bar{g}\}$, i.e.

$$F(x) \subseteq \{\bar{g}\} + S_{(x^*, g^*)}(x)$$

for all $x \in \mathcal{X}$.

On the other hand, if $v \in \{\bar{g}\} + S_{(x^*, g^*)}(x)$, then $v = \bar{g} + v_1$, where $x^*(x) + g^*(v_1) \leq 0$. This implies that

$$\begin{aligned} x^*(x) + g^*(v - \bar{g}) &\leq 0 \\ x^*(x) + g^*(v) &\leq g^*(\bar{g}) = \alpha. \end{aligned}$$

In other words, if $v \in \{\bar{g}\} + S_{(x^*, g^*)}(x)$, then $x^*(x) + g^*(v) \leq \alpha$. Therefore, by Equation (7.2), we have that (x_0, g_0) does not belong to the graph of the affine function $\{\bar{g}\} + S_{(x^*, g^*)}(x)$.

Next assume $g^*(k) = 0$ for all $k \in K$. Equation (7.2) yields the existence of $\beta > 0$ such that for all $(x, g) \in \text{graph}F$

$$x^*(x) + g^*(g) \leq x^*(x_0) + g^*(g_0) - \beta. \tag{7.3}$$

Take (x_1, g_1) and $k \in K$ such that $x_1 \in \text{dom}F$, $g_1 \in F(x_1)$ and $g_1 - k \notin F(x_1)$. This is possible as F is K-proper. Using the separation argument, we obtain $(x_1^*, g_1^*) \in \mathcal{X}^* \times K_-^\circ$ and $\gamma \in \mathbb{R}$ such that for all $(x, g) \in \text{graph}F$

$$x_1^*(x) + g_1^*(g) \leq \gamma < x_1^*(x_1) + g_1^*(g_1 - k). \tag{7.4}$$

If $g_1^*(k) = 0$ and substituting $(x, g) = (x_1, g_1) \in \text{graph}F$ into Equation (7.4), we get a contradiction. Thus, we may conclude that $g_1^*(k) < 0$ for all $k \in K$. Choose

$$s > \max\{0, \frac{1}{\beta}(\gamma - x_1^*(x_0) - g_1^*(g_0))\}.$$

Multiply Equation (7.3) by s and add the result to the left inequality of Equation (7.4). We obtain, for all $(x, g) \in \text{graph}F$

$$\begin{aligned}(x_1^* + sx^*)(x) + (g_1^* + sg^*)(g) &\leq \gamma + s(x^*(x_0) + g^*(g_0) - \beta) \\ &< \gamma + sx^*(x_0) + sg^*(g_0) - \gamma + x_1^*(x_0) + g_1^*(g_0) \\ &= (x_1^* + sx^*)(x_0) + (g_1^* + sg^*)(g_0),\end{aligned}$$

where the last inequality comes from the definition of s. Thus we get Equation (7.2) with x^*, g^* and α replaced by $x_1^* + sx^*$, $g_1^* + sg^*$ and $\gamma + s(x^*(x_0) + g^*(g_0) - \beta)$. Since $g_1^* + sg^* < 0$, the desired conclusion follows as in the first part.

Secondly, assume that F is proper but not K-proper. Since F is not K-proper and in view of Proposition 7.1.6, we have that $F(x) - K \subseteq F(x)$ for all $x \in \text{dom}F$. Assume $g^*(k) < 0$ for all $k \in K$ and let $g \in F(x)$ for some $x \in \text{dom}F$. Then Equation (7.2) yields for all $t > 0$

$$\begin{aligned}x^*(x) + g^*(g - tk) &< x^*(x_0) + g^*(g_0) \\ g^*(k) &> \frac{1}{t}(x^*(x - x_0) + g^*(g - g_0)).\end{aligned}$$

By the Archimedean property, we get that $g^*(k) \geq 0$ for all $k \in K$, a contradiction.

Thus, we only need to consider the case where $g^*(k) = 0$ for all $k \in K$. As F is proper, we have that $F(x) \neq G$ for all $x \in \text{dom}F$. Hence, there exists $x_2 \in \text{dom}F$ and $g_2 \in G$ such that $g_2 \notin F(x_2)$. Using the separation argument, there exists $(x_2^*, g_2^*) \in \mathcal{X}^* \times G^*$ and $\gamma_2 \in \mathbb{R}$ such that

$$x_2^*(x) + g_2^*(g) \leq \gamma_2 < x_2^*(x_2) + g_2^*(g_2).$$

The rest follows by a similar argument as in part 2 of the K-proper case. □

Corollary 7.1.10. *Let $f : \mathcal{X} \to \mathcal{P}_K^c$ be convex.*

(i) The closure of f is convex and maps into $\mathcal{K}_c$.

(ii) If $h : \mathcal{X} \to \mathcal{K}_c$ is closed and convex such that $f(x) \subseteq h(x)$ for all $x \in \mathcal{X}$, then $(\mathrm{cl} f)(x) \subseteq h(x)$ for all $x \in \mathcal{X}$.

(iii) For all $x \in \mathcal{X}$, $(\mathrm{cl} f)(x) \neq G$ if and only if there exists $(x^*, g^*) \in \mathcal{X}^* \times K^\circ_- \backslash \{0\}$ and $g \in G$ such that for all $x \in \mathcal{X}$

$$(\mathrm{cl} f)(x) \subseteq S_{(x^*,g^*)}(x) + \{g\}.$$

(iv) If there exists $x_0 \in \mathcal{X}$ such that $(\mathrm{cl} f)(x_0) = G$, then $(\mathrm{cl} f)(x) = G$ for all $x \in \mathrm{dom}(\mathrm{cl} f) \supseteq \mathrm{dom} f$.

Proof. (i): The closure of a convex set is convex. Hence, $\mathrm{cl}(\mathrm{graph} f) = \mathrm{graph}(\mathrm{cl} f)$ is convex, which implies that $(\mathrm{cl} f)$ is convex.

(ii): This is due to the fact that $\mathrm{graph} f \subseteq \mathrm{graph}\ h$, which implies that

$$\mathrm{graph}(\mathrm{cl} f) \subseteq \mathrm{graph}(\mathrm{cl}\ h).$$

Since graph h is closed, we have that

$$\mathrm{graph}(\mathrm{cl}\ h) \subseteq \mathrm{graph} h.$$

(iii): If $\mathrm{cl} f$ is proper, the result follows from Theorem 7.1.9. If $\mathrm{cl} f$ is not proper, then $x^* = 0$, $g^* \in K^\circ_- \backslash \{0\}$ and $g = 0$ yields a minorant since in this case, $\mathrm{cl} f = \emptyset$. The converse follows from (ii), with h being the affine minorant.

(iv): This follows from Proposition 7.1.6. □

For the duality theory, the definitions of the conjugate and biconjugate are very important. Hamel [77] defines them as follows.

Definition 7.1.11. The *conjugate* $-F^*$ of $F : \mathcal{X} \to \mathcal{K}$ is defined by

$$-F^*(x^*, g^*) = \mathrm{cl} \bigcup_{x \in \mathcal{X}} [F(x) + S_{(x^*,g^*)}(-x)],$$

and the *biconjugate* F^{**} of F by

$$F^{**}(x) = \bigcap_{(x^*,g^*) \in \mathcal{X}^* \times K^\circ_- \backslash \{0\}} [-F^*(x^*, g^*) + S_{(x^*,g^*)}(x)].$$

The minus sign of $-F^*$ should be read as part of the symbol. This notation is used, so that no subtraction needs to be defined and so that $F+K$ and $-F^*$ have the same image space.

For $F:\mathcal{X}\to\mathcal{P}(G)$ and $g^*\in K^\circ_-\backslash\{0\}$, define a function $\varphi_{F,g^*}:\mathcal{X}\to[-\infty,\infty]$ by

$$\varphi_{F,g^*}(x)=\inf_{g\in F(x)}-g^*(g).$$

The classical Legendre-Fenchel convex conjugate of φ_{F,g^*} is given by

$$\varphi^*_{F,g^*}(x^*)=\sup_{x\in\mathcal{X}}\{x^*(x)-\varphi_{F,g^*}(x)\}.$$

We have that

$$\varphi^*_{F,g^*}(x^*)=\sup_{x\in\mathcal{X},g\in F(x)}\{x^*(x)+g^*(g)\}:=\sigma_{\mathrm{graph}F}(x^*,g^*).$$

The function $\sigma_{\mathrm{graph}F}:\mathcal{X}^*\times G^*\to[-\infty,\infty]$ is called the support function of the graph of the set-valued function F. The following lemma relates this support function to $-F^*$.

Lemma 7.1.12. *For $F:\mathcal{X}\to\mathcal{P}(G)$, $x^*\in\mathcal{X}^*$ and $g^*\in K^\circ\backslash\{0\}$, it holds*

$$-F^*(x^*,g^*)=\{g\in G:g^*(g)\leq\sigma_{\mathrm{graph}F}(x^*,g^*)\},\tag{7.5}$$

and for all $x\in\mathcal{X}$,

$$F^{**}(x)=\bigcap_{(x^*,g^*)\in\mathcal{X}^*\times K^\circ_-\backslash\{0\}}\{g\in G:x^*(x)+g^*(g)\leq\sigma_{\mathrm{graph}F}(x^*,g^*)\}.\tag{7.6}$$

Proof. Firstly, we claim that

$$\{g\in G:g^*(g)<\sigma_{\mathrm{graph}F}(x^*,g^*)\}\subseteq\bigcup_{x\in\mathcal{X}}[F(x)+S_{(x^*,g^*)}(-x)].$$

To show this, let $g\in\{g\in G:g^*(g)<\sigma_{\mathrm{graph}F}(x^*,g^*)\}$. Then

$$\begin{aligned}g^*(g)&<\sigma_{\mathrm{graph}F}(x^*,g^*)\\&=\sup_{x\in\mathcal{X},g\in F(x)}\{x^*(x)+g^*(g)\}.\end{aligned}$$

Thus, there exists $(x,\bar{g})\in\mathrm{graph}F$ such that $g^*(g)\leq x^*(x)+g^*(\bar{g})$. Rewriting this, we have that $x^*(-x)+g^*(g-\bar{g})\leq 0$, i.e. $g-\bar{g}\in S_{(x^*,g^*)}(-x)$. Hence, $\bar{g}\in F(x)$ and $g-\bar{g}\in S_{(x^*,g^*)}(-x)$, which proves our claim.

Conversely, we claim that

$$\bigcup_{x\in\mathcal{X}}[F(x)+S_{(x^*,g^*)}(-x)]\subseteq\{g\in G:g^*(g)<\sigma_{\mathrm{graph}F}(x^*,g^*)\}.$$

Let $g \in \bigcup_{x\in\mathcal{X}}[F(x) + S_{(x^*,g^*)}(-x)]$. Hence, there exists $x \in \mathrm{dom}F$, such that $g = \bar{g} + g'$, where $\bar{g} \in F(x)$ and $g' \in S_{(x^*,g^*)}(-x)$. By the definition of $\varphi_{F,g^*}(x)$, we have that $g^*(\bar{g}) \leq -\varphi_{F,g^*}(x)$ and using the definition of $S_{(x^*,g^*)}$, we get $g^*(g') \leq x^*(x)$. Hence,

$$g^*(g) = g^*(\bar{g} + g') \leq x^*(x) - \varphi_{F,g^*}(x) \leq \sigma_{\mathrm{graph}F}(x^*, g^*),$$

which proves our second claim. Thus, we have shown that

$$\{g \in G : g^*(g) < \sigma_{\mathrm{graph}F}(x^*, g^*)\} \subseteq F^*(x^*, g^*) \subseteq \{g \in G : g^*(g) < \sigma_{\mathrm{graph}F}(x^*, g^*)\},$$

which proves the first part.

Let $g \in F^{**}(x)$ and $x^* \in \mathcal{X}^*$, $g^* \in K_-^\circ\backslash\{0\}$. Then there exists $g_1 \in -F^*(x^*, g^*)$ and $g_2 \in S_{(x^*,g^*)}(x)$ such that $g = g_1 + g_2$. Equation (7.5) gives $g^*(g_1) \leq \sigma_{\mathrm{graph}F}(x^*, g^*)$ and the definition of $S_{(x^*,g^*)}$ gives $x^*(x) + g^*(g_2) \leq 0$. Hence,

$$x^*(x) + g^*(g) = g^*(g_1) + x^*(x) + g^*(g_2) \leq \sigma_{\mathrm{graph}F}(x^*, g^*).$$

Conversely, let $x^*(x) + g^*(g) \leq \sigma_{\mathrm{graph}F}(x^*, g^*)$ for $g \in G$, $x \in \mathcal{X}$, $x^* \in \mathcal{X}^*$ and $g^* \in K_-^\circ\backslash\{0\}$. As g^* is surjective, there exists $g_2 \in G$ such that $x^*(x) + g^*(g_2) \leq 0$. Thus, $x^*(x) + g^*(g - g_2) \leq \sigma_{\mathrm{graph}F}(x^*, g^*)$. □

The next three propositions give some of the important properties of the conjugate and biconjugate functions.

Proposition 7.1.13.

(i) The conjugate $-F^*$ of F is a well-defined map from $\mathcal{X}^* \times K_-^\circ$ into $\mathcal{K}_c$. Moreover, $-F^*(x^*, g^*)$ is of the form $\{g\} + S_{(x^*,g^*)}(0)$ for some $g \in G$ or it is an element of $\{G, \emptyset\}$.

(ii) The biconjugate F^{**} of F is a well-defined map from $\mathcal{X}$ into $\mathcal{K}_c$.

See [77] for a proof of this proposition.

Proposition 7.1.14 (Set-valued Young-Fenchel's inequalities)**.** *For each $x \in \mathcal{X}$, $x^* \in \mathcal{X}^*$ and $g^* \in K_-^\circ\backslash\{0\}$, we have*

(i) $F(x) \oplus S_{(x^*,g^*)}(-x) \subseteq -F^*(x^*, g^*)$; and

(ii) $F^{**}(x) \subseteq S_{(x^*,g^*)}(x) \oplus -F^*(x^*, g^*)$.

Proposition 7.1.15. *Let $F_1, F_2, F : \mathcal{X} \to \mathcal{K}$. Then* *(i)* $F_1 \subseteq F_2 \Rightarrow -F_1^* \subseteq -F_2^* \Rightarrow F_1^{**} \subseteq F_2^{**}$; *and*

(ii) $F \subseteq F^{**}$.

The proofs of Proposition 7.1.14 and Proposition 7.1.15 are consequences of the definitions of $-F^*$ and F^{**}.

The biconjugation theorem is stated next.

Theorem 7.1.16. *A proper function $F : \mathcal{X} \to \mathcal{K}_c$ is closed and convex if and only if $F^{**}(x) = F(x)$ for all $x \in \mathcal{X}$.*

Proof. First, let F be a proper closed convex function. We claim that every affine minorant of F is also one of F^{**}. Let $x^* \in \mathcal{X}^*$, $g^* \in K^\circ \backslash \{0\}$ and $g \in G$ such that for all $x \in \mathcal{X}$

$$F(x) \subseteq \{g\} + S_{(x^*,g^*)}(x).$$

Adding $S_{(x^*,g^*)}(-x)$ on both sides and taking the union over all $x \in \mathcal{X}$, we get

$$-F^*(x^*, g^*) \subseteq \{g\}.$$

Once again, adding $S_{(x^*,g^*)}(x)$ to both sides we obtain

$$-F^*(x^*, g^*) + S_{(x^*,g^*)}(x) \subseteq \{g\} + S_{(x^*,g^*)}(x).$$

Then, taking the intersection on both sides over $x^* \in \mathcal{X}^*$, $g^* \in K^\circ \backslash \{0\}$, results in

$$F^{**}(x) \subseteq \{g\} + S_{(x^*,g^*)}(x),$$

proving the claim.

Corollary 7.1.10 ensures that the set of proper affine minorants of F is nonempty and Theorem 7.1.9 shows that F is the pointwise supremum of such minorants. Since, by Proposition 7.1.15, we have $F(x) \subseteq F^{**}(x)$ for all $x \in \mathcal{X}$, the result follows.

Conversely, assume $F(x) = F^{**}(x)$ for all $x \in \mathcal{X}$. Then by Proposition 7.1.13, we have that F^{**} is closed and convex. $\square$

7.2 Duality in tensor products

We let $\mathcal{X} = E\widetilde{\otimes}_l Y$, the completed l-tensor product, where E and Y are Banach lattices and let $0 < e \in E$. Let G be a Riesz subspace of Y. Then the Banach lattice $E\widetilde{\otimes}_l G$ may be considered as an l-normed closed Riesz subspace of $E\widetilde{\otimes}_l Y$.

To extend the ordering induced by the cone K_Y to $E\widetilde{\otimes}_l Y$, we define

$$C = \mathrm{cl}_l(E_+ \otimes K_Y),$$

i.e. C is the l-norm closure of

$$E_+ \otimes K_Y = \left\{ x \in E \otimes Y : x = \sum_{i=1}^{n} x_i \otimes y_i \text{ for } n \in \mathbb{N}, x_1, \ldots, x_n \in E_+, y_1, \ldots, y_n \in K_Y \right\}$$

in $E\widetilde{\otimes}_l Y$. Then, C is a closed convex cone in $E\widetilde{\otimes}_l Y$. Note that if $g \in K$, then $e \otimes g \in C$.

The *polar cone* of C is given by

$$C^\circ = \{x^* \in (E\widetilde{\otimes}_l Y)^* : x^*(x) \geq 0 \text{ for all } x \in C\},$$

and the *negative polar cone* of C by $C^\circ_- = -C^\circ$.

Definition 7.2.1. Let $F : E\widetilde{\otimes}_l Y \to \mathcal{K}$ be a function.

(i) If for all $x \in E\widetilde{\otimes}_l Y$ and $g \in G$,

$$F(x + e \otimes g) = F(x) + \{-g\},$$

then F is called *translative* with respect to $0 < e \in E$.

(ii) If for all $x_1, x_2 \in E\widetilde{\otimes}_l Y$,

$$x_2 - x_1 \in C \quad \Rightarrow \quad F(x_1) \subseteq F(x_2),$$

then F is called *monotone* with respect to a convex cone $C \subseteq E\widetilde{\otimes}_l Y$.

Lemma 7.2.2. *Let $g^* \in K^\circ_- \backslash \{0\}$.*

(i) If $x^* \in C^\circ_-$, then the function $S_{(x^*,g^*)} : E\widetilde{\otimes}_l Y \to \mathcal{K}_c$ is monotone with respect to C.

(ii) If $g^*(g) = x^*(e \otimes g)$ for all $g \in G$, then the function $S_{(x^*,g^*)} : E\widetilde{\otimes}_l Y \to \mathcal{K}_c$ is translative with respect to $0 < e \in E$.

Proof. (i) Consider $x_1, x_2 \in E\widetilde{\otimes}_l Y$ such that $x_2 - x_1 \in C$ and $x^* \in C^\circ_-$. Then, $x^*(x_2 - x_1) \leq 0$, which implies that $0 \in S_{(x^*,g^*)}(x_2 - x_1)$. Then,

$$\begin{aligned} S_{(x^*,g^*)}(x_2) &= S_{(x^*,g^*)}(x_2 - x_1 + x_1) \\ &= S_{(x^*,g^*)}(x_2 - x_1) + S_{(x^*,g^*)}(x_1) \\ &\supseteq S_{(x^*,g^*)}(x_1), \end{aligned}$$

proving the monotonicity.

(ii) Assume $x^*(e \otimes g) = g^*(g)$ for all $g \in G$. Consider $x \in E\widetilde{\otimes}_l Y$ and $g \in G$. Then

$$\begin{aligned} x^*(x + e \otimes g) + g^*(u) &= x^*(x) + x^*(e \otimes g) + g^*(u) \\ &= x^*(x) + g^*(g + u). \end{aligned}$$

Hence,

$$\begin{aligned} S_{(x^*,g^*)}(x + e \otimes g) &= \{u \in G : x^*(x) + g^*(g + u) \leq 0\} \\ &= \{v - g \in G : x^*(x) + g^*(v) \leq 0\} \\ &= \{v \in G : x^*(x) + g^*(v) \leq 0\} + \{-g\} \\ &= S_{(x^*,g^*)}(x) + \{-g\}, \end{aligned}$$

which proves the translativity. □

Define the zero-sublevel set of $F : \mathcal{X} \to \mathcal{K}$ by

$$A_F = \{x \in \mathcal{X} : 0 \in F(x)\} = \{x \in \mathcal{X} : K \subseteq F(x)\}.$$

Note that if F is monotone and $0 \in F(0)$, then $C \subseteq A_F$.

The positive polar cone of A_F is given by

$$A_F^+ = \{x^* \in \mathcal{X}^* : x^*(x) \geq 0 \text{ for all } x \in A_F\}$$

and the negative polar cone by $A_F^- = -A_F^+$.

Proposition 7.2.3.

(i) If F is convex, then A_F is convex.

(ii) If F is positive homogeneous, then A_F is a cone.

Proof. (i) Assume that F is convex. Take $x, y \in A_F$. This implies that $0 \in F(x)$ and $0 \in F(y)$. Since F is convex, we have for $\lambda \in (0, 1)$

$$\lambda F(x) + (1 - \lambda)F(y) \subseteq F(\lambda x + (1 - \lambda)y).$$

Hence, $0 \in F(\lambda x + (1 - \lambda)y)$. Thus, $\lambda x + (1 - \lambda)y \in A_F$, showing that A_F is convex.

(ii) Assume that F is positive homogeneous, i.e. $F(\lambda x) = \lambda F(x)$ for all $\lambda > 0$. Thus, if $0 \in F(x)$, then it must be that $0 \in F(\lambda x)$. In other words, for all $x \in A_F$, we have that $\lambda x \in A_F$ for $\lambda > 0$. This proves that A_F is a cone. □

The next theorem states the representation result for proper, closed, convex functions.

Theorem 7.2.4.

(i) Let $F : E\widetilde{\otimes}_l Y \to \mathcal{K}_c$ *be proper, closed, convex, translative with respect to* $e \in E$*, monotone with respect to* C *and* $0 \in F(0)$*. Then*

$$-F^*(x^*, g^*) = \begin{cases} \mathrm{cl} \bigcup_{x \in A_F} S_{(x^*,g^*)}(-x) & \text{if } (x^*, g^*) \in \mathcal{Y} \\ G & \text{if } (x^*, g^*) \notin \mathcal{Y}, \end{cases} \tag{7.7}$$

where

$$\mathcal{Y} = \{(x^*, g^*) \in C_-^\circ \times K_-^\circ : x^*(e \otimes g) = g^*(g) \text{ for all } g \in G\}.$$

(ii) The dual representation of F *is*

$$F(x) = \bigcap_{(x^*,g^*) \in \mathcal{Y}} \left[S_{(x^*,g^*)}(x) - \mathrm{cl} \bigcup_{x \in A_F} S_{(x^*,g^*)}(-x) \right] \tag{7.8}$$

for all $x \in E\widetilde{\otimes}_l Y$*.*

Proof. Our proof is based on [79, p. 27, Theorem 6.2].

(i) First, note that for $x^* \in (E\widetilde{\otimes}_l Y)^*$ and $g^* \in K_-^\circ \backslash \{0\}$, we have

$$\mathrm{cl} \bigcup_{x \in A_F} S_{(x^*,g^*)}(-x) \subseteq \mathrm{cl} \bigcup_{x \in A_F} [F(x) + S_{(x^*,g^*)}(-x)] \subseteq -F^*(x^*, g^*).$$

For the converse inclusion, consider $(x^*, g^*) \in \mathcal{Y}$, $x \in \mathrm{dom} F$ and $g \in F(x)$. Then

$$S_{(x^*,g^*)}(-x) + \{g\} = S_{(x^*,g^*)}(-x - e \otimes g).$$

We claim that

$$S_{(x^*,g^*)}(-x - e \otimes g) \subseteq \mathrm{cl} \bigcup_{x \in A_F} S_{(x^*,g^*)}(-x). \tag{7.9}$$

To prove this inclusion, consider $u \in S_{(x^*,g^*)}(-x - e \otimes g)$. Thus, $x^*(-x - e \otimes g) + g^*(u) \leq 0$. Let $y = x + e \otimes g$. Since $g \in F(x)$, we have that $0 \in F(x) + \{-g\}$. As $F(x) + \{-g\} = F(x + e \otimes g)$, we have that $y = x + e \otimes g \in A_F$ and $x^*(-y) + g^*(u) \leq 0$. Hence,

$$u \in \bigcup_{x \in A_F} S_{(x^*,g^*)}(-x),$$

which proves our claim. As (7.9) holds for all $g \in F(x)$, we have

$$S_{(x^*,g^*)}(-x) + F(x) \subseteq \mathrm{cl} \bigcup_{x \in A_F} S_{(x^*,g^*)}(-x).$$

Hence,

$$-F^*(x^*, g^*) \subseteq \operatorname{cl} \bigcup_{x \in A_F} S_{(x^*,g^*)}(-x)$$

whenever $(x^*, g^*) \in \mathcal{Y}$.

It remains to show that if $(x^*, g^*) \notin \mathcal{Y}$, then $-F^*(x^*, g^*) = G$. We will show the converse. Consider $x^* \in \mathcal{X}^*$ and $g^* \in K_-^\circ \backslash \{0\}$. Assume $-F^*(x^*, g^*) \neq G$. By Proposition 7.1.13, $-F^*(x^*, g^*) \subseteq \{g\} + S_{(x^*,g^*)}(0)$ for some $g \in G$. Hence,

$$-\infty < \sup_{g_0 \in -F^*(x^*,g^*)} g^*(g_0) \leq g^*(g) < \infty.$$

Note that because F is proper, we know that $-F^*(x^*, g^*) \neq \emptyset$.

Since F is translative, we have for arbitrary $w \in G$,

$$\begin{aligned} -F^*(x^*, g^*) &= \operatorname{cl} \bigcup_{x \in E \widetilde{\otimes}_l Y} [F(x) + S_{(x^*,g^*)}(-x)] \\ &= \operatorname{cl} \bigcup_{x \in E \widetilde{\otimes}_l Y} \Big[F(x + e \otimes w) + S_{(x^*,g^*)}(-x - e \otimes w)\Big] \\ &\subseteq -F^*(x^*, g^*) + \{-w\} + S_{(x^*,g^*)}(-e \otimes w). \end{aligned}$$

This results in

$$\sup_{g_0 \in -F^*(x^*,g^*)} g^*(g_0) \leq \sup_{g_0 \in -F^*(x^*,g^*)} g^*(g_0) - g^*(w) + x^*(e \otimes w).$$

Hence, $x^*(e \otimes w) - g^*(w) \geq 0$ for all $w \in G$. This is only possible, if $g^*(w) = x^*(e \otimes w)$.

As F is C-monotone and $0 \in F(0)$, we have that $C \subseteq A_F$.

By the Young-Fenchel inequality, the definition of A_F and the fact that $C \in A_F$, we have for all $x \in C$,

$$S_{(x^*,g^*)}(-x) \subseteq F(x) + S_{(x^*,g^*)}(-x) \subseteq -F^*(x^*, g^*).$$

Since C is a cone, if $x \in C$, then $tx \in C$ for all $t > 0$. Thus,

$$S_{(x^*,g^*)}(-tx) \subseteq -F^*(x^*, g^*).$$

Applying g^* to this inclusion, we get for all $t > 0$,

$$tx^*(x) \leq \sup_{g_0 \in -F^*(x^*,g^*)} g^*(g_0) \in \mathbb{R}.$$

The Archimedean property of $\mathbb{R}$ implies that $x^*(x) \leq 0$. Since, x is an arbitrary element in C, we have $x^* \in C_-^\circ$. Putting all this together, we get $(x^*, g^*) \in \mathcal{Y}$.

(ii) Finally, the dual representation formula (7.8) is a consequence of Equation (7.7) and Theorem 7.1.16. □

Theorem 7.2.4 can be specialised to functions that are also positively homogeneous. This is stated in the next theorem.

Theorem 7.2.5.

(i) Let $F : E\widetilde{\otimes}_l Y \to \mathcal{K}_c$ be proper, closed, convex, translative with respect to $e \in E$, monotone with respect to C, $0 \in F(0)$ and positively homogeneous, then

$$-F^*(x^*, g^*) = \begin{cases} S_{(x^*,g^*)}(0) & \text{if } (x^*, g^*) \in \mathcal{Y} \text{ and } x^* \in A_F^- \\ G & \text{if } (x^*, g^*) \notin \mathcal{Y} \text{ or } x^* \notin A_F^-. \end{cases} \tag{7.10}$$

(ii) The dual representation of F is

$$F(x) = \bigcap_{(x^*,g^*)\in\mathcal{Y},\ x^*\in A_F^-} S_{(x^*,g^*)}(x) \tag{7.11}$$

for all $x \in E\widetilde{\otimes}_l Y$.

Proof. Our proof is based on [79, p. 27, Theorem 6.2].

(i) For the case where F is also positive homogeneous, if $(x^*, g^*) \in \mathcal{Y}$, we have by Theorem 7.2.4, that $-F^*(x^*, g^*) = \operatorname{cl}\bigcup_{x\in A_F} S_{(x^*,g^*)}(-x)$. Since $0 \in A_F$, we have $S_{(x^*,g^*)}(0) \subseteq \operatorname{cl}\bigcup_{x\in A_F} S_{(x^*,g^*)}(-x)$.

On the other hand, if $x^* \in A_F^-$, then $x^*(x) \le 0$ for all $x \in A_F$. Hence,

$$g \in S_{(x^*,g^*)}(-x) = \{u \in G : g^*(u) \le x^*(x)\}$$

implies $g \in S_{(x^*,g^*)}(0)$. This gives $\operatorname{cl}\bigcup_{x\in A_F} S_{(x^*,g^*)}(-x) \subseteq S_{(x^*,g^*)}(0)$.

Now, assume $x^* \notin A_F^-$. Then, there exists $x \in A_F$ such that $x^*(x) > 0$. Because F is positively homogeneous, we have by Proposition 7.2.3 that A_F is a cone. Therefore, $tx \in A_F$ for $t > 0$. Thus,

$$\operatorname{cl}\bigcup_{x\in A_F} S_{(x^*,g^*)}(-x) \supseteq \bigcup_{t>0}\{g \in G : g^*(g) \le tx^*(x)\} = G.$$

Hence, $\operatorname{cl}\bigcup_{x\in A_F} S_{(x^*,g^*)}(-x) = G$, whenever $x^* \notin A_F^-$.

(ii) The dual representation formula (7.11) is a direct consequence of Equation (7.10) and Theorem 7.1.16. □

7.3 Risk measures on the l-tensor product

Definition 7.3.1. Let $0 < e \in E$. A function $\rho_e : E\widetilde{\otimes}_l Y \to \mathcal{K}$ is called a *set-valued monetary e-risk measure* if it satisfies

(R0) Normalisation, i.e. $K \subseteq \rho_e(0)$ and $\rho_e(0) \cap (-\mathrm{int}K) = \emptyset$;

(R1) C-monotonicity, i.e. $x - y \in C$ implies $\rho_e(y) \subseteq \rho_e(x)$;

(R2) Cash additivity (also called translation invariance), i.e. $\rho_e(x + e \otimes u) = \rho_e(x) + \{-u\}$ for all $x \in E\widetilde{\otimes}_l Y$ and $u \in G$.

If, in addition, ρ_e satisfies

(R3) Convexity, i.e. $\lambda\rho_e(x) + (1-\lambda)\rho_e(y) \subseteq \rho_e(\lambda x + (1-\lambda)y)$ for $\lambda \in (0, 1)$,

then ρ_e is called a *set-valued convex e-risk measure.*
If ρ_e satisfies (R0), (R1) and (R2) and

(R4) Positive homogeneity, i.e. $\rho_e(cx) = c\rho_e(x)$ for all $x \in E\widetilde{\otimes}_l Y$ and $c > 0$;

(R5) Subadditivity, i.e. $\rho_e(x_1) + \rho_e(x_2) \subseteq \rho_e(x_1 + x_2)$ for all $x_1, x_2 \in E\widetilde{\otimes}_l Y$;

then ρ_e is called a *set-valued coherent e-risk measure.*

If $E = L_p(P)$, where $1 \le p \le \infty$, and $e = \mathbf{1} : \Omega \to \mathbb{R}$, is defined by $\mathbf{1}(\omega) = 1$ a.e., we denote $\rho_{\mathbf{1}}$ by ρ. The properties of $\rho_{\mathbf{1}}$ can be interpreted as in the scalar-valued case, see for example [60].

For the definition of the acceptance sets, we require some form of closedness. Hamel et al. [78] introduced the following notion. A set $A \in E\widetilde{\otimes}_l Y$ is *directionally closed* with respect to $0 < e \in E$, if for any $x \in E\widetilde{\otimes}_l Y$ and any sequence $(u_k)_{k\in\mathbb{N}} \subseteq G$ with $\lim\limits_{k\to\infty} u_k = 0$, it follows from $x + e \otimes u_k \in A$ for all $k \in \mathbb{N}$, that $x \in A$.

Definition 7.3.2. Let $0 < e \in E$.

(i) An *e-acceptance set* is a set $A_e \subseteq E\widetilde{\otimes}_l Y$ which satisfies

- **(A0)** $u \in K$ implies $e \otimes u \in A_e$ and $u \in -\mathrm{int}K$ implies $e \otimes u \notin A_e$;
- **(A1)** A_e is directionally closed with respect to e; and
- **(A2)** $A_e + C \subseteq A_e$.

(ii) If, in addition, A_e is convex, then A_e is called a *convex e-acceptance set.*

(iii) If A_e satisfies (A0), (A1), (A2) and is a convex cone, then A_e is called a *coherent e-acceptance set.*

Note that (A2) and the definition of C imply that $A_e + e \otimes u \in A_e$ for $u \in K$. Moreover, by (A0), we have $0 \in A_e$ and hence, by (A2), $C \subseteq A_e$.

The next two theorems show the one-to-one relationship between set-valued risk measures and acceptance sets.

Theorem 7.3.3.

(i) Let $\rho_e : E\widetilde{\otimes}_l Y \to \mathcal{K}$ *be a monetary e-risk measure. Then*

$$A_{\rho_e} = \{x \in E\widetilde{\otimes}_l Y : 0 \in \rho_e(x)\}$$

is an acceptance set.

(ii) If ρ_e *is convex, then so is* A_{ρ_e}.

(iii) If ρ_e *is coherent, then* A_{ρ_e} *is a coherent e-acceptance set.*

Proof. (i) First we show (A0). Let $u \in K$. Then by (R0), $u \in \rho_e(0)$. By (R2), $\rho_e(e \otimes u) = \rho_e(0 + e \otimes u) = \rho_e(0) + \{-u\}$. Since $u \in \rho_e(0)$, it must be that $0 \in \rho_e(e \otimes u)$. Hence, $e \otimes u \in A_{\rho_e}$. Next, let $u \in -\text{int}K$. Then by (R0), $u \notin \rho_e(0)$. Similarly as above, by (R2), $0 \notin \rho_e(e \otimes u)$ and $e \otimes u \notin A_{\rho_e}$.

Secondly, we show (A1). Let $x \in E\widetilde{\otimes}_l Y$, $(u_k)_{k\in\mathbb{N}} \subseteq G$ with $\lim_{k\to\infty} u_k = 0$ and $x + e \otimes u_k \in A_{\rho_e}$. Then, by the definition of A_{ρ_e} and (R1), we have $0 \in \rho_e(x + e \otimes u_k) = \rho_e(x) + \{-u_k\}$. In other words, $u_k \in \rho_e(x)$ for all $k \in \mathbb{N}$. Since, ρ_e maps into $\mathcal{K}$, $\rho_e(x)$ is closed. Hence, $0 \in \rho_e(x)$, implying that $x \in A_{\rho_e}$. Thus, A_{ρ_e} is radially closed with respect to e.

Lastly, we check (A2). Let $x_1 \in A_{\rho_e}$ and $x_2 \in C$. Then $(x_1 + x_2) - x_1 \in C$ and by (R1) $0 \in \rho_e(x_1) \subseteq \rho_e(x_1 + x_2)$. Thus, $x_1 + x_2 \in A_{\rho_e}$, as desired.

(ii) This was proved in Proposition 7.2.3.

(iii) If ρ_e is coherent, then by Proposition 7.2.3, we have that A_{ρ_e} is a cone. As positive homogeneity and subadditivity imply convexity, we have by (ii), that A_{ρ_e} is convex. □

Theorem 7.3.4.

(i) Let $A_e \subseteq E\widetilde{\otimes}_l Y$ *be an e-acceptance set. Then* ρ_{A_e}*, given by*

$$\rho_{A_e}(x) = \{u \in G : x + e \otimes u \in A_e\},$$

is a monetary risk measure.

(ii) If A_e *is convex, then so is* ρ_{A_e}.

(iii) If A_e is a coherent acceptance set, then ρ_{A_e} is a coherent e-risk measure.

Proof. (i) First, we show that ρ_{A_e} maps into $\mathcal{K}$, i.e. we need to show that $\rho_{A_e}(x) + K \subseteq \rho_{A_e}(x)$ and $\rho_{A_e}(X)$ is closed for each $x \in E\widetilde{\otimes}_l Y$. Let $u \in \rho_{A_e}(x)$ and $v \in K$. Then, by the definition of ρ_{A_e}, $X + e \otimes u \in A_e$. By (A2), we have $x + e \otimes (u + v) \in A_e$, and hence, $u + v \in \rho_{A_e}(x)$. Consider a sequence $(u_k)_{k\in\mathbb{N}} \subseteq \rho_{A_e}(x)$ with $\lim_{k\to\infty} u_k = u$. By the definition of ρ_{A_e},

$$x + e \otimes u_k = (x + e \otimes u) + e \otimes (u_k - u) \in A_e,$$

for all $k \in \mathbb{N}$. Since A_e is directionally closed, this implies that $x + e \otimes u \in A_e$, which gives $u \in \rho_{A_e}(x)$. This proves that $\rho_{A_e}(x)$ is closed for each $X \in E\widetilde{\otimes}_l Y$.

Now, we can show (R0). By (A0) and the definition of ρ_{A_e}, we get $u \in \rho_{A_e}(0)$ if $u \in K$ and $u \notin \rho_{A_e}(0)$ if $u \in -\text{int}K$.

Next, we need to show (R1). Let $x_1, x_2 \in E\widetilde{\otimes}_l Y$ such that $x_2 - x_1 \in C$. Consider $u \in \rho_{A_e}(x_1)$. Then, $x_1 + e \otimes u \in A_e$. As $A_e + C \subseteq A_e$, we have that

$$x_1 + e \otimes u + (x_2 - x_1) \in A_e.$$

This implies that $x_2 + e \otimes u \in A_e$, i.e. $u \in \rho_{A_e}(x_2)$. Therefore, $\rho_{A_e}(x_1) \subseteq \rho_{A_e}(x_2)$.

Lastly, let us show (R2). Let $X \in E\widetilde{\otimes}_l Y$ and $u \in G$. By the definition of ρ_{A_e}, we have

$$\begin{aligned}\rho_{A_e}(x + e \otimes u) &= \{v \in G : x + e \otimes u + e \otimes v \in A_e\}\\ &= \{v \in G : x + e \otimes (u + v) \in A_e\}\\ &= \{w \in G : x + e \otimes w \in A_e\} + \{-u\}\\ &= \rho_{A_e}(x) + \{-u\}.\end{aligned}$$

(ii) Now, assume A_e is convex. Let $\lambda \in (0, 1)$ and $u \in \lambda\rho_{A_e}(x) + (1 - \lambda)\rho_{A_e}(y)$. Then, we can write u as $u = \lambda u_1 + (1 - \lambda)u_2$ for some $u_1 \in \rho_{A_e}(x)$ and $u_2 \in \rho_{A_e}(y)$. Therefore, $x + e \otimes u_1 \in A_e$ and $y + e \otimes u_2 \in A_e$. As A_e is convex, we have

$$\lambda(x + e \otimes u_1) + (1 - \lambda)(y + e \otimes u_2) \in A_e.$$

But,

$$\begin{aligned}&\lambda(x + e \otimes u_1) + (1 - \lambda)(y + e \otimes u_2)\\ &\qquad = \lambda x + (1 - \lambda)y + \lambda(e \otimes u_1) + (1 - \lambda)(e \otimes u_2)\\ &\qquad = \lambda x + (1 - \lambda)y + e \otimes (\lambda u_1 + (1 - \lambda)u_2)\\ &\qquad = \lambda x + (1 - \lambda)y + e \otimes u.\end{aligned}$$

Hence, $\lambda x + (1-\lambda)y + e \otimes u \in A_e$, thus, $u \in \rho_{A_e}(\lambda x + (1-\lambda)y)$, proving that ρ_{A_e} is convex.

(iii) If A_e is a coherent acceptance set, then A_e is a convex cone. By (ii), we have that ρ_{A_e} is convex. Consider $x, y \in E\widetilde{\otimes}_l Y$ and $u \in \rho_{A_e}(x) + \rho_{A_e}(y)$. Then, $u = u_1 + u_2$, where $u_1 \in \rho_{A_e}(x)$ and $u_2 \in \rho_{A_e}(y)$. This means that $x + e \otimes u_1 \in A_e$ and $y + e \otimes u_2 \in A_e$. Since A_e is a cone, we have that

$$x + e \otimes u_1 + y + e \otimes u_2 = x + y + e \otimes u \in A_e.$$

Thus, $u \in \rho_{A_e}(x+y)$, proving that ρ_{A_e} is subadditive. Using the subadditivity repeatedly, we can show that ρ_{A_e} is positive homogeneous and hence coherent. □

Hamel and Heyde [78] consider a set-valued expectation and a set-valued upper expectation in the setting of $L_p(P, \mathbb{R}^d)$. We extend these notions to the framework under consideration.

A *dual pair* is a triple $(U, V, \langle \cdot, \cdot \rangle)$, consisting of vector spaces U and V and a bilinear mapping $\langle \cdot, \cdot \rangle : U \times V \to \mathbb{R}$ such that

- $\langle u, v \rangle = 0$ for all $u \in U$ implies $v = 0$, and
- $\langle u, v \rangle = 0$ for all $v \in V$ implies $u = 0$.

We recall from [142] that if E is a Banach lattice and $0 < e \in E$, then e is a *quasi interior point* of E provided that $\mathrm{cl}_E\{y \in E : |y| \leq Me \text{ for some } M \in \mathbb{R}_+\} = E$.

It is well known that $\mathbf{1}$ is a quasi interior point of $L_p(P)$ for $1 \leq p < \infty$.

Definition 7.3.5. Consider the dual pair $(E\widetilde{\otimes}_l Y, (E\widetilde{\otimes}_l Y)^*, \langle \cdot, \cdot \rangle)$, where $\langle x, x^* \rangle = x^*(x)$, and let $0 < e \in E$. For all $x^* \in (E\widetilde{\otimes}_l Y)^*$, define $\mathbb{E}^{x^*}_{G,e}$ and $\mathbb{F}^{x^*}_{G,e}$ respectively, for all $x \in E\widetilde{\otimes}_l Y$, by

$$\mathbb{E}^{x^*}_{G,e}(x) = \{g \in G : \langle x - e \otimes g, x^* \rangle = 0\}$$

and

$$\mathbb{F}^{x^*}_{G,e}(x) = \{g \in G : \langle x - e \otimes g, x^* \rangle \leq 0\}.$$

In the case $E = L_p(P)$ and $e = \mathbf{1} : \Omega \to \mathbb{R}$, defined by $\mathbf{1}(\omega) = 1$ a.e., we denote $\mathbb{F}^{x^*}_{G,\mathbf{1}}$ by $\mathbb{F}^{x^*}_{G}$.

We can now state the dual representation result first for set-valued coherent e-risk measures defined on the completed l-tensor product.

Theorem 7.3.6. *Let $\rho_e : E\widetilde{\otimes}_l Y \to \mathcal{K}_c$ be proper and closed. The function ρ_e is a coherent e-risk measure if and only if, for all $x \in E\widetilde{\otimes}_l Y$,*

$$\rho_e(x) = \bigcap_{Z^* \in C^\circ \cap A^+_{\rho_e}} \mathbb{F}^{Z^*}_{G,e}[-x]. \tag{7.12}$$

Proof. Let $\rho_e : E\widetilde{\otimes}_l Y \to \mathcal{K}_c$ be a proper closed coherent e-risk measure.

If $Z^* \in (E\widetilde{\otimes}_l Y)^*$, define a function $z^* : G \to \mathbb{R}$ by

$$z^*(g) = Z^*(e \otimes g) \text{ for all } g \in G.$$

Then $z^* = Z^* \circ j$, where $j : G \to E\widetilde{\otimes}_l Y$ is the canonical embedding, given by $j(g) = e \otimes g$. As $j(K) \subseteq C$, it follows that $Z^* \in C^\circ$ implies $z^* \in K^\circ$.

Since $(-Z^*, -z^*) \in \mathcal{Y}$ for $Z^* \in C^\circ$, Theorem 7.2.5 implies that for all $x \in E\widetilde{\otimes}_l Y$,

$$\rho_e(x) = \bigcap_{Z^* \in C^\circ \cap A^+_{\rho_e}} S_{(-Z^*,-z^*)}(x).$$

Furthermore, if $Z^* \in C^\circ$, then as already noted, $z^* \in K^\circ$, and we get

$$\begin{aligned} S_{(-Z^*,-z^*)}(x) &= \Big\{ g \in G : \langle x, -Z^* \rangle - \langle g, z^* \rangle \leq 0 \Big\} \\ &= \Big\{ g \in G : \langle -x, Z^* \rangle - \langle e \otimes g, Z^* \rangle \leq 0 \Big\} \\ &= \Big\{ g \in G : \langle -x - e \otimes g, Z^* \rangle \leq 0 \Big\} \\ &= \mathbb{F}^{Z^*}_{G,e}[-x], \end{aligned}$$

proving the representation.

Conversely, we need to show that (7.12) is a coherent e-risk measure.

First, we show that (R0) holds. Note that

$$\mathbb{F}^{Z^*}_{G,e}[0] = S_{(-Z^*,-z^*)}(0) = \{ g \in G : z^*(g) \geq 0 \}.$$

As $z^* \in K^\circ$, we have that $z^*(k) \geq 0$ for all $k \in K$. Hence, $K \subseteq \mathbb{F}^{Z^*}_{G,e}[0]$ for all $Z^* \in C^\circ \cap A_{\rho_e^+}$. Therefore, $K \subseteq \rho_e(0)$.

Assume $\rho_e(0) \cap (-\text{int}K) \neq \emptyset$. Select $g \in G$ such that $g \in \rho_e(0)$ and $g \in -\text{int}K$. From $g \in -\text{int}K$ and $z^* \in K^\circ$, we get $z^*(g) \leq 0$. From $g \in \rho_e(0)$, we get

$$g \in \mathbb{F}^{Z^*}_{G,e}[0] \text{ for all } Z^* \in C^\circ \cap A^+_{\rho_e}.$$

Thus, $z^*(g) \geq 0$, a contradiction. In other words, we must have $\rho_e(0) \cap (-\text{int}K) = \emptyset$.

Next, note that

$$\begin{aligned}\mathbb{F}^{Z^*}_{G,e}[-x-e\otimes g] &= \{u\in G : -Z^*(x+e\otimes g)-z^*(u)\le 0\}\\ &= \{u\in G : -Z^*(x)-z^*(u+g)\le 0\}\\ &= \{v-g\in G : -Z^*(x)-z^*(v)\le 0\}\\ &= \mathbb{F}^{Z^*}_{G,e}[-x]+\{-g\}.\end{aligned}$$

Hence, $\rho_e(x+e\otimes g)=\rho_e(x)+\{-g\}$, proving translativity with respect to e.

Lastly, consider $x,y\in E\widetilde{\otimes}_l Y$ such that $x-y\in C$. Let $g\in \mathbb{F}^{Z^*}_{G,e}(-y)$. Then

$$-Z^*(y)-z^*(g)\le 0.$$

Hence, $-Z^*(x)+Z^*(x-y)-z^*(g)\le 0$ and

$$-Z^*(x)-z^*(g)\le -Z^*(x-y)\le 0 \text{ as } Z^*\in C^\circ,$$

showing that $g\in \mathbb{F}^{Z^*}_{G,e}(-x)$, i.e. $\mathbb{F}^{Z^*}_{G,e}(-y)\subseteq \mathbb{F}^{Z^*}_{G,e}(-x)$ for all $Z^*\in C^\circ$. Therefore, $\rho_e(y)\subseteq \rho_e(x)$, proving the C-monotonicity.

Using the properties of $S_{(-Z^*,-z^*)}$, we know that

$$\mathbb{F}^{Z^*}_{G,e}[tx]=t\mathbb{F}^{Z^*}_{G,e}[x] \text{ for } t\neq 0$$

and

$$\mathbb{F}^{Z^*}_{G,e}[x_1+x_2]=\mathbb{F}^{Z^*}_{G,e}[x_1]+\mathbb{F}^{Z^*}_{G,e}[x_2].$$

Hence, ρ_e is both positive homogeneous and subadditive. □

Next, we state and prove the representation of set-valued convex e-risk measures defined on the completed l-tensor product.

Theorem 7.3.7. *Let $\rho_e : E\widetilde{\otimes}_l Y\to \mathcal{K}_c$ be proper. If ρ_e is a closed convex e-risk measure, then for all $x\in E\widetilde{\otimes}_l Y$,*

$$\rho_e(x)=\bigcap_{Z^*\in C^\circ}\Big[\mathbb{F}^{Z^*}_{G,e}[-x]+\alpha_{min}(Z^*)\Big], \tag{7.13}$$

where the penalty function $\alpha_{min}: C^\circ\to G$ is given by

$$\alpha_{min}(Z^*)=\operatorname{cl}\bigcup_{x_0\in A_{\rho_e}}\mathbb{F}^{Z^*}_{G,e}[x_0] \ \text{ for } Z^*\in C^\circ.$$

Moreover, α_{min} is the minimal penalty function that represents ρ_e, i.e. any penalty function α for which (7.13) holds, satisfies $\alpha_{min}(Z^*)\subseteq \alpha(Z^*)$ for all $Z^*\in C^\circ$.

Proof. (i): Let $\rho_e : E\widetilde{\otimes}_l Y \to \mathcal{K}_c$ be a proper closed convex e-risk measure. Define z^* as in the proof of Theorem 7.3.6.

Then, Theorem 7.2.4 implies that for all $x \in E\widetilde{\otimes}_l Y$,

$$\rho_e(x) = \bigcap_{Z^* \in C^\circ} \left[S_{(-Z^*,-z^*)}(x) + \operatorname{cl} \bigcup_{x_0 \in A_{\rho_e}} S_{(-Z^*,-z^*)}(-x_0) \right].$$

We also have by the proof of Theorem 7.3.6, that $S_{(-Z^*,-z^*)}(x) = \mathbb{F}^{Z^*}_{G,e}[-x]$.

Let

$$\alpha_{min}(Z^*) = \operatorname{cl} \bigcup_{x_0 \in A_{\rho_e}} \mathbb{F}^{Z^*}_{G,e}[x_0],$$

giving us the required representation.

Finally, let α be any penalty function for ρ_e. Then, for all $Z^* \in C^\circ$, we have that

$$\rho_e(x) \subseteq \mathbb{F}^{Z^*}_{G,e}[-x] + \alpha(Z^*).$$

Hence,

$$\alpha(Z^*) \supseteq \rho_e(x) + \mathbb{F}^{Z^*}_{G,e}[x] \text{ for all } x \in E\widetilde{\otimes}_l Y,$$

i.e.

$$\begin{aligned}
\alpha(Z^*) &\supseteq \operatorname{cl} \bigcup_{x \in E\widetilde{\otimes}_l Y} \left[\rho_e(x) + \mathbb{F}^{Z^*}_{G,e}[x] \right] \\
&\supseteq \operatorname{cl} \bigcup_{x \in \mathcal{A}_{\rho_e}} \left[\rho_e(x) + \mathbb{F}^{Z^*}_{G,e}[x] \right] \\
&\supseteq \operatorname{cl} \bigcup_{x \in \mathcal{A}_{\rho_e}} \mathbb{F}^{Z^*}_{G,e}[x] = \alpha_{min}(Z^*),
\end{aligned}$$

as required. □

Next, we present applications of our main result, Theorem 7.3.7.

Corollary 7.3.8. *Suppose that Y^* has the Radon-Nikodým property.*

(i) Let $1 \le p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ and $\rho_e : L_p(P,Y) \to \mathcal{K}_c$. In addition to the descriptions of ρ_e in Theorem 7.3.7, C° and $\mathbb{F}^{Z^*}_{G,e}$ are respectively, given by

$$C^\circ = \{Z^* \in L_q(P,Y^*) : Z^*(c) \ge 0 \text{ for all } c \in C\}$$

and, for all $x \in L_p(P,Y)$,

$$\mathbb{F}^{Z^*}_{G,e}[x] = \left\{ g \in G : \int_\Omega x(\omega) Z^*(\omega)\, dP \le z^*(g) \right\},$$

where $z^* = Z^* \circ j$ and $j : G \to E\widetilde{\otimes}_l Y$ is the canonical embedding, given by $j(g) = e \otimes g$.

(ii) *Let* (Ψ, Ψ^*) *be complementary finite Young functions and* $\rho_e : H_\Psi(P,Y) \to \mathcal{K}_c$*. In addition to the descriptions of* ρ_e *in Theorem 7.3.7,* C° *and* $\mathbb{F}^{Z^*}_{G,e}$ *are respectively, given by*

$$C^\circ = \{Z^* \in L_{\Psi^*}(P)\widetilde{\otimes}_l Y^* : Z^*(c) \geq 0 \text{ for all } c \in C\}$$

and, for all $x \in H_\Psi(P,Y)$*,*

$$\mathbb{F}^{Z^*}_{G,e}[x] = \left\{ g \in G : \int_\Omega x(\omega)Z^*(\omega)\,dP \leq z^*(g) \right\},$$

where $z^* = Z^* \circ j$ *and* $j : G \to E\widetilde{\otimes}_l Y$ *is the canonical embedding, given by* $j(g) = e\otimes g$*.*

(iii) *Let* E *be a separable Banach lattice with order continuous norm and a quasi interior point* $0 < e \in E$ *and* $\rho_e : E\widetilde{\otimes}_l Y \to \mathcal{K}_c$*. In addition to the descriptions of* ρ_e *in Theorem 7.3.7,* C° *and* $\mathbb{F}^{Z^*}_{G}$ *are respectively, given by*

$$C^\circ = \{Z^* \in E^*\widetilde{\otimes}_l Y^* : Z^*(c) \geq 0 \text{ for all } c \in C\}$$

and, for all $x \in E\widetilde{\otimes}_l Y$*,*

$$\mathbb{F}^{Z^*}_{G,e}[x] = \left\{ g \in G : \sum_{i,j} x_i^*(x_j)y_i^*(y_j) \leq \sum_i x_i^*(e)y_i^*(g) \right\}.$$

where $x = \sum_j x_j \otimes y_j$*, for some sequences* $(x_i) \subseteq E$ *and* $(y_i) \subseteq F$ *such that* $\| \sum_i |x_j| \|_E < \infty$ *and* $\lim_{n\to\infty} \|y_j\|_Y = 0$*, and* $Z^* = \sum_j x_j^* \otimes y_j^*$*, for some sequences* $(x_i^*) \subseteq E^*$ *and* $(y_i^*) \subseteq Y^*$ *such that* $\| \sum_i |x_j^*| \|_{E^*} < \infty$ *and* $\lim_{n\to\infty} \|y_j^*\|_{Y^*} = 0$*.*

Proof. (i) Let Y^* be a Banach space with the Radon-Nikodým property. If $(\Omega, \mathcal{F}, P)$ is a probability space, $1 \leq p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, then $((L_p(P,Y), (L_q(P,Y^*), \langle \cdot, \cdot \rangle)$, where $\langle \cdot, \cdot \rangle : L_p(P,Y) \times L_q(P,Y^*) \to \mathbb{R}$ is given by

$$\langle u, v \rangle = \int_\Omega u(t)v(t)\,dP,$$

is a dual pair (see [52]).

By the respective definitions of C° and $\mathbb{F}^{Z^*}_{G,e}$, it follows that

$$C^\circ = \{Z^* \in L_q(P,Y^*) : Z^*(c) \geq 0 \text{ for all } c \in C\}$$

and, for all $x \in L_p(P,Y)$,

$$\mathbb{F}^{Z^*}_{G,e}[x] = \left\{ g \in G : \int_\Omega x(\omega)Z^*(\omega)\,dP \leq z^*(g) \right\},$$

where $z^* = Z^* \circ j$ and $j : G \to E\widetilde{\otimes}_l Y$ is the canonical embedding, given by $j(g) = e \otimes g$.

(ii) Let Y^* be a Banach space with the Radon-Nikodým property. If $(\Omega, \mathcal{F}, P)$ is a probability space, and Ψ a finite Young function with complementary finite Young function Ψ^*, then $((H_\Psi(P,Y), (L_{\Psi^*}(P,Y^*), \langle \cdot,\cdot \rangle)$, where $\langle \cdot,\cdot \rangle : (H_\Psi(P,Y) \times L_{\Psi^*}(P,Y^*) \to \mathbb{R}$ is given by

$$\langle u, v \rangle = \int_\Omega u(t)v(t)\, dP,$$

is a dual pair (see [114]). The rest follows as in (i).

(iii) If E is a separable Banach lattice with order continuous norm and a quasi interior point $0 < e \in E$, then $(E\widetilde{\otimes}_l Y, E^*\widetilde{\otimes}_l Y^*, \langle \cdot,\cdot \rangle)$, where $\langle \cdot,\cdot \rangle : E\widetilde{\otimes}_l Y \times E^*\widetilde{\otimes}_l Y^* \to \mathbb{R}$ is given by

$$\langle u, v \rangle = \sum_{i,j} x_i^*(x_j) y_i^*(y_j),$$

and $u = \sum_j x_j \otimes y_j$, for some sequences $(x_i) \subseteq E$ and $(y_i) \subseteq Y$ such that $\| \sum_i |x_j| \|_E < \infty$ and $\lim_{n\to\infty} \|y_j\|_Y = 0$, and $v = \sum_j x_j^* \otimes y_j^*$, for some sequences $(x_i^*) \subseteq E^*$ and $(y_i^*) \subseteq Y^*$ such that $\| \sum_i |x_j^*| \|_{E^*} < \infty$ and $\lim_{n\to\infty} \|y_j^*\|_{Y^*} = 0$, is a dual pair (see [31]).

By the respective definitions of C° and $\mathbb{F}_{G,e}^{Z^*}$, it follows that

$$C^\circ = \{Z^* \in E^*\widetilde{\otimes}_l Y^* : Z^*(c) \geq 0 \text{ for all } c \in C\}$$

and, for all $x \in E\widetilde{\otimes}_l Y$,

$$\mathbb{F}_{G,e}^{Z^*}[x] = \left\{ g \in G : \sum_{i,j} x_i^*(x_j) y_i^*(y_j) \leq \sum_i x_i^*(e) y_i^*(g) \right\},$$

where $x = \sum_j x_j \otimes y_j$, for some sequences $(x_i) \subseteq E$ and $(y_i) \subseteq F$ such that $\| \sum_i |x_j| \|_E < \infty$ and $\lim_{n\to\infty} \|y_j\|_Y = 0$, and $Z^* = \sum_j x_j^* \otimes y_j^*$, for some sequences $(x_i^*) \subseteq E^*$ and $(y_i^*) \subseteq Y^*$ such that $\| \sum_i |x_j^*| \, \|_{E^*} < \infty$ and $\lim_{n\to\infty} \|y_j^*\|_{Y^*} = 0$. □

Remark 7.3.9. Consider the case $Y = \mathbb{R}^d$ and $G = \mathbb{R}^m$. Let $E^i = \mathbf{1}$ for all $1 \leq i \leq d$ and identify $g = (g_1, \ldots, g_m) \in \mathbb{R}^m$ and $(g_1, \cdots, g_m, 0 \ldots, 0) \in \mathbb{R}^d$. Then

$$(\mathbf{1} \otimes g) = \sum_{i=1}^m g_i E^i.$$

Consequently, for all $Z^* \in L_q(P, \mathbb{R}^d)$ and $X \in L_p(P, \mathbb{R}^d)$ (respectively, $Z^* \in L_\Psi(P)\widetilde{\otimes}_l \mathbb{R}^d$ and $X \in H_\Psi(P, \mathbb{R}^d)$) we may replace the expression for $\mathbb{F}_G^{Z^*}[x]$ in Corollary 7.3.8 (i) (respectively (ii)) by

$$\mathbb{F}_G^{Z^*}[x] = \left\{ u \in \mathbb{R}^m : \int_\Omega \left(x - \sum_{i=1}^m u_i E^i \right) \cdot Z^* \, dP \leq 0 \right\}.$$

This yields the expression as considered by Hamel et al. [79] for the case $L_p(P, \mathbb{R}^d)$.

Appendix A

Appendix: Functional analysis background

A.1 Riesz spaces

To properly understand the concept of duality, it is necessary to know some important properties about certain Riesz spaces. This section will cover these properties. The reader is referred to [32], [54], [167] and [168] for more details. Most theorems in this chapter are taken from [168].

Definition A.1.1. $(X, \leq)$ is a *partially ordered set* if $\leq$ is a partial ordering, i.e.

(i) $x \leq x$ for all $x \in X$,

(ii) $x \leq y$ and $y \leq z$ implies $x \leq z$ and

(iii) $x \leq y$ and $y \leq x$ implies $x = y$.

We will denote the partially ordered set $(X, \leq)$ by X if it is clear which partial ordering is meant.

Definition A.1.2. Let X be a partially ordered set.

(i) If every subset of X consisting of two elements has a supremum and an infimum then X is called a *lattice*.

(ii) We denote $\sup\{x, y\}$ by $x \vee y$ and $\inf\{x, y\}$ by $x \wedge y$ for all $x, y \in X$.

(iii) We write $x^+ = x \vee 0$, $x^- = (-x) \vee 0$ and $|x| = x^+ + x^-$.

(iv) $x_n \uparrow x$ denotes that (x_n) is an increasing sequence and $x = \sup\{x_n : n \in \mathbb{N}\}$.

Definition A.1.3. The real vector space E is called an *ordered vector space* if E has a partial ordering satisfying

(i) $f \leq g \Rightarrow f + h \leq g + h$ for every $f, g, h \in E$ and

(ii) $f \geq 0 \Rightarrow \alpha f \geq 0$ for every $\alpha \geq 0$ in $\mathbb{R}$.

If E is also a lattice with respect to the partial ordering, then E is called a *Riesz space* or a *vector lattice*.

Riesz spaces allow for the following algebraic structures.

Definition A.1.4. Let E be a Riesz space.

(i) $R \subseteq E$ is called a *Riesz subspace* if R is a vector subspace of E and for all $x, y \in R$ we have $x \wedge y \in R$ and $x \vee y \in R$.

(ii) $S \subseteq E$ is called *solid* if $f \in S$ implies $\left[\text{-}|f|, |f|\right] \subseteq S$.

(iii) $A \subseteq E$ is called an *ideal* if A is a solid vector subspace.

(iv) An ideal $B \subseteq E$ is called a *band* if for all $D \subseteq B$ with $f = \sup D$, we have $f \in B$.

Definition A.1.5. Let E be a Riesz space. Then E is called

(i) *Archimedean* if $\inf\{\frac{u}{n} : n \in \mathbb{N}\} = 0$ for all $u \in E^+$.

(ii) *Dedekind complete* if every non-empty subset of E that is bounded above (bounded below) has a supremum (infimum).

Definition A.1.6. Let E be a Riesz space. We say that $f, g \in E$ are *disjoint* if $|f| \wedge |g| = 0$ and we write $f \perp g$. If D is a non empty subset of E, then the set

$$D^d = \{f \in E : f \perp g \ \ \forall\, g \in D\}$$

is called the *disjoint complement* of D. This set is sometimes denoted by $D^{\perp}$.

Definition A.1.7. A map $P : E \to E$ is a *projection* if $P^2 x = P(Px) = Px$ for all $x \in E$.

Definition A.1.8. A band B in a Riesz space E is called a *projection band* if

$$E = B \oplus B^d.$$

Let B be a projection band of a Riesz space E. The map $P_B : E \to B$, defined by $P_B(x) = x_1$, where $x = x_1 + x_2$, $x_1 \in B$ and $x_2 \in B^d$, is a linear projection and a Riesz homomorphism, i.e. $|P_B x| = P_B|x|$. Moreover, P_B is *order continuous*, i.e. if $0 \leq x_\alpha \uparrow x$ in E, then $P_B(x_\alpha) \uparrow P_B(x)$. Furthermore, for all $x \in E^+$

$$P_B(x) = \sup\{y \in B : 0 \leq y \leq x\}.$$

If $B \oplus B^d = E$ for every band B in E, then E is said to have the *projection property.* It is well-known that Archimedean Riesz spaces have the projection property and Riesz spaces with the projection property are Dedekind complete.

For any nonempty subset D of the Riesz space E, the intersection of all bands containing D is called the *band generated by* D, denoted by B_D. In particular, if D consists of one element u, then we write $B_D = B_u$ and B_u is called the *principle band* generated by u.

The projection determined by a principal projection band B_u is denoted by P_u. It is known, see [167], that for all $x \in E^+$,

$$P_u(x) = \sup\{x \wedge n|u| : n \in \mathbb{N}\}. \tag{A.1}$$

Definition A.1.9. An element $e > 0$ in the Riesz space E is called a *weak order unit* if $B_e = E$.

If e is a weak order unit of E, then (A.1) implies $x \wedge ne \uparrow x$ for all $x \in E^+$. The next result shows that the positivity assumption on x is not required.

Lemma A.1.10. *If the Riesz space E has a weak order unit $e \in E^+$, then*

$$x \wedge ne \uparrow x \qquad \text{for all } x \in E.$$

Proof. For any $x \in E$

$$\begin{aligned} x \wedge ne &= (x \wedge ne)^+ - (x \wedge ne)^- \\ &= (x \wedge ne) \vee 0 - (-(x \wedge ne)) \vee 0 \\ &= (x \vee 0) \wedge (ne \vee 0) - ((-x) \vee (-ne)) \vee 0 \\ &= x^+ \wedge ne - ((-x) \vee (-ne) \vee 0) \\ &= x^+ \wedge ne - ((-x) \vee 0) \\ &= x^+ \wedge ne - x^-. \end{aligned}$$

Since $x^+ \geq 0$, using the above we know that $x^+ \wedge ne \uparrow x^+$. Therefore,

$$x \wedge ne = x^+ \wedge ne - x^- \uparrow x^+ - x^- = x,$$

which completes the proof. □

Consider a vector subspace A of a vector space V. We can divide V into equivalence classes modulo A by introducing an equivalence relation in V defined by saying that f and g in V are equivalent if and only if $f - g \in A$. The set of all equivalence classes is called the *quotient space of V modulo A* and is denoted by V/A.

The quotient space V/A is a vector space if we define the vector space operations by

$$[f] + [g] = [f + g] \text{ and}$$
$$\alpha[f] = [\alpha f]$$

for all $[f], [g] \in V/A$ and $\alpha \in \mathbb{R}$.

Theorem A.1.11. *If A is an ideal in the Riesz space E, then the quotient space E/A is a Riesz space with respect to the following partial ordering: given $[f]$ and $[g]$ in E/A, we write $[f] \leq [g]$ whenever there exists elements $f_1 \in [f]$ and $g_1 \in [g]$ satisfying $f_1 \leq g_1$.*

A.2 Order duality

Now that we know the basics of Riesz spaces we can start looking at the definitions of algebraic and order duals and look at some of their important properties.

Definition A.2.1. Let X and Y be vector spaces.

(i) A *linear operator* is a map $T : X \to Y$ that satisfies $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ for each $\alpha, \beta \in \mathbb{R}$ and $x, y \in X$.

(ii) We shall denote by $L(X, Y)$ the vector space of all linear operators from X into Y.

(iii) In the case where $Y = \mathbb{R}$, we shall denote $L(X, Y)$ by $X^\#$. The elements of $X^\#$ are called *linear functions* and $X^\#$ is called the *algebraic dual* of X.

Definition A.2.2. Let E and F be Riesz spaces and let $T : E \to F$ be a linear operator.

(i) T is said to be a *positive* operator if $x \geq 0$ in E implies $T(x) \geq 0$ in F.

(ii) T is said to be *order bounded* if T maps any order interval in E into an order interval in F. We will denote the set of all order bounded linear operators by $L^b(E, F)$.

(iii) T is said to be *regular* if T can be written as $T = T_1 - T_2$ where both T_1 and T_2 are positive. We will denote the set of all regular linear operators by $L^r(E, F)$.

(iv) T is called *order continuous* if for any downwards directed set $D \subseteq E$ with $D \downarrow 0$ we have $\inf\{|T(f)| : f \in D\} = 0$ in F. The vector space of all order continuous operators in $L^b(E, F)$ is denoted by $L^n(E, F)$.

(v) T is called *σ-order continuous* if for any monotone sequence (f_n) in E with $f_n \downarrow 0$ we have $\inf\{|T(f_n)| : n \in \mathbb{N}\} = 0$ in F. The vector space of all σ-order continuous operators in $L^b(E, F)$ is denoted by $L^c(E, F)$.

Theorem A.2.3. *Let E be a Riesz space and F a Dedekind complete Riesz space. Then $L^b(E, F)$ is a Dedekind complete Riesz space.*

The following result shows how the spaces $L^r(E, F)$ and $L^b(E, F)$ are related.

Theorem A.2.4. *If E and F are Riesz spaces with F Dedekind complete and $T : E \to F$ is a linear operator, then T is regular if and only if T is order bounded, i.e. then $L^r(E, F) = L^b(E, F)$.*

Since $\mathbb{R}$ is a Dedekind complete Riesz space, we have that if $F = \mathbb{R}$, then order bounded linear functionals on E are the same as regular linear functionals on E. This space of all order bounded linear functionals is called the *order dual* of E and is denoted by $E^\sim$.

We will denote the set of all order continuous linear functional of E by $E_n^\sim$ and the set of all σ-order continuous linear functionals of E by $E_c^\sim$. Their disjoint complements will be denoted by $E_t^\sim$ and $E_s^\sim$ respectively. The elements of $E_s^\sim$ are called *singular* elements.

Since $E^\sim$ is Dedekind complete, it has the projection property and we have the following theorem taken from [167].

Theorem A.2.5. *Let E be a Riesz space. Then*

(i) $E_c^\sim$ is a projection band in $E^\sim$, i.e.

$$E^\sim = E_c^\sim \oplus E_s^\sim.$$

Therefore, for any $\phi \in (E^\sim)^+$, there exists a unique decomposition

$$\phi = \phi_c + \phi_s,$$

with $\phi_c \in (E_c^\sim)^+$ and $\phi_s \in (E_s^\sim)^+$.

(ii) Similarly for $E_n^\sim$ and $E_t^\sim$, i.e.

$$E^\sim = E_n^\sim \oplus E_t^\sim.$$

The following result, also taken from [167], gives a complete description of ϕ_c and ϕ_s for $\phi \geq 0$.

Theorem A.2.6. *For any $\phi \geq 0$ in $E^\sim$ and $u \geq 0$ in E, the σ-order continuous component ϕ_c of ϕ satisfies*

$$\phi_c(u) = \inf\{\lim \phi(u_n) : 0 \leq u_n \uparrow u\}.$$

Hence, ϕ is singular, i.e. $\phi = \phi_s$, if and only if for any $\epsilon > 0$, there exists a sequence $0 \leq u_n \uparrow u$ in E such that $0 \leq \phi(u_n) < \epsilon$ for all n.

A.3 Banach lattices

The next step is to introduce norms on Riesz spaces which are compatible with the order structure.

Definition A.3.1. Let X be a real vector space. A map $\|\cdot\| : X \to \mathbb{R}$ is called a *norm* if

(i) $\|f\| \geq 0$ for all $f \in X$ and $\|f\| = 0$ if and only if $f = 0$,

(ii) $\|\alpha f\| = |\alpha|\|f\|$ for all $f \in X$ and $\alpha \in \mathbb{R}$ and

(iii) $\|f + g\| \leq \|f\| + \|g\|$ for all $f, g \in X$.

The pair $(X, \|\cdot\|)$ is called a *normed space.* If $(X, \|\cdot\|)$ is complete with respect to the norm, i.e. every norm Cauchy sequence has a limit in X, then $(X, \|\cdot\|)$ is called a *Banach space.*

We denote the normed space $(X, \|\cdot\|)$ by X and call X a Banach space, if it is clear which norm is meant. The next theorem shows that if a vector space is complete with respect to two norms, then the two norms must be equivalent.

Theorem A.3.2. *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on a vector space X. If $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ are Banach spaces, then $\|\cdot\|_1$ is equivalent to $\|\cdot\|_2$, i.e. there exist $\alpha, \beta \in \mathbb{R}$ such that $\alpha\|\cdot\|_1 \leq \|\cdot\|_2 \leq \beta\|\cdot\|_2$.*

Definition A.3.3. Let E be a Riesz space.

(i) If E is equipped with a norm $\|\cdot\|$, then $\|\cdot\|$ is called a *Riesz norm* if for all $f, g \in E$ with $|f| \leq |g|$, we have that $\|f\| \leq \|g\|$.

(ii) A Riesz space E equipped with a Riesz norm is called a *normed Riesz space.*

(iii) If a normed Riesz space E is complete with respect to the norm, then E is called a *Banach lattice*.

Definition A.3.4. The normed Riesz space E is said to have an *order continuous norm* if for any subset $D \downarrow 0$ in E, we have $\inf\{\|f\| : f \in D\} = 0$. The norm is said to be σ*-order continuous* if for any sequence $f_n \downarrow 0$ in E, we have $\|f_n\| \downarrow 0$.

We next define some special types of Banach lattices, which we will encounter at a later stage.

Definition A.3.5. Let E denote a normed Riesz space.

(i) $(E, \|\cdot\|)$ is called an L*-normed* space if $\|\cdot\|$ satisfies $\|x+y\| = \|x\| + \|y\|$ for all $x, y \in E^+$. An L-normed Banach lattice is called an AL-space.

(ii) $(E, \|\cdot\|)$ is called an M*-normed* space if $\|\cdot\|$ satisfies $\|x \vee y\| = \|x\| \vee \|y\|$ for all $x, y \in E^+$. An M-normed Banach lattice is called an AM-space.

A.4 Norm duality

In this section, we will define the Banach dual of normed Riesz spaces and relate Banach duals to order duals.

Definition A.4.1. Let X and Y be normed spaces. A linear operator $T : X \to Y$ is said to be *norm bounded* if there exists a constant $A > 0$ such that $\|Tx\| \leq A\|x\|$ for all $x \in X$.

Definition A.4.2. Let X and Y be normed spaces.

(i) We define the normed space $\mathcal{L}(X, Y)$ by

$$\mathcal{L}(X, Y) := \{T \in L(X, Y) : \text{T is norm bounded}\}$$

together with the *operator norm* $\|\cdot\|$, defined by

$$\|T\| = \sup\{\ \|Tx\| : \|x\| \leq 1\} \text{ for all } T \in \mathcal{L}(X, Y).$$

(ii) In the case where $Y = \mathbb{R}$, we will denote $\mathcal{L}(X, Y)$ by X^*. The elements of X^* are called *linear functionals* and X^* is called the *continuous dual space* or the *Banach dual space* of X.

If X is a normed space and Y is a Banach space, then $\mathcal{L}(X,Y)$ is also a Banach space with respect to the operator norm. In particular, we have that X^* is a Banach space.

Theorem A.4.3. *Let E and F be Banach lattices with F Dedekind complete. Then, for all $T \in \mathcal{L}(E,F)$ and $T \geq 0$,*

$$\|T\| = \sup\{\|Tx\| : 0 \leq x \in E,\ \|x\| \leq 1\}.$$

Theorem A.4.4. *If E is a Banach lattice, then $E^* = E^\sim$.*

Definition A.4.5. Let X and Y be Banach spaces and $T : X \to Y$ a bounded linear map. Then the *adjoint* $T^* : Y^* \to X^*$ of T is given by

$$\left[T^*(y^*)\right](x) = y^*(Tx),$$

for all $x \in X$ and $y^* \in Y^*$.

A.5 Semi-M-spaces

In this section, we explain what is meant by a semi-M-space and characterise these spaces in terms of AL-spaces.

Definition A.5.1. Let $(E, \|\cdot\|)$ be a normed Riesz space. The space E is called a *semi-M-space* if it has the property that for all u_1 and u_2, positive elements in E with $\|u_1\| = \|u_2\| = 1$ and for any sequence (v_n) with $u_1 \vee u_2 \geq v_n \downarrow 0$, we have $\lim \|v_n\| \leq 1$.

Example A.5.2.

1. If E is an M-space, then E is also a semi-M-space, as in this case

$$\|u_1\| = \|u_2\| = 1 \Rightarrow \|u_1 \vee u_2\| = 1.$$

2. If the norm $\|\cdot\|$ is σ-order continuous, then E is a semi-M-space, as in this case

$$v_n \downarrow 0 \Rightarrow \lim \|v_n\| = 0.$$

The following is an important characterisation of semi-M-spaces.

Theorem A.5.3. *The normed Riesz space E is a semi-M-space if and only if the band E_s^* in the Banach dual E^* is an AL-space.*

Proof. [167, Thm 119.2] We will only prove the one direction. The proof of the other direction is not necessary in this thesis but can be found in the above reference.

Assume that E is a semi-M-space. The norms of E and E^* will be denoted by $\|\cdot\|$ and $\|\cdot\|^*$ respectively. Let ϕ_1 and ϕ_2 be positive elements of E_s^*. Then, $\|\phi_1 + \phi_2\|^* \leq \|\phi_1\|^* + \|\phi_2\|^*$. We need to prove this inequality in the other direction.

Let $\epsilon > 0$ be given. Since E is a semi-M-space, it follows from Theorem A.4.3, that there exists positive elements u_1 and u_2 in E such that $\|u_1\| = \|u_2\| = 1$,

$$\phi_1(u_1) > \|\phi_1\|^* - \epsilon \quad \text{and} \quad \phi_2(u_2) > \|\phi_2\|^* - \epsilon.$$

Set $u = u_1 \vee u_2$ and $\phi = \phi_1 + \phi_2$. Then $u \in E^+$ and $0 \leq \phi \in E_s^*$. Hence, by Theorem A.2.6, there exists a sequence $0 \leq w_n \uparrow u$ in E such that $\phi(w_n) < \epsilon$ for all n. Writing $v_n = u - w_n$ for all n, we have that $u_1 \vee u_2 = u \geq v_n \downarrow 0$ and so $\lim \|v_n\| \leq 1$ as we assumed that E is a semi-M-space. Therefore, there exists a natural number n_0 such that $\|v_n\| < 1 + \epsilon$ for all $n \geq n_0$. For these values of n, we get

$$\begin{aligned}
\|\phi_1 + \phi_2\|^* &= \|\phi\|^* \\
&\geq \phi\Big(\frac{v_n}{1+\epsilon}\Big) \\
&= \phi\Big(\frac{u - w_n}{1+\epsilon}\Big) \\
&\geq \phi\Big(\frac{u}{1+\epsilon}\Big) - \frac{\epsilon}{1+\epsilon} \\
&= \frac{1}{1+\epsilon}\Big\{\phi_1(u) + \phi_2(u) - \epsilon\Big\} \\
&\geq \frac{1}{1+\epsilon}\Big\{\phi_1(u_1) + \phi_2(u_2) - \epsilon\Big\} \\
&\geq \frac{1}{1+\epsilon}\Big\{\ \|\phi_1\|^* + \|\phi_2\|^* - 3\epsilon\Big\}.
\end{aligned}$$

Hence, for all $\epsilon > 0$, we have $\|\phi_1 + \phi_2\|^* \geq \frac{1}{1+\epsilon}\Big\{\ \|\phi_1\|^* + \|\phi_2\|^* - 3\epsilon\Big\}$, which implies that $\|\phi_1 + \phi_2\|^* \geq \|\phi_1\|^* + \|\phi_2\|^*$. This proves that $\|\cdot\|^*$ is 1-additive on E_s^*. Since E_s^* is a band in the Banach space E^*, it is norm complete and therefore, E_s^* is an AL-space. □

The proof of the theorem that E is a semi-M-space if and only if E_s^* is an AL-space is due to de Jonge [37]. However, he assumed that E has the principal projection property to prove it. Fremlin noted that the principle projection property is superfluous.

Orlicz spaces are an example of non-trivial semi-M-spaces [167].

A.6 Measure Theory

Let Ω be a non-empty set. A collection S of subsets of Ω is called an *algebra* on Ω if S has the following properties:

(i) $\Omega \in S$,

(ii) if $A \in S$ then $A^c \in S$, and

(iii) if $A_1, A_2, \ldots, A_n \in S$, then $\bigcup_{i=1}^n A_i \in S$.

An algebra S on Ω is called a *σ-algebra* if it is closed under countable unions, i.e. if $A_k \in S$ for $k \in \mathbb{N}$, then $\bigcup_{k=1}^\infty A_k \in S$. It is easily seen that a σ-algebra is closed under countable intersections.

If Ω is a non-empty set and S a σ-algebra on Ω, then (Ω, S) is called a *measurable space* and the sets in S are called *measurable sets*.

Let (Ω, S) be a measurable space. A function $\mu : S \to [0, \infty]$ is called a *measure* if $\mu(\emptyset) = 0$ and μ is σ-additive, i.e. if $(A_k)_{k\in\mathbb{N}}$ is a sequence of disjoint sets in S, then

$$\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k).$$

The triplet (Ω, S, μ) is called a *measure space*. If $\mu(\Omega) = 1$, then μ is called a *probability measure* and (Ω, S, μ) is called a *probability space*. If $\mu(\Omega) < \infty$, then μ is called a *finite* measure. A measure μ is called *σ-finite* if there exists a sequence $(X_k)_{k\in\mathbb{N}}$ of sets in S such that $\Omega = \bigcup_{k=1}^\infty X_k$ and $\mu(X_k) < \infty$ for every $k \in \mathbb{N}$.

Financial mathematics is based on the theory of stochastic processes. Hence, it is very important to understand what a stochastic process is and the concepts surrounding it. Most of this section is taken from [86, 87, 123].

Definition A.6.1. A *stochastic process* $(X_t)_{t\in I}$ is a collection of real-valued random variables $X_t : \Omega \to \mathbb{R}$, indexed by $t \in I$, where I is some index set. When $I = \mathbb{N}$, then X is called a *discrete-time* process and when $I = \mathbb{R}^+$, then X is called a *continuous-time* process.

The following definition classifies a widely-used type of stochastic process.

Definition A.6.2. A function $x : \mathbb{R}^+ \to \mathbb{R}$ is called *right continuous with left limits* or more commonly *càdlàg*, if the left limit x_{t-} and the right limit x_{t+} are finite and $x_t = x_{t+}$ for all $t \in \mathbb{R}^+$. A process is called càdlàg if all its sample paths are càdlàg almost surely.

In financial mathematics, the idea of the flow of information plays an important role. This is formalised through the concept of filtrations.

Definition A.6.3. Consider a probability space $(\Omega, \mathcal{F}, P)$. A family of σ-algebras $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{R}^+}$ is called a *filtration* if $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for all $s \leq t \in \mathbb{R}^+$. If we endow a probability space with a filtration $\mathbb{F}$, such that

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_T = \mathcal{F},$$

then the combined structure $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is called a *filtered probability space.*

The filtration $\mathbb{F}$ is used to model the flow of information. As time passes, an observer knows more and more detailed information, that is finer and finer partitions of Ω.

Definition A.6.4. A random variable X on $(\Omega, \mathcal{F})$ is called a $\mathcal{F}_t$-*measurable* function if for any Borel set $B \in \mathcal{B}$, the set $\{\omega : X_t(\omega) \in \mathcal{B}\}$ is a member of $\mathcal{F}_t$.

The definition of $\mathcal{F}_t$-measurable is equivalent to $\{\omega : X_t(\omega) \leq x\} \in \mathcal{F}_t$ for all $x \in \mathbb{R}$.

Definition A.6.5.

(i) Consider a probability space $(\Omega, \mathcal{F}, P)$. The σ-algebra $\mathcal{F}$ is called P-*complete* if and only if $A \subseteq B$ with $B \in \mathcal{F}$ such that $P(B) = 0$ implies that $A \in \mathcal{F}$.

(ii) A filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is *complete* if the σ-algebra $\mathcal{F}$ is P-complete and $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$.

Definition A.6.6. A filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is said to satisfy the *usual conditions* if it is complete and the filtration is *right continuous*, i.e.

$$\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{t>s} \mathcal{F}_s.$$

The right continuity implies that any information known immediately after t is also known at t.

Definition A.6.7. A nonnegative random variable τ, which is allowed to take the value ∞, is called a *stopping time* with respect to the filtration $\mathbb{F}$ if

$$\{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \in \mathbb{R}^+.$$

Definition A.6.8. Let τ be a stopping time. The stopped process $M_{t\wedge\tau}$ is defined by

$$M_{t\wedge\tau} = M_t^\tau = \mathbb{1}_{\{t<\tau\}} X_t + \mathbb{1}_{\{t\geq\tau\}} X_\tau \qquad \text{for all } t \in \mathbb{R}^+.$$

Definition A.6.9.

(i) A process X is said to be *adapted* to $\mathbb{F}$ if X_t is $\mathcal{F}_t$-measurable for every $t \in \mathbb{R}^+$.

(ii) A process X is said to be *predictable* with respect to $\mathbb{F}$, if it is one of the following:

  (a) a left-continuous adapted process, in particular, a continuous adapted process.

  (b) a limit (almost sure, in probability) of left-continuous adapted processes.

  (c) a regular right-continuous process such that, for any stopping time τ, X_τ is $\mathcal{F}_\tau$-measurable, where $\mathcal{F}_\tau$ is the σ-algebra generated by the sets $A \cap \{T < t\}$ for $A \in \mathcal{F}_t$.

  (d) a Borel-measurable function of a predictable process.

(iii) A process X is *locally bounded* if $\sup_{0\leq t\leq T} |X_t| < \infty$, i.e. if there exists a sequence (τ_n) of stopping times, increasing almost surely to ∞, such that the stopped processes $X_t^{\tau_n}$ are uniformly bounded for each $n \in \mathbb{N}$.

We next define processes with finite variation, as this will be needed to define certain types of martingales.

Definition A.6.10. Let $t \in \mathbb{R}^+$. A sequence of partitions $(\pi_t^n)_{n\in\mathbb{N}}$ of the interval $[0, t]$, with $\pi_t^n = \{t_0^n, \ldots, t_{m_n}^n\}$ for each $n \in \mathbb{N}$ and $\lim_{n\to\infty} \|\pi_t^n\| = 0$, is called a *Riemann sequence*.

Definition A.6.11. Let X be a càdlàg adapted process and for each $t \in \mathbb{R}^+$, choose a Riemann sequence $(\pi_t^n)_{n\in\mathbb{N}}$. For $p > 0$ and for the partition π_t^n, define

$$S^{(p)}(X, \pi_t^n) = \sum_{i=0}^{m_n-1} |X_{t_{i-1}^n} - X_{t_i^n}|^p.$$

If for each $t \in \mathbb{R}^+$

$$V_t^{(p)}(X) := \lim_{n\to\infty} S^{(p)}(X, \pi_t^n) < \infty \quad \text{a.s.},$$

then the process $V^{(p)}(X)$ is well-defined and is called the p^{th} *variation* of X.

Definition A.6.12. A càdlàg adapted process A is of *finite variation* if $V_t^{(1)}(A) < \infty$ a.s. for all $t \in \mathbb{R}^+$. A càdlàg adapted process A is of bounded variation if $\sup_{t\in\mathbb{R}^+} V_t^{(1)}(A) < \infty$.

A.7 L_p-spaces

Let $(\Omega, \mathcal{F}, P)$ be a measure space and denote by $\mathcal{L}_0(P)$ the set of measurable functions $f : \Omega \to \mathbb{R}$. It is well-known that $\mathcal{L}_0(P)$ is an ordered vector space under pointwise addition, scalar multiplication and ordering. Moreover, $\mathcal{L}_0(P)$ is a Dedekind complete Riesz space, in which the following identities hold:

$$(f \vee g)(\omega) = \max\{f(\omega), g(\omega)\} \text{ a.e. and } (f \wedge g)(\omega) = \min\{f(\omega), g(\omega)\} \text{ a.e.}$$

for all $f, g \in \mathcal{L}_0(P)$. Two measurable real functions f and g on Ω are *equivalent* if they agree almost everywhere, i.e. $f \sim g$ if and only if $f = g$ a.e.. It follows easily that $\sim$ is an equivalence relation.

The space $L_0(P)$ is defined to be the collection of equivalence classes of (P-a.e. equal) measurable functions, i.e.

$$L_0(P) = \mathcal{L}_0(P)/\sim\ .$$

Endowed with the canonical addition, scalar multiplication and order of quotient spaces, $L_0(P)$ is a Dedekind complete Riesz space (see [168]).

The space $L_\infty(P)$ comprises of the equivalence classes of measurable functions, which are essentially bounded, i.e. $\inf\{M > 0 : |f| \leq M \text{ a.e.}\} < \infty$. Then $L_\infty(P)$ is a Banach lattice with respect to the *essential sup norm*, defined by

$$\|f\|_\infty = \inf\{M > 0 : |f| \leq M\}.$$

For $1 \leq p < \infty$, the space $L_p(P)$ is defined to be the collection of equivalence classes of measurable functions f for which $\int_\Omega |f|^p dP < \infty$. Then $L_p(P)$ is an ideal in $L_0(P)$ (and therefore a Dedekind complete Riesz space in its own right). Moreover, $L_p(P)$ is a Banach lattice with respect to the *p-norm* $\|\cdot\|_p$ defined by

$$\|f\|_p = \Big(\int_\Omega |f|^p dP\Big)^{\frac{1}{p}}.$$

It is well known that $\mathbf{1}$ is a weak order unit in $L_p(P)$ for $p = 0$ and $1 \leq p \leq \infty$.

What is interesting about the duals of the L_p-spaces is that they are themselves L_p-spaces. The Riesz representation theorem, stated next, shows how these duals can be represented.

Theorem A.7.1. *Suppose $1 \leq p < \infty$, P is a σ-finite positive measure on Ω, and Φ is a bounded linear functional on $L_p(P)$. Then there exists a unique $g \in L_q(P)$, where q is the exponent conjugate of p, i.e. $\frac{1}{p} + \frac{1}{q} = 1$, such that*

$$\Phi(f) = \int_\Omega fg dP \quad \text{for all } f \in L_p(P).$$

Moreover, if Φ *and* g *are related as above, then we have*

(i) $\|\Phi\|_p = \|g\|_q$ *and*

(ii) $\Phi \geq 0$ *if and only if* $g \geq 0$*.*

In other words, $L_p(P)$ is isometrically and order isomorphic to $(L_q(P))^*$, under the stated conditions.

The following is one of the fundamental results of measure and integration theory.

Theorem A.7.2 (Radon-Nikodým theorem)**.** *Let* P *and* Q *be positive bounded measures on a* σ*-algebra* $\mathcal{F}$ *in a set* Ω*. Then there exists a unique* $g \in L_1(P)$ *such that*

$$Q(E) = \int_E g dP \quad \text{for all } E \in \mathcal{F}.$$

The function g *is called the* Radon-Nikodým derivative *of* Q *with respect to* P*, usually denoted by* $\frac{dQ}{dP}$*.*

Note that $\frac{dQ}{dP}$ is called the *density* of the measure Q if $\int_\Omega \frac{dQ}{dP}\, dP = 1$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_1$ a sub σ-algebra of $\mathcal{F}$. If $f \in L_1(P)$, define

$$Q(A) = \int_A f(\omega) dP(\omega) \text{ for all } \omega \in \mathcal{F}_1.$$

Then $Q : \mathcal{F}_0 \to \mathbb{R}$ is a measure which is absolutely continuous with respect to P, i.e. if $A \in \mathcal{F}_1$ and $P(A) = 0$, then $Q(A) = 0$, denoted by $Q << P$. By the Radon-Nikodým theorem, there exists $g \in L_1(P)$ which is P-almost surely unique, $\mathcal{F}_1$-measurable and

$$\int_A f(\omega) dP(\omega) = \int_A g(\omega) dP(\omega) \text{ for all } \omega \in \mathcal{F}_1.$$

As is customary, g is called the *conditional expectation of* f *relative to* $\mathcal{F}_1$, and is denoted by $\mathbb{E}[f|\mathcal{F}_1]$ (or alternatively by $\mathbb{E}_{\mathcal{F}_1}[f]$).

If viewed as an operator on $L_1(P)$, it is well-known that $\mathbb{E}[\,\cdot\,|\mathcal{F}_1]$ has the following properties.

Theorem A.7.3. *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and* $\mathcal{F}_1$ *a sub* σ*-algebra of* $\mathcal{F}$*. Then the map* $\mathbb{E}[\,\cdot\,|\mathcal{F}_1] : L_1(P) \to L_1(P)$ *has the following properties:*

(i) $\mathbb{E}[\,\cdot\,|\mathcal{F}_1]$ *is linear,*

(ii) $\mathbb{E}[\,\cdot\,|\mathcal{F}_1]$ *is positive,*

(iii) $\mathbb{E}[\mathbf{1}|\mathcal{F}_1] = \mathbf{1}$ *a.s.,*

(iv) $\mathbb{E}\big[\,\mathbb{E}[\,\cdot\,|\mathcal{F}_1]\,\big|\mathcal{F}_1\big] = \mathbb{E}[\,\cdot\,|\mathcal{F}_1]$*; i.e.* $\mathbb{E}[\,\cdot\,|\mathcal{F}_1]$ *is idempotent,*

(v) if $\mathcal{F}_2$ *is a sub* σ*-algebra of* $\mathcal{F}$ *and* $\mathcal{F}_1 \subseteq \mathcal{F}_2$*, then*

$$\mathbb{E}\big[\,\mathbb{E}[\,\cdot\,|\mathcal{F}_1]\,\big|\mathcal{F}_2\big] = \mathbb{E}\big[\,\mathbb{E}[\,\cdot\,|\mathcal{F}_2]\,\big|\mathcal{F}_1\big] = \mathbb{E}[\,\cdot\,|\mathcal{F}_1],$$

(vi) $\mathbb{E}[\,\cdot\,|\mathcal{F}_1]$ *is order continuous,*

(vii) if $\phi : \mathbb{R} \to \mathbb{R}$ *is convex, then* $\phi(\mathbb{E}[\,\cdot\,|\mathcal{F}_1]) \leq \mathbb{E}[\phi(\cdot)|\mathcal{F}_1]$*, which is known as* Jensen's inequality*, and*

(viii) $\|\,\mathbb{E}[\,\cdot\,|\mathcal{F}_1]\,\| = 1$*.*

Moreover, if $1 \leq p < \infty$*, the restriction of* $\mathbb{E}[\,\cdot\,|\mathcal{F}_1]$ *to* $L_p(P)$*, again denoted by* $\mathbb{E}[\,\cdot\,|\mathcal{F}_1]$*, is a map* $\mathbb{E}[\,\cdot\,|\mathcal{F}_1] : L_p(P) \to L_p(P)$ *such that* $\|\,\mathbb{E}[\,\cdot\,|\mathcal{F}_1]\,\| = 1$*.*

A.8 Martingales

The theory of martingales is central to all of financial mathematics. In this section we will define some of the basic concepts.

Definition A.8.1. A process M is called a *P-martingale* if

1. M is adapted to $\mathbb{F}$,
2. M_t is integrable for all $t \in \mathbb{R}^+$ and
3. $\mathbb{E}_P[\,M_t|\mathcal{F}_s\,] = M_s$ for all $s \leq t \in \mathbb{R}^+$.

If it is clear from the context what the measure is, we call a P-martingale a martingale.

Definition A.8.2. A P-martingale M is *uniformly integrable* if

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}^+} \int_{\{|M_t|\geq n\}} |M_t|\, dP = 0.$$

Definition A.8.3. An adapted process M_t is called a *local P-martingale* if there exists a sequence of stopping times τ_n, such that $\tau_n \uparrow \infty$ and for each n the stopped process $M_{t\wedge\tau}$ is a uniformly integrable martingale in t.

The most general processes, for which stochastic calculus has been developed, are semi-martingales.

Definition A.8.4. A regular càdlàg adapted process S is a *semimartingale* if it can be represented as a sum of two processes: a local martingale M_t with $M_0 = 0$ and a process of finite variation A_t with $A_0 = 0$, and

$$S_t = S_0 + M_t + A_t.$$

If we consider a process S with unbounded jumps, then we need to introduce another type of martingale, which allows for the fundamental theorem of asset pricing to be defined in a more general setting. The concept of sigma-martingales was introduced by Chou [28] and Émery [58], although the name was first coined by Delbaen and Schachermayer [40].

Definition A.8.5. A $\mathbb{R}^d$-valued semimartingale S is called a *sigma-martingale* if there exists a $\mathbb{R}^d$-valued martingale M and an M-integrable predictable $\mathbb{R}^+$-valued process φ such that $S = \varphi \cdot M$.

The following characterisation of sigma-martingales was noted by Émery [58].

Proposition A.8.6. *Let S be a d-dimensional semimartingale on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, Q)$. The following conditions are equivalent:*

(i) S is a sigma-martingale.

(ii) There exists (scalar) processes K^i with paths, which Q-a.s. never touch zero, such that K^i is predictable and S^i-integrable under Q and $K^i \cdot S^i$ is a local martingale.

(iii) There exists a d-dimensional martingale N and a positive (scalar) process φ, with each φ^i predictable and N^i-integrable under Q, such that $S^i = \varphi^i \cdot N^i$.

We will denote the set of sigma-martingales measures absolutely continuous with respect to P by $\mathbb{M}_\sigma(S)$, i.e.

$$\mathbb{M}_\sigma(S) = \{Q << P : S \text{ is a sigma-martingale under } Q\}.$$

Note that if S is bounded (resp. locally bounded), then

$$\mathbb{M}_\sigma(S) = \{Q << P : S \text{ is a martingale (resp. local martingale) under } Q\}.$$

The following example of sigma-martingales is taken from [40].

Example A.8.7.

1. A local martingale is a sigma-martingale.

2. Let $(\Omega, \mathcal{F}, P)$ be such that there are two independent stopping times T and U defined on it, both having an exponential distribution with parameter 1. Define M by

$$M_t = \begin{cases} 0 & \text{for } t < T \wedge U \\ 1 & \text{for } t \geq T \wedge U \text{ and } T = T \wedge U \\ -1 & \text{for } t \geq T \wedge U \text{ and } T = T \wedge U. \end{cases}$$

It is easy to verify that M is almost surely well-defined and is a martingale with respect to the filtration $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$ generated by M. Let $\varphi_t = \frac{1}{t}$. Then, φ_t is predictable (as it is deterministic) and it is M-integrable. Define $X = \varphi \cdot M$, i.e.

$$X_t = \begin{cases} 0 & \text{for } t < T \wedge U \\ \frac{1}{T \wedge U} & \text{for } t \geq T \wedge U \text{ and } T = T \wedge U \\ -\frac{1}{T \wedge U} & \text{for } t \geq T \wedge U \text{ and } T = T \wedge U. \end{cases}$$

The process X is well-defined but fails to be a martingale as $\mathbb{E}[|X_t|] = \infty$ for all $t > 0$. X also fails to be a local martingale as $\mathbb{E}[|X_T|] = \infty$ for each stopping time T that is not identically zero (see [58] for more details). But of course, X is a sigma-martingale.

Definition A.8.8. A probability measure Q on $(\Omega, \mathcal{F})$ is called an *equivalent (local/sigma) martingale measure* for S if Q is equivalent to P, i.e. $P(A) = 0$ if and only if $Q(A) = 0$ (they agree on impossible events), and S is a (local/sigma) martingale under Q.

Appendix B

Appendix: Tensor products on Banach spaces

B.1 Definition

In this section, we assume that X, Y and Z are vector spaces.

Definition B.1.1. A mapping $A : X \times Y \to Z$ is *bilinear* if it is linear in each variable, i.e.

$$A(a_1x_1 + a_2x_2, y) = a_1A(x_1, y) + a_2A(x_2, y) \text{ and}$$
$$A(x, b_1y_1 + b_2y_2) = b_1A(x, y_1) + b_2A(x, y_2)$$

for all $x_i, x \in X$, $y_i, y \in Y$ and scalars a_i, b_i, $i = 1, 2$.

We denote the vector space of bilinear mappings from $X \times Y$ into Z by $B(X \times Y, Z)$. If Z is the scalar field, then we write $B(X \times Y)$.

Definition B.1.2. A *tensor product* of the vector spaces X and Y is a pair (T, t) consisting of a vector space T and a bilinear mapping $t : X \times Y \to T$ which satisfies the following universal mapping property:

(UMP) If (G, A) is a pair consisting of a vector space G and a bilinear mapping $A : X \times Y \to G$, then there exists a unique linear mapping $A^l : T \to G$ such that $A = A^l \circ t$.

It is well known, and easy to prove, that the pair (T, t) is essentially unique. The pair is denoted by $(X \otimes Y, \otimes)$, and $X \otimes Y$ is referred to as the *tensor product* of X and Y.

For $x \in X$ and $y \in Y$, define

$$(x \otimes y)(A) = A(x, y) \text{ for all } A \in B(X \times Y).$$

As $x \otimes y$ is a linear map on $B(X \times Y)$, the space $X \otimes Y$ is (under isomorphism) the vector subspace in

$$(B(X \times Y))' = \{f : B(X \times Y) \to \mathbb{R} : f \text{ is linear}\}$$

spanned by $\{x \otimes y : x \in X, y \in Y\}$, which shows that $X \otimes Y$ exists.

A typical tensor in $X \otimes Y$ has the form

$$u = \sum_{i=1}^{n} \lambda_i x_i \otimes y_i, \tag{B.1}$$

where $n \in \mathbb{N}$, $\lambda \in \mathbb{R}$, $x_i \in X$ and $y_i \in Y$. The representation of u is not necessarily unique.

The mapping $(x, y) \to x \otimes y$ can be seen as a type of multiplication on $X \times Y$, which has the following properties:

(i) $(x_1 + x_2) \otimes y = x_1 \otimes y + x_2 \otimes y$,

(ii) $x \otimes (y_1 + y_2) = x \otimes y_1 + x \otimes y_2$,

(iii) $\lambda(x \otimes y) = (\lambda x) \otimes y = x \otimes (\lambda y)$,

(iv) $0 \otimes y = x \otimes 0 = 0$.

Because of (iii), the representation of a tensor u given in (B.1) can be rewritten as

$$u = \sum_{i=1}^{n} x_i \otimes y_i.$$

The primary purpose of tensor products is to linearise bilinear mappings.

Proposition B.1.3. *For every bilinear map $A : X \times Y \to Z$, there exists a unique linear mapping $\tilde{A} : X \otimes Y \to Z$ such that $A(x, y) = \tilde{A}(x \otimes y)$ for all $x \in X$ and $y \in Y$. Thus, $L(X \otimes Y, Z)$ is isomorphic to $B(X \times Y, Z)$. In particular, if $Z = \mathbb{R}$, then we have $B(X \times Y) = (X \otimes Y)^*$.*

In other words, the special bilinear map $(x, y) \to x \otimes y$ acts as a 'universal' bilinear map, i.e. every other bilinear map on $X \times Y$ factors through this one via a linear mapping on the tensor product. This is graphically represented by Figure B.1.

An important consequence of this result is that the tensor product $X \otimes Y$ of vector spaces X and Y always exists and is unique up to isomorphism.

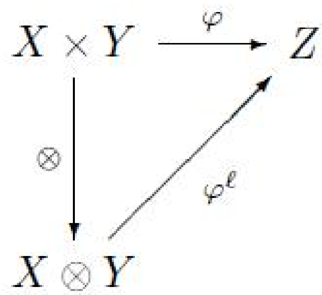

Figure B.1: Tensor maps

B.2 Norms on tensor products

Next we consider the question of how to define a norm on the tensor product of two Banach spaces? Let X and Y be Banach spaces. It is natural to require that

$$||x \otimes y|| \leq ||x||\,||y||.$$

Let $u \in X \otimes Y$. If $u = \sum_{i=1}^{n} x_i \otimes y_i$, then it follows from the triangle inequality that

$$||u|| \leq \sum_{i=1}^{n} ||x_i||\,||y_i||.$$

This must hold for each representation of u, thus

$$||u|| \leq \inf\{\sum_{i=1}^{n} ||x_i||\,||y_i||\},$$

where the infimum is taken over all representations of u.

Definition B.2.1. The *projective norm* is defined as

$$\pi(u) = \inf\left\{\sum_{i=1}^{n} ||x_i||\,||y_i|| : u = \sum_{i=1}^{n} x_i \otimes y_i\right\}.$$

Proposition B.2.2. *Let X and Y be Banach spaces. Then, π is a norm on $X \otimes Y$ and $\pi(x \otimes y) = \|x\|\,\|y\|$ for $x \in X$ and $y \in Y$.*

The tensor product endowed with the projective norm π is denoted by $X \otimes_\pi Y$. Unless X and Y are finite-dimensional, this space is not complete. Thus, we denote its completion by $X \widetilde{\otimes}_\pi Y$.

Now, let E be a Banach lattice and Y a Banach space. A norm introduced by Chaney and Schaefer, which creates what is called an l-tensor product, is defined by

$$\|u\|_l = \inf\left\{\left\| \sum_{i=1}^n \|b_i\|\,|a_i|\right\| \;:\; u = \sum_{i=1}^n a_i \otimes b_i\right\} \text{ for all } u \in E \otimes Y.$$

Then $\|\cdot\|_l$ is a norm on $E \otimes Y$ such that $\|a \otimes b\|_l = \|a\|\,\|b\|$ for all $a \in E$ and $b \in Y$. Moreover, if Y is a Banach lattice, then the norm completion $E\widetilde{\otimes}_l Y$ of $E \otimes Y$ with respect to $\|\cdot\|_l$ is a Banach lattice, with positive cone $(E\widetilde{\otimes}_l Y)_+$ given by the closure of the projective cone

$$E_+ \otimes Y_+ = \Big\{\sum_{i=1}^n x_i \otimes b_i : n \in \mathbb{N}, x_1, \cdots x_n \in E_+ \text{ and } b_1, \cdots b_n \in Y_+\Big\},$$

with respect to $\|\cdot\|_l$.

B.3 The Bochner space $L_p(P, Y)$ as a tensor product

Let $(\Omega, \mathcal{F}, P)$ denote a probability space.

A function $s : \Omega \to Y$ is *simple* if there exist $y_1, y_2, \ldots, y_n \in Y$ and sets $A_1, A_2, \ldots, A_n \in \mathcal{F}$ such that $s = \sum_{i=1}^n y_i \chi_{A_i}$. Here, χ_{A_i} denotes the characteristic function of A_i, given by $\chi_{A_i}(\omega) = 1$ when $\omega \in A_i$ and $\chi_{A_i}(\omega) = 0$ when $\omega \notin A_i$.

A function $f : \Omega \to Y$ is called *P-measurable* if there exists a sequence of simple functions (s_n) with $\lim_{n\to\infty} \|s_n - f\|_Y = 0$ P-a.s..

For $1 \le p < \infty$ and Y a Banach space, let $L_p(P, Y)$ denote the space of (classes of a.e. equal) Bochner p-integrable functions $f : \Omega \to Y$ and denote the *Bochner norm* on $L_p(P, Y)$ by Δ_p, i.e.

$$\Delta_p(f) = \left(\int_\Omega \|f\|_Y^p \, dP\right)^{1/p} \text{ for all } f \in L_p(P, Y).$$

For $p = \infty$, let $L_\infty(P, Y) = \{f : \Omega \to Y \text{ } P\text{-measurable } : ess\sup \|f\| < \infty\}$ and

$$\Delta_\infty(f) = ess\sup \|f\| \text{ for all } f \in L_\infty(P, Y).$$

Consider the bilinear map $\psi : L_p(P) \times Y \to L_p(P, Y)$, given by

$$\psi(f, y)(\omega) = f(\omega)y \quad \text{for all } \omega \in \Omega.$$

Then the induced linear map $\psi^l : L_p(P) \otimes Y \to L_p(P, Y)$ is described by

$$\psi^l(f \otimes y)(\omega) = f(\omega)y \quad \text{for all } \omega \in \Omega.$$

The map ψ^l is injective. Thus, $L_p(P,Y)$ contains a copy of $L_p(P) \otimes Y$ and we may induce the Bochner norm. The normed space $(L_p(P) \otimes Y, ||\cdot||_p)$ is denoted by $L_p(P) \otimes_{\Delta_p} Y$.

Chaney and Schaefer extended $L_p(P,Y)$-spaces by means of an appropriate tensor product of a Banach lattice E and a Banach space Y. They achieved this by using the well known fact that $L_p(P,Y)$ is isometrically isomorphic to the norm completion $L_p(P)\widetilde{\otimes}_{\Delta_p} Y$ of $L_p(P) \otimes_{\Delta_p} Y$, where Δ_p is a reasonable cross norm on $L_p(P) \otimes Y$ (see [52]). It is known that if E is a Banach lattice and Y a Banach space, and if the tensor product $E \otimes Y$ of E and Y is endowed with the norm $\|\cdot\|_l$, then

(i) $\|a \otimes b\|_l = \|a\| \, \|b\|$ for all $a \in E$ and $b \in Y$ (see [23, 31, 110, 111, 142]), and

(ii) if $E = L_p(P)$ and $1 \leq p < \infty$, then $\|\cdot\|_l = \Delta_p$ (see [23, 31, 111, 142]).

Appendix C

Appendix: Convex analysis

In this section, we will introduce some notions and results on convex sets and functions. Consider a vector space $\mathcal{X}$.

Definition C.1.

(i) A set A is *convex* if for all $\lambda \in [0,1]$ and $x, y \in A$, $\lambda x + (1-\lambda)y \in A$.

(ii) A set C in a vector space is said to be a *cone with vertex at the origin* if $x \in C$ implies that $\alpha x \in C$ for all $\alpha \geq 0$. A *cone with vertex* p is defined as a translation $p + C$ of a cone C with vertex at the origin. If the vertex of a cone is not explicitly mentioned then it is assumed to be the origin.

(iii) The *convex hull* of the subset A is the set

$$\text{co}A = \bigcap \{C \subseteq \mathcal{X} : A \subseteq C, C \text{ convex}\}.$$

(iv) The *closed convex hull* of the set $A \subseteq \mathcal{X}$ is the set $\overline{\text{co}}A = \text{cl}(\text{co}A)$, i.e. the smallest closed convex set containing A.

The convex hull can also be represented as

$$\text{co}A = \Big\{ \sum_{i=1}^{n} \lambda_i x_i : n \in \mathbb{N},\, (\lambda_i) \subseteq \mathbb{R}^+,\, (x_i) \subseteq A,\, \sum_{i=1}^{n} \lambda_i = 1 \Big\}.$$

We will denote by $\text{int}(A)$ the interior of a subset A with respect to the norm-topology, i.e. the open set which is the union of all open subsets of A. By $\text{core}(A)$ we denote the algebraic interior, i.e. the set of all points $x \in A$ with the property that for every $y \in A$, there exists $\epsilon > 0$ such that $x + ty \in A$ for all $t \in [0, \epsilon]$. We always have $\text{int}(A) \subseteq \text{core}(A)$.

Theorem C.2. *Let C be a convex subset of the topological vector space $\mathcal{X}$. Then the following holds.*

(i) The closure of C, denoted by $\mathrm{cl}(C)$, is convex.

(ii) The interior of C is convex.

(iii) If $\mathrm{int}(C) \neq \emptyset$, then $\mathrm{cl}(\mathrm{int}(C)) = \mathrm{cl}(C)$ and $\mathrm{int}(\mathrm{cl}(C)) = \mathrm{int}(C)$.

(vi) If $\mathrm{int}(C) \neq \emptyset$, then $\mathrm{core}(C) = \mathrm{int}(C)$.

Theorem C.3 (Eidelheit)**.** *Let A and B be two non-empty convex subsets of the topological vector space $\mathcal{X}$. If $\mathrm{int}(A) \neq \emptyset$ and $B \cap \mathrm{int}(A) = \emptyset$, then there exist $x^* \in \mathcal{X}^* \backslash \{0\}$ and $\alpha \in \mathbb{R}$ such that for all $x \in A$ and $y \in B$*

$$\langle x, x^* \rangle \leq \alpha \leq \langle y, x^* \rangle ,$$

or equivalently,

$$\sup_{x \in A} x^*(x) \leq \inf_{y \in B} x^*(y).$$

In the case of locally convex spaces, we have the following.

Theorem C.4. *Let $\mathcal{X}$ be a locally convex space and $A, B \subseteq \mathcal{X}$ be two non-empty convex sets. If A is closed, B is compact and $A \cap B = \emptyset$, then there exist $x^* \in \mathcal{X}^* \backslash \{0\}$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ such that for all $x \in A$ and $y \in B$*

$$\langle x, x^* \rangle \leq \alpha_1 < \alpha_2 \leq \langle y, x^* \rangle ,$$

or equivalently,

$$\sup_{x \in A} x^*(x) < \inf_{y \in B} x^*(y).$$

Definition C.5. Let A be a subset of the normed space $\mathcal{X}$. The *positive polar cone* of A is given by

$$A^\circ = \{x^* \in \mathcal{X}^* : \langle x, x^* \rangle \geq 0 \text{ for all } x \in A\}.$$

and the *negative polar cone* of A by

$$A^\circ_- = \{x^* \in \mathcal{X}^* : \langle x, x^* \rangle \leq 0 \text{ for all } x \in A\}.$$

Both A° and A°_- are nonempty, convex cones.

Next, we define some notations and definitions for extended real-valued convex functions. For a function $f : \mathcal{X} \to [-\infty, \infty]$, we denote the *domain* of f by

$$\text{dom} f = \{x \in \mathcal{X} : f(x) < \infty\}$$

and the *epigraph* of f by

$$\text{epi} f = \{(x, t) \in \mathcal{X} \times \mathbb{R} : f(x) \leq t\}.$$

Definition C.6.

(i) The function $f : \mathcal{X} \to [-\infty, \infty]$ is *proper* if $\text{dom} f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in \mathcal{X}$.

(ii) The function $f : \mathcal{X} \to [-\infty, \infty]$ is *convex* if for all $x, y \in \text{dom} f$ and $\lambda \in (0, 1)$

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

(iii) The function $f : \mathcal{X} \to [-\infty, \infty]$ is *lower semi-continuous* if $\{x \in E : f(x) > k\}$ is an open set for all $k \in \mathbb{R}$.

(iv) The function $f : \mathcal{X} \to [-\infty, \infty]$ is *subdifferentiable* if for all $x \in E$, the set

$$\partial f(x) = \{r^* \in E^* \ : f(x) - f(y) \leq r^*(x - y) \ \text{ for all } y \in E\},$$

is non-empty.

Note that a function is lower semi-continuous if for any net $(x_\alpha) \subseteq E$ converging to some $x \in E$, $f(x) \leq \liminf_\alpha f(x_\alpha)$. This definition is equivalent to (iii) above.

The following theorem states some properties of convex functions. For a proof of this theorem see [169, Theorem 2.1.1].

Theorem C.7. *Let $f : \mathcal{X} \to [-\infty, \infty]$. The following statements are equivalent.*

(i) The function f is convex.

(ii) The domain of f is a convex set.

(iii) The epigraph of f is a convex subset of $\mathcal{X} \times \mathbb{R}$.

Theorem C.8. *If $f_i : \mathcal{X} \to [-\infty, \infty]$ is convex for every $i \in I$ ($I \neq \emptyset$), then $\sup_{i \in I} f_i$ is convex. Moreover,*

$$\text{epi}(\sup_{i \in I} f_i) = \bigcap_{i \in I} \text{epi} f_i.$$

Definition C.9. The *lower semi-continuous envelope* or *lower semi-continuous regularisation* $\bar{f}$ of the function $f : \mathcal{X} \to [-\infty, \infty]$ is defined by

$$\bar{f}(x) = \inf\{t \in \mathbb{R} : (x,t) \in \mathrm{cl}(\mathrm{epi} f)\}.$$

Since $\mathrm{cl}(\mathrm{epi} f)$ is closed, we have that $\mathrm{epi}\bar{f} = \mathrm{cl}(\mathrm{epi} f)$.

Theorem C.10. *Let $f : \mathcal{X} \to [-\infty, \infty]$ be a convex function. Then the following holds.*

(i) The function $\bar{f}$ is convex.

(ii) If $g : \mathcal{X} \to [-\infty, \infty]$ is convex, lower semi-continuous and $g \leq f$, then $g \leq \bar{f}$.

(iii) The function $\bar{f}$ does not take the value $-\infty$ if and only if f is bounded from below by a continuous affine function.

The lower semi-continuous and convex function, which is naturally associated with the function $f : \mathcal{X} \to [-\infty, \infty]$ is called the *lsc convex hull* of f, is denoted by $\overline{\mathrm{co}} f$ and defined by $\mathrm{epi}(\overline{\mathrm{co}} f) = \mathrm{cl}(\mathrm{co}(\mathrm{epi} f))$. We have that $\overline{\mathrm{co}} f \leq \bar{f} \leq f$.

Theorem C.11. *If the convex function f is bounded above on a neighbourhood of a point of its domain, then f is continuous on the interior of its domain. Moreover, if f is not proper, then f is identically $-\infty$ on* $\mathrm{int}(\mathrm{dom} f)$.

Corollary C.12. *Let $f : \mathcal{X} \to [-\infty, \infty]$ be a convex function. Then f is continuous on* $\mathrm{int}(\mathrm{dom} f)$ *if and only if* $\mathrm{int}(\mathrm{epi} f)$ *is non-empty in $\mathcal{X} \times \mathbb{R}$.*

The following result is proved by Ekeland [56, Proposition I.2.5].

Proposition C.13 ([56])**.** *Let $f : \mathcal{X} \to [-\infty, \infty]$ be a convex function. The following statements are equivalent:*

(i) There exists a non-empty open set $\mathcal{O}$ on which f is not everywhere equal to $-\infty$ and is bounded above by a constant $c < \infty$.

(ii) f is a real-valued function, and it is continuous over the interior of its effective domain, which is not empty.

The next proposition is taken from [139, Proposition 3.1].

Proposition C.14. *Suppose that $\mathcal{X}$ is a Banach lattice and $f : \mathcal{X} \to [-\infty, \infty]$ is proper, convex and monotonic. Then, f is continuous and subdifferentiable on the interior of its domain.*

Definition C.15. The *conjugate* $f^* : \mathcal{X}^* \to [-\infty, \infty]$ of a function $f : \mathcal{X} \to [-\infty, \infty]$ is given by

$$f^*(x^*) = \sup_{x \in \mathcal{X}} \{\langle x, x^* \rangle - f(x)\}.$$

The conjugate is sometimes also known as the *Fenchel conjugate*. In the next theorem, we state some properties of conjugate functions.

Theorem C.16. *Let $f, g : \mathcal{X} \to [-\infty, \infty]$.*

(i) The conjugate f^* is convex.

(ii) The Young-Fenchel inequality holds, i.e. for all $x \in \mathcal{X}$ and $x^* \in \mathcal{X}^*$

$$f(x) + f^*(x^*) \geq \langle x, x^* \rangle .$$

(iii) $f \leq g$ implies that $g^* \leq f^*$.

(iv) $f^* = \bar{f}^* = (\overline{\mathrm{co}} f)^*$ and $f^{**} \leq \overline{\mathrm{co}} f \leq \bar{f} \leq f$.

Theorem C.17. *Let $f : \mathcal{X} \to \overline{\mathbb{R}}$ be proper, lower semi-continuous and convex. Then f^* is lower semi-continuous in the weak* topology (w*-lsc), proper and convex and $f^{**} = f$.*

Theorem C.18. *Let $f : \mathcal{X} \to \overline{\mathbb{R}}$ have non-empty domain. Let $\overline{\mathrm{co}} f$ denote the lower semi-continuous convex hull of f and $\bar{f}$ the lower semi-continuous hull of f, i.e.* $\mathrm{epi}\bar{f} = \mathrm{cl}(\mathrm{epi} f)$.

(i) If $\overline{\mathrm{co}} f$ is proper, then $f^{**} = \overline{\mathrm{co}} f$; if $\overline{\mathrm{co}} f$ is not proper, then $f^{**} = -\infty$.

(ii) Suppose that f is convex. If f is lower semi-continuous at $\bar{x} \in \mathrm{dom} f$, then $f(\bar{x}) = f^{**}(\bar{x})$; moreover if $f(\bar{x}) \in \mathbb{R}$, then $f^{**} = \bar{f}$ and $\bar{f}$ is proper.

Proof. (i): The function $\overline{\mathrm{co}} f$ is convex and lsc. If $\overline{\mathrm{co}} f$ is proper, using Theorem C.17 and Theorem C.16 (iv), we have that

$$\overline{\mathrm{co}} f = (\overline{\mathrm{co}} f)^{**} = f^{**}.$$

If $\overline{\mathrm{co}} f$ is not proper, since $\mathrm{dom}(\overline{\mathrm{co}} f) \supseteq \mathrm{dom} f \neq \emptyset$, $\overline{\mathrm{co}} f$ takes the value $-\infty$. Hence, $f^* = (\overline{\mathrm{co}} f)^* = \infty$, and so $f^{**} = -\infty$.

(ii): Since f is convex, we have $\overline{\text{co}}f = \bar{f}$. Also, since f is lower semi-continuous at $\bar{x}$, we have that $\bar{f}(\bar{x}) = f(\bar{x})$. If $f(\bar{x}) = -\infty$, then it is obvious that $f^{**}(\bar{x}) = f(\bar{x})$. Let $f(\bar{x}) \in \mathbb{R}$. Then, $\bar{f}(\bar{x}) \in \mathbb{R}$ and so $\bar{f}$ is proper. From the first part, we have that $f^{**} = f$, as $f^{**}(\bar{x}) = \bar{f}(\bar{x}) = f(\bar{x})$. □

For any convex functional f defined on a convex set C in a vector space $\mathcal{X}$, define $[f, C]$ by

$$[f, C] = \{(r, x) \in \mathbb{R} \times \mathcal{X} : x \in C, f(x) \leq r\}.$$

Note that $[f, C]$ is a convex set in $\mathbb{R} \times \mathcal{X}$. If you think of the $\mathbb{R}$ axis as being the vertical axis in the $\mathbb{R} \times \mathcal{X}$ space, then the set $[f, C]$ can be thought of as the region above the graph of f. The set $[f, C]$ is sometimes called the *epigraph* of f over C.

Proposition C.19. *The function f defined on the convex domain C is convex if and only if $[f, C]$ is a convex set.*

For a proof, see [119].

Next, we need to analyse, when this set $[f, C]$ contains interior points.

Proposition C.20. *If f is a convex function on the convex domain C in a normed space and C has nonempty relative interior int(C), then the convex set $[f, C]$ has a relative interior point (r_0, x_0) if and only if f is continuous at the point $x_0 \in$ int(C).*

Since the utility function is a concave function, the theory for concave functions is considered. Given a concave function g defined on a convex set D of a vector space, define the set

$$[g, D] = \{(r, x) : x \in D, r \leq g(x)\}.$$

The set $[g, D]$ is convex and the result of Proposition C.20 can be extended to it.

Definition C.21. Let f be a convex function and g a concave function defined, respectively, on the convex sets C and D in the normed space $\mathcal{X}$. The *conjugate sets* C^* and D^* are defined by

$$C^* = \{x^* \in \mathcal{X}^* : \sup_{x \in C}\{\langle x, x^* \rangle - f(x)\} < \infty\},$$

$$D^* = \{x^* \in \mathcal{X}^* : \inf_{x \in D}\{\langle x, x^* \rangle - g(x)\} > -\infty\}.$$

The functions f^* and g^* *conjugate* to f and g respectively are defined by

$$f^*(x^*) = \sup_{x\in C}\{\langle x, x^*\rangle - f(x)\},$$
$$g^*(x^*) = \inf_{x\in D}\{\langle x, x^*\rangle - g(x)\}.$$

It is relatively easy to verify that both C^*, f^* and D^* are convex and g^* is concave.

Applications of these conjugate functions to optimisation problems are looked at next. The problem under consideration is

$$\inf_{C\cap D}\{f(x) - g(x)\},$$

where f is convex over C and g is concave over D. In standard minimisation problems, g is usually zero.

The following result plays an important role in the thesis.

Theorem C.22 (Fenchel duality theorem)**.** *Let $\mathcal{X}$ be a normed space and assume that*

(i) f is a convex function on the convex set $C \subseteq \mathcal{X}$,

(ii) g is a concave functions on the convex sets $D \subseteq \mathcal{X}$,

(iii) $C \cap D$ contains points in the relative interior of C and D,

(iv) either $[f, C]$ or $[g, D]$ has nonempty interior, and

(v) $\inf_{x\in C\cap D}\{f(x) - g(x)\}$ is finite.

Then

$$\inf_{x\in C\cap D}\{f(x) - g(x)\} = \max_{x^*\in C^*\cap D^*}\{g^*(x^*) - f^*(x^*)\},$$

where the maximum on the right is attained by some $x_0^* \in C^* \cap D^*$.

Moreover, if the infimum on the left is attained by some $x_0 \in C \cap D$, then

$$\max_{x\in C}\{\langle x, x_0^*\rangle - f(x)\} = \langle x_0, x_0^*\rangle - f(x_0)$$

and

$$\min_{x\in D}\{\langle x, x_0^*\rangle - g(x)\} = \langle x_0, x_0^*\rangle - g(x_0).$$

See [119] for a proof of this theorem.

The problem we will be considering requires finding the supremum of a concave function. Thus, setting $f(x) = 0$ in Fenchel's duality theorem above, yields

$$\sup_{x\in D} g(x) = \min_{x^*\in D^*} -g^*(x^*). \tag{C.1}$$

Note as well that in our problem D is a convex cone and hence D^* will be the polar cone of D.

Bibliography

[1] Ando, T. (1960): Linear functionals on Orlicz spaces. *Nieuw Archief voor Wiskunde,* 3, 1 - 16.

[2] Ansel, J.P. and C. Stricker (1990): Quelques remarques sur un théorème de Yan. *Séminaire de Probabilités,* 24, 266 - 274.

[3] Ansel, J.P. and C. Stricker (1994): Couverture des actifs contingents et prix maximum. *Annales de l'I.H.P* 30, 303 - 315.

[4] Artzner, P., F. Delbaen, J.M. Eber and D. Heath (1996): A characterisation of measures of risk. *Technical Report No 1186, Cornell University.*

[5] Artzner, P., F. Delbaen, J.M. Eber and D. Heath (1999): Thinking coherently. *RISK,* 10, 68 - 71.

[6] Artzner, P., F. Delbaen, J.M. Eber and D. Heath (1999): Coherent measures of risk. *Mathematical Finance,* 9, 203 - 228.

[7] Artzner, P. and D. Heath (1995): Approximate completeness with multiple martingale measures. *Mathematical Finance,* 5, 1 - 11.

[8] Ash, R.B. (1972): Measure, integration and functional analysis. *Academic Press, Inc.*

[9] Bachelier, L. (1900): Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure,* 3, 21 - 86.

[10] Badescu, A., R.J. Elliott and T.K. Siu (2009): Esscher transforms and consumption-based models. *Insurance: Mathematics and Economics,* 45, 337 - 347.

[11] Bellini, F. and M. Frittelli (2002): On the existence of minimax martingale measures. *Mathematical Finance,* 12, 1 - 21.

[12] Biagini, S. and M. Frittelli (2004): On the super-replication price of unbounded claims. *The Annals of Applied Probability,* 14, 1970 - 1991.

[13] Biagini, S. and M. Frittelli (2005): Utility maximisation in incomplete markets for unbounded processes. *Finance and Stochastics,* 9, 493 - 517.

[14] Biagini, S. (2008): An Orlicz spaces duality for utility maximisation in incomplete markets. *Progress in Probability,* 59, 445 - 455.

[15] Biagini, S. and M. Frittelli (2008): A unified framework for utility maximisation problems: An Orlicz space approach. *Annals of Applied Probability,* 18, 929 - 966.

[16] Biagini, S. and M. Frittelli (2009): On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures. *Optimality and Risk - Modern Trends in Mathematical Finance,* 1 - 28.

[17] Biagini, S., M. Frittelli and M. Grasselli (2010): Indifference price with general semimartingales. *Mathematical Finance,* 21, 423 - 446.

[18] Birnbaum, Z. and W. Orlicz (1931): Über die Verallgemeinerung des Begriffes der zueinander Konjugierten Potenzen. *Studia Mathematica,* 3, 1 - 67.

[19] Bombal, F. (1987): On l_1 subspaces of Orlicz vector-valued function spaces. *Math. Proc. Cambridge Philos. Soc.,* 101, 107 - 112.

[20] Borwein, J.M. (1987): Automatic continuity and openness of convex relations. *Proceedings of the American Mathematical Society,* 99, 49 - 55.

[21] Bühlmann, H. (1980): An economic premium principle. *ASTIN Bulletin*, 11, 52 - 60.

[22] Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci and L. Montrucchio (2010): Risk measures: Rationality and diversification. *Mathematical Finance,* 21, 743 - 774.

[23] Chaney, J. (1972): Banach lattices of compact maps. *Mathematische Zeitschrift,* 129, 1 - 19.

[24] Chateauneuf, A., R. Kast and A. Lapied (1996): Choquet pricing for financial markets with frictions. *Mathematical Finance,* 6, 323 - 330.

[25] Cheridito, P., F. Delbaen and M. Kupper (2006): Coherent and convex monetary risk measures for bounded càdlàg proccesses. *Finance and Stochastics,* 10, 427 - 448.

[26] Cheridito, P. and T. Li (2007): Risk measures on Orlicz hearts. *Mathematical Finance,* 19, 189 - 214.

[27] Cheridito, P. and T. Li (2008): Dual characterisation of properties of risk measures on Orlicz hearts. *Mathematics and Financial Economics,* 2, 29 - 55.

[28] Chou, C.S. (1977): Caractérisation d'une classe de semimartingales. *Seminaire de Probabilités,* 13, 250 - 252.

[29] Choquet, G. (1954): Theory of capacities. *Annals de l'Institute Fourier,* 5, 131 - 295.

[30] Cullender, S.F. and C.C.A Labuschagne (2006): A description of norm-convergent martingales on vector valued L_p-spaces. *Journal of Mathematical Analysis and Applications,* 323, 119 - 130.

[31] Cullender, S.F. and C.C.A. Labuschagne (2008): Convergent martingales of operators and the Radon Nikodým property in Banach spaces. *Proceedings of the American Mathematical Society,* 136, 3883 - 3893.

[32] Cullender, S.(2007): Generalised martingale and stopping time techniques in Banach spaces. *PhD thesis, University of the Witwatersrand.*

[33] Cvitanic, J. and I. Karatzas (1998): On dynamic measures of risk. *Finance and Stochastics,* 3, 451 - 482.

[34] Davis, M.H.A. and A.R. Norman (1990): Portfolio selection with transaction costs. *Mathematics of Operations Research,* 15, 676 - 713.

[35] Davis, M.H.A. (1997): Option pricing in incomplete markets. *Mathematics of Derivative Securities*, Cambridge University Press.

[36] Dalang, R.C., A. Morton and W. Willinger (1990): Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics and Stochastic Reports,* 29, 185 - 201.

[37] De Jonge, E. (1977): The semi-M-property for normed Riesz spaces. *Compositio Mathematica,* 34, 147 - 172.

[38] Delbaen, F. (1992): Representing martingale measures when asset prices are continuous and bounded. *Mathematical Finance,* 2, 107 - 130.

[39] Delbaen, F. and W. Schachermayer (1994): A general version of the fundamental theorem of asset pricing. *Mathematische Annalen,* 300, 463 - 520.

[40] Delbaen, F. and W. Schachermayer (1998): The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen,* 312, 215 - 250.

[41] Delbaen, F. and W. Schachermayer (2006): The mathematics of arbitrage. *Springer, Berlin-Heidelberg.*

[42] Delbaen, F. (2002): Coherent Risk measures on general probability spaces. *Advances in Finance and Stochastics,* 39 - 56.

[43] Denneberg, D. (1994): Non-Additive Measure and Integral. *Kluwer Academic Publishers, Boston.*

[44] Dhaene, J., M. Denuit, M.J. Goovaerts, R. Kaas and D. Vyncke (2002): The concept of comonotonicity in acturial science and finance: theory. *Insurance: Mathematics and Economics,* 31, 3 - 33.

[45] Dinculeanu, N. (1966): Integral representation of linear operators I. *Stud. Cerc. Mat.,* 18, 349 - 385.

[46] Dinculeanu, N. (1966): Integral representation of linear operators II. *Stud. Cerc. Mat.,* 18, 483 - 536.

[47] Diestel, J. (1968): An approach to the theory of Orlicz spaces of Lebesgue-Bochner measurable functions and to the theory of Orlicz spaces of finitely additive vector-valued set functions with applications to the representation of multilinear continuous operators. *PhD thesis, The Catholic University of America.*

[48] Diestel, J. (1970): An approach to the theory of Orlicz spaces of Lesbesgue-Bochner measurable functions. *Mathematische Annalen,* 186, 20 - 33.

[49] Diestel, J. (1970): On the representation of bounded, linear operators from Orlicz spaces of Lebesgue-Bochner measurable functions to any Banach space. *Bull. Acad. Polon. Sci. Sr. Sci. Math. Astronom. Phys.,* 18, 375 - 378.

[50] Diestel, J. (1971): Some remarks on subspaces of Orlicz spaces of vector-valued finitely additive functions. *Studia Mathematica,* 39, 161 - 164.

[51] Diestel, J., J.H. Fourie and J. Swart (2008): The metric theory of tensor products. Grothendieck's résumé revisited. *American Mathematical Society,* Providence, Rhode Island.

[52] Diestel J. and J.J. Uhl (1977): Vector Measures. *American Mathematical Society Surveys,* 15, Providence, Rhode Island.

[53] Dybvig, P.H. and S.A. Ross (1987): Arbitrage. *J. Eatwell, M. Milgate and P. Neuman, eds., The New Palgrave: a Dictionary of Economics*, London, Macmillan, 100 - 106.

[54] Edgar, G.A. and L. Sucheston (1992): Stopping times and directed processes. *Cambridge University Press.*

[55] Embrechts, P., A. McNeil and D. Straumann (2002): Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond, Cambridge University Press.*

[56] Ekeland, I. and R. Temam (1976): Convex analysis and variational problems. *North-Holland Publishing Company, Amsterdam.*

[57] El Karoui, N. and C. Ravanelli (2009): Cash sub-additive risk measures under interest rate ambiguity. *Mathematical Finance,* 19, 561 - 590.

[58] Émery, M. (1980): Compensation de processus à variation finie non localement intégrables. *Seminaire de Probabilités,* 14, 152 - 160.

[59] Esscher, F. (1932): On the probability function in the collective theory of risk. *Skandinavisk Aktuarietidskrift,* 15, 175 - 95.

[60] Föllmer, H. and A. Schied (2002): Convex measures of risk and trading constraints. *Financial and Stochastics,* 6, 429 - 447.

[61] Föllmer, H. and A. Schied (2002): Stochastic finance: An introduction in discrete time. *Walter de Gruyter, Berlin.*

[62] Frittelli, M. and E. Rosazza Gianin (2002): Putting order in risk measures. *Journal of Banking and Finance,* 26, 1473 - 1486.

[63] Frittelli, M. (2004): Some remarks on arbitrage and preferences in securities market models. *Mathematical Finance,* 14, 351 - 357.

[64] Frittelli, M. (2007): No arbitrage and preferences. *Instituto Lombardo - Accademia di Scienze e Lettere,* 179 - 199.

[65] Frittelli M. and M. Maggis (2011): Dual representation of quasiconvex conditional maps. *SIAM Journal of Financial Mathematics,* 2, 357 - 382.

[66] Gerber, H.U. (1974): On additive premium calculation principles. *Astin Bulletin,* 7, 215 - 222.

[67] Gerber, H.U. and E.S.W. Shiu (1994): Option pricing by Esscher transforms. *Transactions of Society of Actuaries,* 46, 99 - 140.

[68] Goovaerts, M.J., F. De Vijlder and J. Haezendonck (1984): Insurance Premiums. *North Holland Publishing,* Amsterdam.

[69] Goovaerts, M.J., R. Kaas, R.J.A. Laeven and Q. Tang (2004): A comonotonic image of independence for additive risk measures. *Insurance: Mathematics and Economics,* 35, 581 - 594.

[70] Goovaerts, M.J. and R.J.A. Laeven (2008): Actuarial risk measures for financial derivative pricing. *Insurance: Mathematics and Economics,* 42, 540 - 547.

[71] Göpfert, A., H. Riahi, C. Tammer and C. Zalinescu (2003): Variational methods in partially ordered spaces. *Springer-Verlag.*

[72] Greco, G. (1982): Sulla rappresentazione di funzionali mediante integrali. *Rendiconti del Seminario Matematico della Universit di Padova,* 66, 21 - 42.

[73] Greub W.H. (1967): Multilinear algebra. *Die Grundlehren der Mathematischen Wissenschaften, Band 136,* Springer-Verlag New York.

[74] Grothendieck, A. (1973): Topological Vector Spaces. *Gordon and Breach, Science Publishers Ltd.*

[75] Halmos, P.R. and L.J. Savage (1949): Application of the Radon-Nikodým theorem to the theory of sufficient statistics. *Annals of Mathematical Statistics,* 20, 225 - 241.

[76] Hamada, M. and M. Sherris (2003): Contingent claim pricing using probability distortion operators. *Applied Mathematical Finance,* 10, 19 - 47.

[77] Hamel. A.H. (2009): A duality theory for set-valued functions I: Fenchel Conjugation Theory. *Set-valued Analysis,* 17, 153 - 182.

[78] Hamel, A.H. and F. Heyde (2010): Duality for set-valued measures of risk. *SIAM Journal of Financial Mathematics,* 1, 66 - 95.

[79] Hamel, A.H., F. Heyde and M. Höhne (2007): Set-valued measures of risk. *Technical report, Princeton University.*

[80] Hamel, A.H. and W. Rudloff (2008): Continuity and finite-valuedness of set-valued risk measures. Festschrift in celebration of Prof. Dr, Wilfried Grecksch's 60th birthday. *Shaker-Verlag, Germany,* 49 - 64.

[81] Harrison, J.M. and D.M. Kreps (1979): Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory,* 20, 381 - 408.

[82] Harrison, J.M. and S.R. Pliska (1981): Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications,* 11, 215 - 260.

[83] Harrison, J.M. and S.R. Pliska (1983): A stochastic calculus model of continuous trading: complete markets. *Stochastic Processes and their Applications,* 15, 313 - 316.

[84] Heath, D. (2000): Back to the future. *Plenary lecture at the First World Congress of the Bachelier Society, Paris.*

[85] Henderson, V. and D. Hobson (2004): Utility indifference pricing - An overview. *Volume of Indifference Pricing, Princeton University Press.*

[86] Jacod, J. and A.N. Shiryaev (2003): Limit theorems for stochastic processes, 2nd Edition. *Springer-Verlag.*

[87] Jacod, J. and M. Yor (1977): Étude de solutions extrémales et représentation intégrale des solutions pour certains problèmes de martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete,* 38, 83 - 125.

[88] Jarrow, R.A., X. Jin and D.B. Madan (1999): The second fundamental theorem of asset pricing. *Mathematical Finance,* 9, 255 - 273.

[89] Jaschke, S. and U. Küchler (2001): Coherent risk measures and good-deal bounds. *Finance and Stochastics* 5, 181 - 200.

[90] Jeurnink, G.A.M. (1982): Integration of functions with values in a Banach lattice. *PhD thesis, University of Nijmegen,* Netherlands.

[91] Jouini E., M. Meddeb and N. Touzi (2004): Vector-valued coherent risk measures. *Finance and Stochastics,* 8, 531 - 552.

[92] Global Derivatives Study Group (1993): Derivatives: Practices and principles.

[93] Kabanov, Y.M. (1997): On the FTAP of Kreps-Delbaen-Schachermayer. *Statistics and control of stochastic processes. The Liptser Festschrift. Proceedings of Steklov Mathematical Institute Seminar, World Scientific, Singapore,* 191 - 203.

[94] Kakutani, S. (1941): Concrete representation of abstract (L)-spaces and the mean ergodic theorem. *Annals of Mathematics,* 42, 523 - 537.

[95] Kakutani, S. (1941): Concrete representation of abstract (M)-spaces. *Annals of Mathematics,* 42, 994 - 1024.

[96] Karatzas, I., J.P. Lehoczky, S.E. Shreve and G. Xu (1991): Martingale and duality methods for utility maximisation in an incomplete market. *SIAM Journal of Control and Optimisation,* 29, 702 - 730.

[97] Kijima, M. and Y. Muromachi (2007): An extension of the Wang transform derived from Bühlmanns economic premium principle for insurance risk. *Insurance: Mathematics and Economics,* 42, 887 - 896.

[98] Klebaner, F.C. (2007): Introduction to stochastic calculus with applications. *Imperial College Press.*

[99] Klein, I. and W. Schachermayer (1996): A quantitative and a dual version of the Halmos-Savage theorem with applications to mathematical finance. *The Annals of Probability,* 24, 867 - 881.

[100] Klein, I. (2006): A comment on market free lunch and free lunch. *Mathematical Finance,* 16, 583 - 588.

[101] Konstantinides, D.G. and C.E. Kountzakis (2011): Risk measures in ordered normed linear spaces with non-empty cone-interior. *Insurance: Mathematics and Economics,* 48, 111 - 122

[102] Kozek, A. (1977): Orlicz spaces of functions with values in Banach spaces. *Annales Societatis Mathematicae Polonae, Series 1: Commentationes Mathematicae XIX.*

[103] Kozek, A. (1979): Convex integral functionals on Orlicz spaces. *Annales Societatis Mathematicae Polonae, Series 1: Commentationes Mathematicae XXI.*

[104] Kramkov, D. and W. Schachermayer (1998): The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annals of Applied Probability,* 9, 904 - 950.

[105] Kramkov, D. and W. Schachermayer (2003): Necessary and sufficient conditions in the problem of optimal investment in incomplete markets. *Annals of Applied Probability,* 13, 1504 - 1516.

[106] Krasnosel'skii M.A. and Y.B. Rutickii (1961): Convex functions and Orlicz spaces. *Groningen, P. Noordhoff.*

[107] Kreps, D.M. (1981): Arbitrage and equilibrium in economics with infinitely many commodities. *Journal of Mathematical Economics,* 8, 15 - 35.

[108] Kusuoka, S. (1993): A remark on arbitrage and martingale measures. *Publ. RIMS, Kyoto University,* 29, 833 - 840

[109] Labuschagne C.C.A. (1994): A note on the order continuity of the norm of $E\widetilde{\otimes}_m F$. *Archiv der Mathematik,* 62, 335 - 337.

[110] Labuschagne, C.C.A (2004): Riesz reasonable cross norms on tensor products of Banach lattices. *Quaestiones Mathematicae,* 27, 243 - 266.

[111] Labuschagne, C.C.A (2004): Characterizing the one-sided tensor norms Δ_p and ${}^t\Delta_p$. *Quaestiones Mathematicae*, 27, 339 - 363.

[112] Labuschagne, C.C.A (2006): A Dodds-Fremlin property for Asplund and Radon-Nikodým operators. *Positivity,* 10, 391 - 407.

[113] Labuschagne, C.C.A. and T.M. Offwood (2010): A note on the connection between the Esscher-Girsanov transform and the Wang transform. *Insurance: Mathematics and Economics,* 47, 385 - 390.

[114] Labuschagne, C.C.A. and T.M. Offwood (2010): A description of Banach space-valued Orlicz hearts. *Central European Journal of Mathematics,* 8, 1109 - 1119.

[115] Labuschagne, C.C.A. and T.M. Offwood (2010): On the fundamental theorems of asset pricing. *The Thailand Econometrics Society*, 2, 49 - 58.

[116] Labuschagne, C.C.A. and T.M. Offwood (2011): A tensor product approach to representations of set-valued risk measures. *Submitted.*

[117] Labuschagne, C.C.A. and T.M. Offwood (2011): Pricing exotic options using the Wang transform. *North American Journal of Economics and Finance,* in print.

[118] Lindenstrauss, Y. and L. Tzafriri (1979): Classical Banach spaces I and II. *Springer Verlag,* Berlin-Heidelberg-New York.

[119] Luenberger, D.G. (1997): Optimisation by vector space methods. *Wiley.*

[120] Luxemburg, W.A.J. (1955): Banach function spaces. *PhD thesis, Technische Hogeschool te Delft.*

[121] Luxemburg, W.A.J. and A.C. Zaanen (1971): Riesz Spaces, Vol. I. *North-Holland Publishing Company, Amsterdam.*

[122] Margrabe, W. (1978): The value of an option to exchange one asset for another. *The Journal of Finance,* 33, 177 - 186.

[123] McWalter, T.A. (2006): Quadratic criteria for optimal martingale measures in incomplete markets. *MSc thesis, University of the Witwatersrand.*

[124] Merton, R.C. (1992): Continuous-time finance. *Blackwell Publishers.*

[125] Meyer-Nieberg, P. (1991): Banach lattices, *Springer Verlag,* Berlin-Heidelberg-New York.

[126] Nagel, R. (1973): Ordnungsstetigkeit in Banachverbänden. *Manuscripta Math.,* 9, 9 - 27.

[127] Norstad, J. (2005): An introduction to utility theory.
homepage.mac.com/j.norstad/finance/util.pdf

[128] Orlicz, W. (1932): Über eine gewisse Klasse von Räumen vom Typus. *Bulletin of the International Academy of Polisch Science Series A,* 8/9, 207 - 220.

[129] Parker, J.M (1996): A note on comonotonic additivity. *Glasgow Mathematical Journal,* 38, 199 - 205.

[130] Pelsser, A. (2008): On the applicability of the Wang transform for pricing financial risks. *ASTIN Bulletin,* 31, 171 - 181.

[131] Pliska, S.R. (1986): A stochastic calculus model of continuous trading: Optimal portfolios. *Mathematics of Operations Research,* 11, 371 - 382.

[132] Popa, N. (1979): Die Permanenzeigenschaften der Tensorprodukte von Banachverbänden. *Romanian-Finnish Seminar on Complex Analysis, Lecture Notes in Mathematics,* 743, 627 - 647.

[133] Rao, M.M. (1968): Linear functionals on Orlicz spaces: General theory. *Pacific Journal of Mathematics,* 25, 553 - 585.

[134] Rao, M.M. and Z.D. Ren (1991): Theory of Orlicz spaces. *Marcel Dekker Inc., New York.*

[135] Rockafellar, R.T. (1968): Integrals which are convex functionals. *Pacific Journal of Mathematics,* 24, 525 - 539.

[136] Rockafellar, R.T. (1971): Integrals which are convex functionals, II. *Pacific Journal of Mathematics,* 39, 439 - 469.

[137] Rogers, L.C.G. (2001): Duality in constrained optimal investment and consumption problems: A synthesis. *University of Cambridge.*

[138] Rudin, W. (1974): Real and complex analysis. *McGraw-Hill, Inc, New York.*

[139] Ruszczyński A. and A. Shapiro (2006): Optimisation of convex risk functions. *Mathematics of Operations Research,* 31, 433 - 452.

[140] Ryan, R.A. (2002): Introduction to tensor products of Banach spaces. *Springer-Verlag,* London.

[141] Schachermayer, W. (2000): Optimal investment in incomplete markets when wealth may become negative. *Annals of Applied Probability,* 11, 694 - 734.

[142] Schaefer, H.H. (1974): Banach lattices and positive operators. *Springer-Verlag,* Berlin-Heidelberg-New York.

[143] Schmeidler, D. (1986): Integral representation without additivity. *Proceedings of the American Mathematical Society,* 97, 255 - 261.

[144] Shreve, S.E. (2005): Stochastic calculus for finance I. The binomial asset pricing model. *Springer-Verlag.*

[145] Skala, H.J. (1999): Comonotonic additive operators and their representations. *Glasgow Mathematical Journal,* 41, 191 - 196.

[146] Skala, H.J. (2002): Generalised comonotonically additive operators: Representations by Choquet integrals. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems,* 10, 329 - 346.

[147] Song, Y. and J. Yan (2006): The representation of two types of functionals on $L_\infty(\Omega, \mathcal{F})$ and $L_\infty(\Omega, \mathcal{F}, \mathbb{P})$. *Science in China Series A: Mathematics,* 49, 1376 - 1382.

[148] Song, Y. and J. Yan (2009): Risk measures with comonotonic subadditivity or convexity and respecting stochastic orders. *Insurance: Mathematics and Economics,* 45, 459 - 465.

[149] Stoica, G. (2006): Relevant coherent measures of risk. *Journal of Mathematical Economics,* 42, 794 - 806.

[150] Stricker, C. (1990): Arbitrage et lois de martingale. *Annales de l'I.H.P., Section B,* 26, 451 - 460.

[151] Sundaresan, K. (1977): The Radon-Nikodým theorem for Lebesgue-Bochner function spaces. *Journal of Functional Analysis,* 24, 276 - 279.

[152] Talagrand, M. (1983): La structure des espaces de Banach réticulés ayant la propriété de Radon-Nikodým. *Israel Journal of Mathematics,* 44, 213 - 220.

[153] Talagrand, M. (1984): Pettis integral and measure theory. *Memoirs of the American Mathematical Society,* 307.

[154] Turett, B. and J.J. Uhl (1976): $L_p(\mu, X)$ $(1 < p < \infty)$ has the Radon-Nikodým property if X does by martingales. *Proceedings of the American Mathematical Society,* 61, 347 - 350.

[155] Von Neumann, J. (1928): Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen,* 100, 295 - 320.

[156] Von Neumann, J. and O. Morgenstern (1947): Theory of games and economic behaviour, 2nd Edition. *Princeton University Press,* Princeton.

[157] Wang, S.S. (1996): Premium calculation by transforming the layer premium density. *ASTIN Bulletin,* 26, 71 - 92.

[158] Wang, S.S., V.R. Young and H.H. Panjer (1997): Axiomatic characterisation of insurance prices. *Insurance: Mathematics and Economics,* 21, 173 - 183.

[159] Wang, S.S. (2000): A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance,* 67, 15 - 36.

[160] Wang, S.S. (2002): A universal framework for pricing financial and insurance risks. *ASTIN Bulletin,* 32, 213 - 234.

[161] Wang, S.S. (2003): Equilibrium pricing transforms: New results using Bühlmann's 1980 economic model. *ASTIN Bulletin,* 33, 57 - 73.

[162] Wang, S. (2007): Normalized exponential tilting: Pricing and measuring multivariate risks. *North American Actuarial Journal,* 11, 89 - 99.

[163] Wang, T. (1999): A class of dynamic risk measures. *Working Paper, University of British Columbia.*

[164] Yaari, M.E. (1987): A dual theory of choice under risk. *Econometrica,* 55, 95 - 115.

[165] Yan, J.A. (1980): Caractérisation d'une classe d'ensembles convexes de L_1 ou H_1. *Séminaire de Probabilités XIV, Springer Lecture Notes in Mathematics,* 784, 220 - 222.

[166] Zaanen, A.C. (1953): Linear analysis. Measure and integral, Banach and Hilbert space, linear integral equations. *Interscience Publishers,* New York.

[167] Zaanen, A.C. (1983): Riesz Spaces, Vol. II. *North-Holland Publishing Company, Amsterdam.*

[168] Zaanen, A.C. (1991): Introduction to operator theory in Riesz spaces. *Springer-Verlag Berlin Heidelburg New York.*

[169] Zalinescu, C. (2002): Convex analysis in general vector spaces. *World Scientific Publishing Company.*